Published in final edited form as:

Behav Genet. 2018 January ; 48(1): 22–33. doi:10.1007/s10519-017-9882-y.

Cross-lagged analysis of interplay between differential traits in sibling pairs: Validation and application to parenting behavior and ADHD symptomatology

Arden Moscati^{1,2}, Brad Verhulst³, Kevin McKee², Judy Silberg², and Lindon Eaves²

¹Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA

²Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University School of Medicine, Richmond, VA, USA

³Department of Psychology, Michigan State University, East Lansing, Michigan, USA

Abstract

Understanding the factors that contribute to behavioral traits is a complex task, and partitioning variance into latent genetic and environmental components is a useful beginning, but it should not also be the end. Many constructs are influenced by their contextual milieu, and accounting for background effects (such as gene-environment correlation) is necessary to avoid bias. This study introduces a method for examining the interplay between traits, in a longitudinal design using differential items in sibling pairs. The model is validated via simulation and power analysis, and we conclude with an application to paternal praise and ADHD symptoms in a twin sample. The model can help identify what type of genetic and environmental interplay may contribute to the dynamic relationship between traits using a cross-lagged panel framework. Overall, it presents a way to estimate and explicate the developmental interplay between a set of traits, free from many common sources of bias.

Introduction

Univariate exploration of behavioral traits in families provides a great deal of insight into the relative proportions of genetic and environmental contributions to variance, but without further elaboration, it is unable to partition these factors into direct and indirect pathways. Some genetic factors may play a direct causal role in the etiology of variance in a trait, such as social religiosity (1), through effects on metabolic pathways or neural development patterns, with the extent of their actions contained fully within the physiology of the individual. Other genetic factors may act proximally on a different trait, perhaps conscientiousness, which influences behavior in such a way that it tends to attract treatment or experiences that tend to predispose a person to increased levels of the first trait, social religiosity. These types of indirect sequences of causality are undoubtedly relevant to the origins of a wide variety of outcomes, but are not commonly studied. Therefore, we propose a structural model designed to bring out this kind of interplay between traits over time using longitudinal data on sibling pairs.

One relevant situation for which the model was designed is when a genetic trait makes a person more likely to be exposed to a particular salient environmental exposure. There are many forms of interaction and correlation between genetic and environmental factors that act to contribute to variation in human behavioral traits; and recognizing these kinds of interplay is necessary to avoid drawing misleading conclusions about etiology (2). This interplay can also profoundly complicate attempts to disentangle the relative roles of genetic and environmental factors in giving rise to trait variance. One of the primary forms of gene-environment interplay is known as gene-environment correlation (G-E correlation or rGE) (3).

G-E correlation is the tendency for the genetic factors that influence a phenotype to be systematically related to the environmental factors that influence the phenotype. Accordingly, an individual's genotype is related both to the phenotype of interest and to the environmental factors that affect the manifestation of phenotype. There are several mechanisms that can induce such a correlation between genetic and the environmental factors, the three most commonly discussed of which are passive, active, and evocative (or reactive) rGE (3). First, passive rGE occurs when the genotypes of an individual's relatives "create" an environment that can influence that individual's phenotypes. This is most common when parental genotypes (one allele each of which, it should be remembered, will generally be passed down to biological children) contribute to parental phenotypes that affect the rearing environments for their children. These rearing environments can potentially promote similar phenotypes in the offspring indirectly, such as depressed parents providing depressogenic home environments, in addition to direct effects on the phenotype from those same genotypes present in the children. Second, active rGE occurs when individuals seek out specific environmental experiences, indirectly motivated in part by their genotype. For example, children with ADHD having delinquent friends that then encourage delinquent behaviors. Finally, evocative rGE denotes the genetically-influenced tendency for individuals to induce particular kinds of responses from those around them, which might be the case for children with anxiety, who might cause others to act ambivalently towards them, which would further increase the anxiety (2).

Passive rGE can be directly assessed with some study designs that incorporate information from multiple biological relatives such as sibling interaction models (4), extended family models (5), or in mixed biological/adoptive family designs that assess the correlations between family environment and child phenotype in adoptive and biological families (3), such as when there are similar proportions of high individual risk environments between the two family types.

Active and evocative rGE are yet more difficult to assess. One of the difficulties with assessing active and evocative rGE is tied to the fact that in most cross-sectional designs the contribution of environmental exposure correlated with an individual's genotype is confounded with shared environmental factors statistically independent from the genotype (for correlations between shared environment and additive genetic effects) and the genetic effects themselves (for correlations between unique environment and additive genetic effects).

Longitudinal designs provide additional information to disentangle the complex relationship between the genotype and the environment. In the proposed model specifically, the examination of the pattern of changes in the phenotype over time allows the components contributing to its correlation with a certain environmental factor to be deconstructed, which can provide information about the marginal contribution of one facet of rGE to the development of the phenotype.

Importantly, there have recently been a number of other efforts targeted at the elusive variance attributable to rGE. Dolan et al. have constructed an elaboration of the ACE simplex model (6), which posits autoregressive paths between the phenotype of children and later unique environmental experiences (E) of both themselves and their sibling. They note that a number of constraints are required for identification, and that their model has low power to fit ACE data, but do find that fitting an ACE or ADE (as appropriate to the pattern of twin correlations) simplex model to data simulated by an AE model that includes rGE influences (such as their phenotype to unique environment paths) causes the spurious finding of significant C or D influences (given positive or negative transmission, respectively). Applications of similar modeling intuitions have found that mean trait increases (such as generational increases in average intelligence over time) can arise from phenotype to environment correlation (7). These models provide estimates of within-trait gene-environment interplay; successfully modeling complex phenotypes that self-reinforce over time. The goals, however, are somewhat different from those of the present model, which are focused on testing and explicating the gene-environment interplay between two specific traits in a longitudinal framework.

Beam and Turkheimer (8) have also been working on the rGE issue. They find that even small initial differences in phenotype between siblings can cause reallocation of family resources, setting up a dynamic interaction in which interaction between phenotype and environment causes later increases in total variance. They suggest that such feedback may underlie some cases of increasing heritability over time, and that unstandardized parameter estimates should be examined for evidence of increased phenotypic variance over time. Estimating this type of feedback between phenotype and environment is one of the main goals of the present study.

The present method extends the basic logic of the cross-lagged panel correlation test (9) to test for causal interplay between traits over time, including specific measured contributions to active or evocative rGE. The cross-lagged panel correlation test is a quasi-experimental test for spuriousness, comparing the strengths of causal pathways between variables over time to determine if there is evidence for a causal relationship between them.

In (positive) genotype-environment correlation, there is a tendency for genetic factors contributing to a certain phenotype in a person (a behavioral outcome or psychiatric diagnosis, for instance) also to increase the likelihood of that person experiencing environmental exposures that also tend to predispose towards the same phenotype. Thus genes may act through indirect means mediated by environmental exposure. The longitudinal cross-lagged design attempts to separate the two sides of this feedback loop – the effects of the phenotype on the (exposure to the) environmental factors and the converse

effects of environment on phenotype. While this model uses the phenotypic measure itself as a proxy for the genotypic factors influencing that phenotype, this has the advantages of plausibility (virtually any genes correlated with an environment through active or evocative mechanisms must have that association mediated by some kind of phenotypic effect) and interpretability, by allowing for the influence of the specific environmental factor chosen on the trajectory of the phenotype over time (whereas normally we don't consider environmental factors as causes of genetic effects).

This basic model is applicable to active and evocative rGE, with the difference being in the environmental variable chosen. For the purposes of this model, the difference is semantic. The model tests for active rGE when the relevant environment is at least partially under the intentional control of the individual. Alternatively, the model tests for evocative rGE when the environmental variable is the reactions of others. Indeed, in cases where the relevant environment is a social one, the distinction between active and evocative rGE is ambiguous. The decision of what term to use is certainly dependent on the context and interpretation. Note that there is not a singular parameter corresponding to magnitude of rGE in the model - rather the magnitude and relative strength of the cross-lagged paths (β_{CP} and β_{PC} ; see Figure 1). To the extent that the two traits chosen correspond with a (genetically-influenced) phenotype and an environment, the cross-paths together represent the gene-environment interplay, or rGE, between them.

Another key feature of the model is the use of differential variables (differences between a pair of siblings), which removes common environmental effects. This means that most forms of passive rGE, which are confounded with shared environment, are controlled in the model, but also that passive rGE is not estimable. Differential items also may have better psychometric properties than absolute assessments, such as increased inter-rater reliability and consistency over time (10).

When exploring the change in phenotypes over time it is prudent to account for the presence of factors that exert a constant influence on all measurement occasions (i.e. the stable level of each trait) that may be independent from the accumulating effects of the interaction between the variables. In the current model, these effects are accounted for by the inclusion of two latent factors (one for each variable) with equal loadings on each measurement occasion.

Quite inadvertently, the specification of this model (detailed in the Model Structure section) aligns substantially with other refinements to cross-lagged panel designs in the literature. A number of strategies have been proposed in order to address the limitations of the basic cross-lagged panel model with autoregression, which include implausible modeling of theoretical processes and difficulty in the interpretation of parameter estimates (11) as well as potential errors in the estimates themselves. Past simulations have also suggested possible errors in determining whether notable cross-lagged effects are present, inaccurate comparison of the size of cross-lagged effects, and mistaken signs in cross-lagged relationships (12). The present model is comparable to both the random intercepts cross-lagged panel model, or RI-CLPM (12), and the autoregressive latent trajectory model with structured residual model or ART-SR (11, 13). The main difference that separates the current

model (and these two similar specifications) from the basic cross-lagged panel design is the inclusion of parameters corresponding to differences both within and between families (or persons, in other uses), which, it should be noted, is not a particularly recent innovation, having been used for instance by Phillips and Fulker in a model with both time-specific and time-constant factors (14).

We begin with a brief explanation of the primary structure of the model, then to demonstrate of the soundness of the model, we present a series of simulations conducted with combinations of continuous and ordinal variables as the traits under study. In order to properly gauge the requirements for effective use of the model we then present power analyses for various degrees of correlation between traits and sample sizes. We next provide a case example of the fitting of the model in a cohort sequential sample of child and adolescent twins, using ADHD symptoms and paternal praise as the variables of interest. Finally, we conclude with discussion of the potential use and interpretation of the model.

Model Structure

The model's characteristic features are evident (as with all structural equation models) in its diagrammatic representation: Figure 1, which uses the example traits of 'Parenting' and 'Child Behavior'. Each manifest variable in the model represents a difference between Twin 1 (T_1) and Twin 2 (T_2) of observed raw score z , measuring latent trait i :

$$x_{i,t} = z_{i,T_2} - z_{i,T_1}$$

For either twin, each pair of manifest variables at each time follows the equations:

$$x_{1,t} = \beta_{11}x_{1,t-1} + \beta_{21}x_{2,t-1} + \zeta_1 + \delta_{1,t}$$

$$x_{2,t} = \beta_{22}x_{2,t-1} + \beta_{12}x_{1,t-1} + \zeta_2 + \delta_{2,t}$$

Where $x_{1,t}$ and $x_{2,t}$ represent a pair of variables, in this case, difference scores, at time t in the series. Autoregressive and cross-lagged regression coefficients β_{ij} describe the regression of $x_{j,t}$ onto $x_{i,t-1}$. In the equations above, where $i = j$, β describes the degree of autoregression in a variable over time, and where $i \neq j$, β describes the cross-lagged regression between variables. These represent the influences of a trait at one time point (e.g. a parenting measure of twins age 8–9) on the both the same trait and the other trait (e.g. a child behavior) at the next time point (ages 10–11). All paths with the same independent and dependent variables are constrained to equality between multiple instances, to assess the overall trend of the effect of one occasion on the next: for example, the effect of ADHD symptoms at time 1 on the ADHD symptoms at time 2 would be estimated with the same value as the effect of the symptoms at time 2 on symptoms at time 3, and so on. This equality constraint could perhaps be relaxed to investigate moderation of autoregressive effects by time, but caution must be used to avoid underidentification, if the number of parameters increases significantly.

ζ_i represents the latent factors as trait variance which remains stable across all timepoints, while conversely $\delta_{i,t}$ represents residual variances of the manifest variables at each time, not due to the latent factor or prior time points. These residual variances are estimated with individual values for each time point, to allow for potential increases in residual variance over time. These estimates include measurement error as well. Testing the fit of a constraint holding the variance of each time point in a trait to equality can provide insight into whether the overall average variance is changing over time, for instance if greater differences in a trait are developing in older ages. The latent variables, ζ_i represent any persistent factors influencing the relevant trait equally throughout all measurement occasions (loadings are constrained to 1 from each latent factor to its relevant manifest variables); such as genetic factors active at all ages. The extent to which these factors vary between families is accounted for by the variance of the latent factors (labeled with a ζ for the two factors, estimated separately) and any overlap of latent factors between the two traits is represented by a covariance between them, designated ϕ .

The expected covariance matrix was calculated using RAM notation (15):

$$\Sigma(\Theta) = F(I - A)^{-1} \Omega (I - A)^{-1T} F^T$$

$\Sigma(\Theta)$ represents the expected covariance as a function of the vector of parameters, Θ . The A matrix contains all of the free and fixed path coefficients directionally relating both latent and observed variables. The variances and non-directional covariances of both latent and observed variables are defined in the Ω matrix, including, for instance, the covariance between the latent traits. The F , or filter matrix distinguishes observed from latent variables.

All of these parameters are initially estimated separately for MZ and DZ twin pair groups – this allows for testing of significant differences in each parameter between the two groups. As mentioned above, the traits chosen for use in this model should be differential in nature; that is, signed values of differences between siblings (twins, if possible). These may be asked in a directly differential fashion: “Which sibling displays more trait X?”, or they may be computed from individual ratings of the two siblings (see Figure 2). Differential items, in addition to their psychometric advantages (10), have a couple of interesting properties. A differential rating of a twin pair removes any shared environmental variation, and will (assuming twin order is randomly assigned) naturally have a mean (and generally mode) of 0, within error range (or the middle, balanced response if an ordinal item). Difference scores therefore also remove the sort of passive rGE characterized by correlation between additive genetic effects and the shared environment, since correlation between A and C acts like C (16), and C (or shared environmental factors) are removed in difference scores. In addition, there are fewer variables in the model, as what would be separate twin 1 and twin 2 values are instead collapsed into a single value, increasing ease of computation somewhat. There is a different, and somewhat more limited set of questions one can pose using difference scores, it is worth noting; inquiries into individual trait means are not possible. The effects of the other biometrical variance components based on relative resemblance of MZ and DZ twins on a differential item are also somewhat different (see Table 1). This means, among other implications, that the true variances of DZ differential items should exceed MZ’s,

because when the siblings have the same level of a trait, they should have the mean value (0), and MZ's are more similar than DZ's due to genetic factors at the least, so there will be less expected dispersion around the mean for MZ pair differential values. Also, twice the squared mean absolute value of the difference ratings is an estimate of the intra-pair variance within twin pairs. Thus, under the simple assumptions of the classical twin method, twice the difference between the squared mean absolute values of the difference ratings of MZ and DZ twins is proportional to the heritability of the underlying trait. If the model is effective, the magnitude and relative estimates of MZ and DZ autoregressive paths should indicate the extent to which the two traits under consideration influence each other, and also to what degree genetic factors are responsible for the interplay between the traits, as genetic factors will only generally affect the sequential differential rating occasions for DZ pairs. This may be a direct way to estimate the strength of active or evocative gene-environment correlation (between the genes of a genetic trait and a given environmental exposure characterized by the other trait).

Simulations

In order to test the performance of the model and evaluate its properties, we conducted a number of simulations using data randomly generated around certain input parameters. Firstly, initial simulations directly generated the differential trait values were run, to determine the performance of the model with different types of traits, both continuous and ordinal. Secondly, individual data for simulated twin 1 and twin 2 traits were generated using ACE trait and innovation covariance structures, and differential traits were calculated subsequently for entry into the model.

Simulations were conducted in R, using data generated with `mvrnorm()` from the MASS package and all model analyses were conducted with OpenMx.

Simulation 1

Our first simulation used two continuous variables, each conceived as representing the difference between individual values of a sibling pair on some continuous trait. For simplicity, these are generated to have approximately the same variance (with means fixed to 0, as expected for a differential rating of this type*), and represent two groups of siblings: monozygotic (MZ) and dizygotic (DZ) twin pairs. Each group was simulated with 5000 pairs, and 1000 simulations were run in each condition. Regression paths, and residual variances were equated across time points, and regression paths were also equated across zygosity groups, leaving the total number of estimated parameters at 14.

* Referring to the values approximating X in $X = Y - Z$, where assignment of which sibling provides the 'Y' and 'Z' values is random, or alternatively a directly reported answer to 'which sibling does X more?', properly zero-centered.

We also performed simulations using a combination of ordinal and continuous variables. This would represent a leveled characteristic, such as categories of socioeconomic status, in relationship to a continuous trait, which can still helpfully be conceived as weight. This changes relatively little, so long as the thresholds are constrained to their initial values. If

allowed to vary freely, the ‘scale’ of the estimates related to the ordinal variable is uncertain and hence the resulting values may vary collectively into a linear transformation of the true values. For the purposes of this simulation we created three levels for the ordinal variable (perhaps ‘low Socioeconomic status’, ‘moderate SES’, and ‘high SES’), and again 5000 each of MZ and DZ twin pairs were simulated.

Parameter estimates for all paths and variances were recovered with precision; very consistently for continuous-continuous data, and less so when more of the data is ordinal, see Supplemental Figures 1, 2 and 3, which use the parameter naming conventions introduced in Figure 1. This may be viewed as evidence of the identification of the model, structurally speaking, and encouraged further evaluation of its capabilities.

Simulation 2

The relationship between the partitioned variance components of the classical twin model and the parameters of the present model was also of interest, so a series of simulations were conducted with different combinations of genetic and environmental interplay. The simulation generated two (generic) continuous traits of modest genetic correlation ($r_G = 0.3$) along the same scale (mean 0, variance 1), each with 40% of their variance due to genetic factors (A), 40% due to unique environmental factors (E), and 20% due to shared environmental factors (C).

Each condition was simulated with 5000 MZ pairs and 5000 DZ pairs, and 500 replications. 5 occasions of measurement were generated for each individual in both traits, composed of a combination of time-invariant (latent) trait variance and time-specific innovations. Separate normally distributed A, C and E components were generated for each simulated individual, which were then combined into an individual trait value for each time point. See Supplemental Tables 1–4 for covariance matrices for the latent traits, and innovations.

The alterations of interest in the simulated circumstances were made in the covariance matrix between time-adjacent innovations; that is, the relationship between the innovations of trait 1 and trait 2 at time $t-1$ and the innovations of trait 1 and trait 2 at time t . All traits were assumed to have a small autoregressive impact - the value of an individual at time $t-1$ had a small tendency to compound at time t (a high value at time $t-1$ would encourage higher values at time t and low values at time $t-1$ would tend to be followed by lower values at time t). The differences between conditions were expressed in cross-trait effects between adjacent innovations. Four conditions were explored (see Supplemental Tables 5–12 for the beta matrices which instantiate these conditions):

1. Asymmetric additive genetic interplay (AA), where the genetic variance of trait 1 innovations at time $t-1$ affected the genetic variance of trait 2 innovations at time t . For example, genetic influences on body weight beginning to increase the statistical effect of genetic factors on blood pressure over time.
2. Symmetric gene-shared environmental interplay (AC), where each trait's genetic variance has an impact on the shared environmental variance of the other trait at subsequent time points. This could represent genetic influences on scholastic

ability and income increasing familial SES over time, which itself may have compounding environmental benefits on scholastic ability and income.

3. Symmetric gene-unique environmental interplay (AE), where each trait's genetic variance has an impact on the unique environmental variance of the other trait at subsequent time points. For instance, genetic influences on anxiety and depression evoking aversive treatment from friends and family, which then exacerbates mood and anxiety symptoms.
4. Asymmetric unique environmental interplay (EE), so that the unique environmental variance of trait 1 at time $t-1$ predicts the environmental variance of trait 2 at time t . A sample scenario might be an individual happening to befriend a peer who engages in deviant behavior, which then may increase access to illicit substances.

These conditions were used to generate the innovations across time points as a function of the given type of genetic, environmental or gene-environment interplay. The individual twin values on each trait at each time point were then used to compute a difference score: twin 2's value minus twin 1's value. These difference scores were then used as input data for the cross-lagged model described earlier. The model's parameter estimates have a systematic relationship with the type of interplay in the generated data, as shown in Figure 3.

Notable in Figure 3 are the cross-lagged parameters in each simulation condition that are non-zero (significant). For the asymmetric 'AA' condition, only the DZ 1→2 path is significant, because the difference scores only include additive genetic variance in the DZ group (see Table 1 for differences in variance composition between individual ratings and differential ratings), and only the trait 1 to trait 2 path was included in the simulation parameters. For the symmetric 'AC' condition, all cross-lagged paths in both MZ and DZ are approximately zero, since shared environmental variance is entirely excluded from difference scores. For the symmetric 'AE' condition, only the DZ paths are significant, because the difference scores again only include additive genetic variance in the DZ group. For the asymmetric 'EE' condition, the trait 1 to trait 2 paths are significant in both MZ and DZ groups, and of greater magnitude in the MZ group. This is perhaps a surprising finding, because the simulation parameters were equal between MZ and DZ trait 1 to trait 2 cross-lagged influence. The discrepancy is due to the overall greater variance in the DZ group than the MZ, because in the DZ group, half of trait additive genetic variance is captured in the difference scores, which means that there is greater variation. Thus, the MZ trait 1 to trait 2 regression coefficient is larger because E represents a higher proportion of the overall variance of each time point in the MZ group, so the E factors explain a larger share of the data.

These comparisons between cross-lagged parameters can provide a surprising amount of information about the interplay between the two traits, and what types of origins it may have. To further validate the statistical sensitivity of the model, we also conducted a power analysis.

Power Analysis

In order for the results of a model to be properly interpreted, it is helpful to have some context for its efficiency in detecting significant parameters. As this model is primarily focused at the reciprocal interplay between two longitudinally measured variables, parameters that constituted the links between the two traits chosen were varied in effect from small (0.1) to medium (0.3) to large (0.5) effects, according to Cohen's conventions for correlation coefficients (17), while holding the other parameters to fixed values detailed in Figure 4 (with one genetically-influenced trait – Child Behavior; and one largely-environmentally-determined one – Parenting; both traits were simulated as continuous for this power analysis). Note that these regression coefficients are not correlations - these conventions were used for the simulation conditions only because they are broadly familiar in the literature. For the detection of the cross-lagged autoregression parameters, estimates were constrained to equality between the two groups of twins in this particular analysis, for simplicity. To estimate the power at varying effect sizes of these two paths, we calculated power curves (Supplemental Figures 4–6) based on the comparative likelihood of simulated models (for which 5000 pairs of MZ twins, and 5000 pairs of DZ twins were generated) where the parameter of interest is freely estimated and constrained to 0. We also examined power for different magnitudes of the covariance between the two latent factors underlying the central constructs X and Y (see Supplemental Figure 6). Power was good (at an alpha level of 0.05) even at small magnitudes of each of the parameters considered individually, albeit less strong for autoregression of the non-genetic trait X on the genetically influenced trait Y.

Demonstration Methods

Dataset

The Virginia Twin Sample of Adolescent Behavioral Development (18) is a longitudinal study that ascertained twins between the ages of 8 and 17 by statewide recruitment from Virginia schools; families were followed across three additional waves if the twins were younger than 18 and still enrolled in high school. 1894 families were initially identified, and of those 74.5% (n=1412) elected to participate in the first wave. Participation remained well above 70% of the qualifying for follow-up waves as well. The twins were evaluated used an elaborate psychiatric battery of assessments, and their parents were also asked to complete interviews regarding each twin. This results in data from multiple raters for each individual child. Ages of the twins were stratified into 5 groups based on age to the nearest year: 8–9, 10–11, 12–13, 14–15, or older than 15. Dizygotic pairs of unlike sex were also excluded from this analysis, which is an important limitation, but does not preclude genuine investigation of certain sex-related influences – using differential scores between two twins of the same sex will wipe out any mean effects based on sex, and leave only sex effects on the comparison between the twins. Sex differences will be seen only if the differential score within a male like-sex pair is dissimilar from that within a female like-sex pair in some fashion.

Instrument

The data that is used in the present analysis came from one section of a supervised self-report questionnaire administered during the home interview. A modified version of the Sibling Inventory of Differential Experiences (SIDE) (19) was used to evaluate the extent of differential treatment of twins from the perspective of each twin and their parents. Data from the 437 families who completed at least one questionnaire at two or more time points was included in the present analysis.

Items

With data available for a number of items related to various aspects of sibling treatment, the differential items relating to parental praise were chosen to take advantage of the strengths of differential items while indexing parenting trends thought to be significant to overall developmental outcome. Separate items were available for treatment by the mother and father, and since treatment by both parents is of interest to psychological development, both were examined. Separate ratings by twin, father and mother were also present in the dataset, but the rating by the parent whose behavior was indexed by the item was chosen, to reduce the systematic bias of having an external rater. Since fathers are relatively understudied in the parenting literature, the example parenting behavior of paternal praise was chosen for presentation. Items were scored as 5-point semantic differentials, and recoded to have a central point at 0: with -2 indicating “twin 1 more”, 0 indicating “neither twin more”, and 2 indicating “twin 2 more”, with intermediate responses at -1 and 1. Due to the small number of responses at levels -2 and 2, they were collapsed with the adjacent responses (-1 and 1 respectively) to produce 3 possible values, corresponding with “twin 1 more”, “neither twin more”, and “twin 2 more”. ADHD symptomatology was assessed for each twin as a sum score of 14 symptoms, and then a difference score calculated at each time point: twin 2 symptom count - twin 1 symptom count. It is important to note, therefore, that the two variables have different ranges. See Supplemental Tables 13 and 14 for sample sizes and item information.

Analysis

Data was analyzed using cross-lagged (discrete-time autoregressive with lag 1) modeling. In this framework, the score of a variable at each time point is caused by the time point immediately preceding it, as well as the other variable at the preceding time point and acts itself on the two variables at the time point directly subsequent. All time points of a given trait are also subject to the latent factors acting equally. OpenMx path specification was used for optimization of the model fitting procedure. Each autoregressive path was specified as unidirectional (consistent with the temporality of the time points), and the estimated correlation between latent factors as bidirectional.

The analysis was performed initially allowing free variation of path coefficients and most variances (MZ variance for the latent factors in this model was constrained to 0, given the primary theoretical source of variance in these latent factors that act equally over all time points is genetic difference between cotwins in relevant gene sets - virtually absent in MZ twins) between the subgroups of monozygotic twin pairs and same-sex dizygotic twin pairs. Expected means were not estimated separately for males and females, though this could

easily be done in future uses of the model. To attempt to achieve increased parsimony over this quite elaborate model, path coefficients, variances, and means were constrained between various groups, combinations of variables and time points. Each constrained iteration was compared with the saturated model using the likelihood ratio test, with an alpha level of 0.05. The -2 log likelihood of the saturated model was subtracted from that of the constrained model, the distribution of which is theorized to follow a chi-square with degrees of freedom equal to the difference in number of parameters between the two specifications. If this test is not significant for a model iteration, the null hypothesis that the estimated correlations are the same as the observed correlations could not be rejected.

Results

In fitting the cross-lagged models to the data for the four items in relation to ADHD symptomatology, constraints were added to the saturated model iteratively, when a constraint caused the likelihood ratio test to reach significance, that constraint was dropped and other specifications were tried until the most parsimonious model that did not reach LRT significance was found. The parameter estimates for the most parsimonious model with acceptable fit are presented in Table 6, and the goodness-of-fit tests for a subset of examined models is available in Supplemental Table 15. Most parameters could be dropped without loss of fit, excepting for the manifest (residual) variance of paternal praise, and the manifest variance of ADHD symptoms (which additionally varied by twin type), the autoregressive within trait path of ADHD symptoms, and the effect of paternal praise on ADHD symptoms. The confidence intervals of the praise to ADHD path did include zero (reflecting the small sample size, perhaps), but not including the path worsened the fit considerably.

Discussion

One of the most important functions of structural equation models is to partition variance in traits into particular sources of interest and residual error. In some models, family relatedness is deployed in order to determine what proportion is attributable to genetic influences, and shared upbringing - the variance of interest in the present model, however, constitutes the interplay, which should contain most forms of causal relationship between the two traits. The unique features of the model allow for the separation of factors that would induce correlation in the traits due to common causes or background context. Differential ratings between siblings remove all effects of shared upbringing and related gene-environment correlation. The latent factors account for any intercept differences in the traits between the sibling pairs, as well as influences that act on one or both traits equally over all measured time points.

In the test case, many of the pathways were not significant, due in part no doubt to the small sample size of twins that had responded to the items at two or more time points. The latent factors had negligible effect, and the development of differential paternal praise and ADHD symptomatology over time seemed to be due to the temporal stability of ADHD (a substantial within-trait autoregressive pathway) as well as a small path (with confidence intervals overlapping zero, but required for model fit) constituting the effect of paternal praise on ADHD symptomatology – the twin that was praised more at a given age tended to

be have more ADHD symptomatology than his or her cotwin at the next age. This is a somewhat counter-intuitive finding, but positive attention such as parental praise is, in addition to its role in healthy parenting behavior, also a classic feature of unhealthy parenting behavior that is warm and doting, but lacking discipline, known as permissive parenting. This is theorized to be the explanation for praise leading to increased ADHD symptoms. Parents of children diagnosed with ADHD (and other childhood psychopathology) are often more permissive than parents of children without ADHD (20), especially fathers (21). In the final model, there is some evidence against the relationship between these traits being due to genetic factors – the latent factors were not significant, and autoregressive pathways could be constrained to equality between twin types – indicating no additional effect of additive or dominance genetic influence on making DZ twins more different than MZ twins with respect to the interplay between these two traits, suggesting the primacy of unique environmental factors. Notably, however, the residual variance for ADHD symptomatology was substantial (and substantially different between MZ and DZ twin pairs), though not in the expected direction. The MZ pairs had greater variance, despite their genetic similarity, perhaps pointing to potential tendencies for MZ twin pairs to polarize in terms of ADHD symptomatology (one twins hyperactivity or inattention tends to mute the expression or reporting of similar behavior in the other), due perhaps to some combination of identity formation tendencies (“how am I different than my identical twin?”) or differential MZ parenting strategies (“which one is the obedient one?”)

This model is related to, but distinct from, other developmental models attempting to account for the interplay between genetic and environmental factors in the progression of a phenotype over time. While causal relationships between traits at the phenotypic level are of central interest in this model, it does not explicitly model the proportions of variance in these relationships attributable to latent genetic and environmental factors. The presence of unique environmental effects (if the MZ estimate is significantly different from zero), and genetic effects (if the MZ and DZ correlations are significantly different from each other) in the various paths estimated in the model can be detected through model constraints, however. Due to the differential nature of the traits under study, variance due to shared environmental factors, and any variance that contributes to its estimation, such as rGE between common environmental and additive genetic factors (16) will be completely absent from the estimation of the model, which is an important fact to consider when interpreting its results.

Thus, a number of potential states of affairs can be suggested by the relative magnitude of the parameters of the model as applied to a given sample of longitudinal differential ratings of twin pairs for two traits. As suggested by past work on cross-lagged panel designs, asymmetric causal relationships between traits are reflected in potential asymmetry of the trait 1 to 2 versus trait 2 to 1 paths. If all cross-lagged regression parameters are non-significant, this suggests that no dynamic interplay is at work between them (or that any interplay involves shared environmental factors, see below); they may be uncorrelated or any phenotypic correlation may be due to constant factors, such as etiological overlap in genetic or environmental influences. If the cross-lagged paths are significant, comparing MZ paths to DZ can be additionally informative. Due to the nature of difference scores, if a trait has any genetic etiology, it is expected that the variance of DZ difference scores should be

greater than that of MZ difference scores, assuming all assumptions (such as the equal environments assumption) hold.

Based on the results of Simulation 2, it seems as though the relative magnitude of MZ and DZ cross-lagged paths can shed light about which sources of variance are involved in the interplay between traits, given two traits with the same scale and each with some level of additive genetic variance. If only shared environmental factors are involved, none of the cross-paths will be significant, due to the use of difference scores, which effectively control for shared environmental effects. If genetic factors are active in all of the dynamic interplay between traits, either in relationship to other genetic factors or unique environmental factors, the DZ cross-lagged paths that involve additive genetic factors will be significant, but the MZ paths will not. If the unique environmental factors of one trait are active in relationship with the unique environmental factors of the other trait, both MZ and DZ cross-lagged pathways should be significant, with the magnitude of the MZ paths perhaps exceeding the DZ, assuming the MZ residual variances are smaller than the DZ residual variances (as expected in traits with non-zero heritability). If the MZ and DZ cross-lagged paths can be fixed to equality, there may be a combination of genetic and (unique) environmental interplay at work (again given DZ variances exceed MZ variances). Similar tests could perhaps shed light on within-trait gene-environment interplay over time, though validation of such uses is beyond the scope of the present paper.

While we now know that the vast majority of human behavioral and psychological traits are influenced by both genetic and environmental factors to varying degrees, the relationships between such traits are often harder to characterize. Since correlation can reflect statistical relationships due to unmeasured variables in addition to the presence of direct causal links, the characterization of how many of these psychological constructs interact is still in its infancy, especially when the genetic influences on many traits, such as child behavior, also have indirect causal impacts on circumstances, such as parenting patterns, which may later have environmental influences on those same traits – leading to genetic variation in the environment that only increases over developmental time. The presence of relevant unmeasured or unmodeled variables may of course bias the results of our model, but the primary process of cross-lagged causation will reflect, with respect to the measured traits that are included in the model, their overall dynamic interaction, free from many confounding influences such as correlated shared environmental variance, intercept differences between cotwins in either trait, and passive rGE. This can, as detailed above, suggest what kind of dynamic gene-environment interplay (likely of the active or evocative variety) is at work in the relationship between them.

Biometrical genetics has made great strides in partitioning the variance sources determining all manner of human outcomes and traits, but the determination of what particular processes and mechanisms mediate the overall statistical impact of latent sources of variance on the constructs we measure still will require much more research. Our model offers a way to examine and characterize the developmental relationships between phenotypes that is free of many types of confounding influences.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Vance T, Maes HH, Kendler KS. Genetic and environmental influences on multiple dimensions of religiosity: a twin study. *The Journal of nervous and mental disease*. 2010; 198(10):755–61. Epub 2010/10/06. DOI: 10.1097/NMD.0b013e3181f4a07c [PubMed: 20921867]
2. Rutter M, Silberg J. Gene-environment interplay in relation to emotional and behavioral disturbance. *Annual review of psychology*. 2002; 53:463–90. Epub 2001/12/26. DOI: 10.1146/annurev.psych.53.100901.135223
3. Plomin R, DeFries JC, Loehlin JC. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological bulletin*. 1977; 84(2):309–22. Epub 1977/03/01. [PubMed: 557211]
4. Carey G. Sibling imitation and contrast effects. *Behavior genetics*. 1986; 16(3):319–41. Epub 1986/05/01. [PubMed: 3753368]
5. Keller MC, Medland SE, Duncan LE, Hatemi PK, Neale MC, Maes HH, Eaves LJ. Modeling extended twin family data I: description of the Cascade model. *Twin research and human genetics : the official journal of the International Society for Twin Studies*. 2009; 12(1):8–18. Epub 2009/02/13. DOI: 10.1375/twin.12.1.8 [PubMed: 19210175]
6. Dolan CV, de Kort JM, van Beijsterveldt TC, Bartels M, Boomsma DI. GE covariance through phenotype to environment transmission: an assessment in longitudinal twin data and application to childhood anxiety. *Behavior genetics*. 2014; 44(3):240–53. Epub 2014/05/03. DOI: 10.1007/s10519-014-9659-5 [PubMed: 24789102]
7. de Kort MJ, Dolan VC, Kan K-J, van Beijsterveldt EC, Bartels M, Boomsma ID. Can GE-Covariance Originating in Phenotype to Environment Transmission Account for the Flynn Effect? *Journal of Intelligence*. 2014; 2(3)doi: 10.3390/jintelligence2030082
8. Beam CR, Turkheimer E. Phenotype-environment correlations in longitudinal twin models. *Development and psychopathology*. 2013; 25(1):7–16. Epub 2013/02/13. DOI: 10.1017/S0954579412000867 [PubMed: 23398749]
9. Kenny DA. Cross-Lagged Panel Correlation: A Test for Spuriousness. *Psychological bulletin*. 1975; 82(6):887–903.
10. Carbonneau R, Eaves LJ, Silberg JL, Simonoff E, Rutter M. Assessment of the within-family environment in twins: absolute versus differential ratings, and relationship with conduct problems. *Journal of Child Psychology and Psychiatry*. 2002; 43(8):1064–74. [PubMed: 12455927]
11. Berry D, Willoughby MT. On the Practical Interpretability of Cross-Lagged Panel Models: Rethinking a Developmental Workhorse. *Child Development*. 2016; 88(4):1186–206. DOI: 10.1111/cdev.12660 [PubMed: 27878996]
12. Hamaker EL, Kuiper RM, Grasman RP. A critique of the cross-lagged panel model. *Psychol Methods*. 2015; 20(1):102–16. Epub 2015/03/31. DOI: 10.1037/a0038889 [PubMed: 25822208]
13. Curran PJ, Howard AL, Bainter SA, Lane ST, McGinley JS. The separation of between-person and within-person components of individual change over time: a latent curve model with structured residuals. *J Consult Clin Psychol*. 2014; 82(5):879–94. Epub 2013/12/25. DOI: 10.1037/a0035297 [PubMed: 24364798]
14. Phillips K, Fulker DW. Quantitative genetic analysis of longitudinal trends in adoption designs with application to IQ in the Colorado Adoption Project. *Behavior Genetics*. 1989; 19(5):621–58. DOI: 10.1007/BF01066028 [PubMed: 2803186]
15. McArdle, JJ., Nesselroade, JR. *Longitudinal Data Analysis Using Structural Equation Models*. American Psychological Association; 2014.
16. Purcell S. Variance components models for gene-environment interaction in twin analysis. *Twin Res*. 2002; 5(6):554–71. DOI: 10.1375/136905202762342026 [PubMed: 12573187]
17. Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.

18. Hewitt JK, Silberg JL, Rutter M, Simonoff E, Meyer JM, Maes H, Pickles A, Neale MC, Loeber R, Erickson MT, Kendler KS, Heath AC, Truett KR, Reynolds CA, Eaves LJ. Genetics and Developmental Psychopathology: 1. Phenotypic Assessment in the Virginia Twin Study of Adolescent Behavioral Development. *J Child Psychol Psychiatr.* 1997; 38(8):943–63. [PubMed: 9413794]
19. Daniels D, Plomin R. Differential Experience of Siblings in the Same Family. *Developmental Psychology.* 1985; 21(5):747–60.
20. Hutchison L, Feder M, Abar B, Winsler A. Relations between Parenting Stress, Parenting Style, and Child Executive Functioning for Children with ADHD or Autism. *Journal of Child and Family Studies.* 2016; 25(12):3644–56. DOI: 10.1007/s10826-016-0518-2
21. Stormont-Spurgin M, Zentall SS. Contributing factors in the manifestation of aggression in preschoolers with hyperactivity. *J Child Psychol Psychiatry.* 1995; 36(3):491–509. [PubMed: 7782411]

Figure 1.
Overall structure of the cross-lagged model

Figure 2.
Calculation of the difference scores

Figure 3.

Simulation 2 cross-lagged regression coefficient estimates for all simulation conditions. MZ refers to the monozygotic twin group, DZ refers to the dizygotic twin group. 1-->2 and 2-->1 refer to the regression coefficient between trait 1 at time t-1 and trait 2 at time t, and trait 2 at time t-1 and trait 1 at time t, respectively. 'AA', with green lines marking 0, refers to a simulation condition in which trait 1 A factors at time t-1 influence trait 2 A factors at time t. 'AC', with dark blue lines marking 0, refers to a simulation condition in which trait 1 and 2 A factors at time t-1 influence trait 2 and 1 C factors at time t, respectively. 'AE', with light blue lines marking 0, refers to a simulation condition in which trait 1 and 2 A factors at time t-1 influence trait 2 and 1 E factors at time t, respectively. 'EE', with purple lines marking 0, refers to a simulation condition in which trait 1 E factors at time t-1 influence trait 2 E factors at time t.

Figure 4.
Parameter values in the power analysis

Table 1

Relative contributions of specific sources of variance to individual and differential ratings

Relative Contributions	Additive Genetic Factors	Common Environment Factors	Dominance Genetic Factors	Unique Environmental Factors
Individual Ratings				
rMZ	1	1	1	0
rDZ	0.5	1	0.25	0
Differential Ratings				
MZ	0	0	0	1
DZ	0.5	0	0.75	1

Table 2

Parameter estimates for demonstration data best-fitting model

Parameter	Symbol	Estimate	95% Confidence Interval
Paternal Praise autoregression	β_{PP}	0 *	0 *
Paternal Praise to ADHD	β_{PC}	0.27	-0.04 – 0.57
ADHD to Paternal Praise	β_{CP}	0 *	0 *
ADHD autoregression	β_{CC}	0.64	0.56 – 0.72
Paternal Praise residual variance (MZ & DZ)	δ_P	0.62	0.55 – 0.69
ADHD residual variance MZ	δ_{CMZ}	3.74	3.17 – 4.44
ADHD residual variance DZ	δ_{CDZ}	2.07	1.77 – 2.43
Paternal Praise latent variance (MZ & DZ)	ζ_P	0 *	0 *
Paternal Praise and ADHD latent covariance (MZ & DZ)	ψ	0 *	0 *
ADHD latent variance (MZ & DZ)	ζ_C	0 *	0 *

* = parameter was non-significant, and was fixed to zero.