J. Child Psychol. Psychiat. Vol. 41, No. 5, pp. 667–674, 2000
Cambridge University Press

0021–9630/00 $15.00 + 0.00

Extracting Meaning from Comorbidity: Genetic Analyses That Make Sense

Emily Simonoff

Guy's, King's, and St. Thomas' Medical School, London, U.K.

As behavioral genetic strategies have become part of the arsenal of research in developmental psychopathology, a wide variety of genetic analyses are being applied to child psychiatric data. Multivariate genetic techniques have been used to explore comorbidity among traits or disorders and the main analysis undertaken has been to examine whether comorbidity is due to shared genetic and/or environmental factors. However, this model ignores other possible causes of comorbidity, which are reviewed. In particular, genetic analyses of comorbidity have only infrequently considered the model of phenotypic causality (one disorder directly influencing another), which provides an important alternative with potentially different implications for intervention strategies. Data from a recently published article by Wamboldt, Schmitz, and Mrazek (1998) are used to illustrate the potential difficulties of distinguishing between models of shared genetic/environmental risk and phenotypic causality. Given that the sample sizes required to distinguish between these models are often large, and frequently greater than those of the datasets available, it is argued that researchers should select the models that they test based on other lines of evidence that these models are plausible. Where convincing evidence does not exist, researchers should explore alternative models and determine their power to discriminate between these models.

Keywords: Behavior problems, behavioral genetics, comorbidity, epidemiology, methodology, twins.

Abbreviations: AIC: Akaike Information Criterion; CBCL: Child Behavior Checklist.

In the last two decades, genetic research has become a major part of scientific research. One of the effects of the expansion of genetic research generally has been an enormous increase in behavior genetic research with respect to psychiatric disorders and behavioral problems. Researchers with primary interests in child psychopathology have seen that genetic strategies can be powerful tools in dissecting issues of interest. Such questions include whether genetic variability explains sex differences in rates of disorders (Eaves et al., 1997), whether extremes of traits are distinct from the continuum (Rende et al., 1993; Slutske et al., 1997), whether there are reciprocal and genetic links between parental treatment and child behavior (O'Connor, Hetherington, Reiss, & Plomin, 1995) and whether genetic factors influence the relationship between "environmental" risk factors, such as life events, and psychopathology (Thapar & McGuffin, 1996). One of the issues that has been of great interest in the last few years has been the nature of the links between different measures, and this has been tackled using multivariate genetic analyses. These analyses address the extent to which latent genetic and environmental factors are shared across measures. These measures include symptoms of the same disorder, such as the different types of conduct disorder symptoms (Simonoff, Pickles, Meyer, Silberg, & Maes, 1998), the inattention and impulsivity dimensions of attention deficit hyperactivity disorder (Huziak et al., 1998; Nadder, Silberg, Eaves, Maes, & Meyer, 1998), and different components of intelligence (Cardon & Fulker, 1994; Petrill & Thompson, 1993). Similar strategies have also been used to examine whether the reason traits or closely related disorders are frequently comorbid is because of shared genetic and/or environmental influences. This has been applied to a variety of psychopathological traits such as symptoms of anxiety and depression (Kendler, Heath, Martin, & Eaves, 1987), conduct disorder and hyperactivity (Silberg et al., 1996), and internalizing and externalizing behavior (Gjone & Stevenson, 1997).

Comorbidity in child psychiatry is common and therefore an important issue to address using a range of strategies. There are a variety of reasons why comorbidity may occur, and these have been discussed in detail elsewhere (Caron & Rutter, 1991; Neale & Kendler, 1995). They are therefore only alluded to briefly here. Although the term "disorder" is used here for simplicity, the concept need not be considered either pathological or categorical.

(1) *Chance*. The frequency with which the two disorders occur together is not significantly greater than that expected by chance. Comorbidity occurring at greater than chance levels is demonstrated, in systematically assessed populations, by the frequency or prevalence of the comorbid state being greater than the product of the frequencies of the single disorders. For continuous or semi-

Requests for reprints to : Emily Simonoff, Department of Child and Adolescent Psychiatry and Psychology, Guy's, King's and St. Thomas' Medical School, Guy's Hospital, Munro Centre, Snowsfield's, London SE1 3SS, U.K.
(E-mail: e.simonoff@iop.kcl.ac.uk).

continuous traits, comorbidity may be measured by a correlation coefficient significantly different from zero.

(2) *Sampling bias*. Features of the ascertainment process may increase the probability of obtaining comorbid cases. This may be the case for clinically ascertained samples (Berkson, 1946; Pauls, Leckman, & Cohen, 1993), where treatment-seeking may be greater in comorbid cases. Sampling bias can be ruled out in the case of complete ascertainment and the likelihood of it decreased when a variety of alternative ascertainment procedures are employed.

(3) *Population stratification*. Two comorbid disorders may have completely different risk factors, but the prevalence of the risk factors is greatest in the same population groups, so that both disorders are increased in these groups. Even when population sampling is unbiased, this effect may lead to comorbidity.

(4) *Symptom overlap*. The two disorders have shared symptoms, producing overlap only because certain symptoms occur in both. A question that then arises is whether there is evidence to support the differentiation of the two disorders. Amongst the factors determining whether separate classification is correct are the extent to which there are differences in aetiology, response to intervention, and outcome. For example, several symptoms overlap between post-traumatic stress disorder and depression. However, aetiological factors are considered different in the two disorders.

(5) *Correlated error variance*. All measures of disorder will have some measurement error. When the measurement error for the two disorders is correlated, an association based only on the measurement error can occur. This is more likely to occur when the same person is rating both disorders and when evaluations are relative rather than absolute, as in the case of behavioral ratings (Fergusson & Horwood, 1987a, b; Simonoff, Pickles, Hervas, et al., 1998; Simonoff et al., 1995).

(6) *A distinct group*. Comorbid conditions could represent a distinct entity from the two simpler conditions. In determining the presence of a distinct entity, the principles underlying classification would apply. The ICD-10 diagnosis of hyperkinetic conduct disorder could be considered qualitatively distinct on the basis of increased rates of family history of antisocial behavior and more negative outcomes (although it could also be argued that the differentiation of comorbidity and severity is difficult) (Faraone, Biederman, Jetton, & Tsuang, 1997).

(7) *Shared risk factors*. Risk factors for the two disorders overlap. Because the risk factors influence both disorders, the disorders will co-occur at greater than chance levels. These risk factors may be conceptualized as genetic or environmental risk factors, and also as a third trait or disorder with its own genetic and environmental risk factors.

(8) *Phenotypic causality*. One disorder confers a risk for the other. This may occur in one direction only, or reciprocally, with each disorder increasing the risk of the other. An example of this would be hypertension and atherosclerosis, where hypertension is a direct cause of atherosclerosis.

Because the implications of these varying reasons for comorbidity are quite different, the alternatives need to be considered seriously and a number of different types of data collection and analysis may be required to evaluate the likely role of each as an explanatory mechanism. For example, the analysis of the role of risk factors in population subgroups could also alert the researcher to heterogeneity across the population (Neale & Cardon, 1992; Neale & Kendler, 1995). Psychometric methods for exploring the question of symptom overlap are highlighted by Neale and Kendler. The role of correlated error variance can be examined when more than one data source is available, e.g. Simonoff et al. (1995).

Within the model of shared risk factors, discrimination of the nature of the shared risk factor is potentially important for intervention. In some ways the model of a third, measured entity mediating shared risk is always the ultimate mechanism, as both genes and environment exert their effects through potentially measurable intermediate effects. The extent to which these intermediate effects can be identified varies but is often important in designing treatments. For example, the risk of mental retardation in phenylketonuria is conferred genetically, but effective intervention occurs at the intermediate step of reducing plasma phenylalanine levels. In exploring the links between low IQ and child psychiatric problems (Goodman, Simonoff, & Stevenson, 1995; Goodman & Yude, 1996), it is important to know whether the comorbidity is due to shared susceptibility genes and/or suboptimal environments that directly influence both problems, or whether it is mediated by neurological impairment. In practice, it may not always be possible to measure potential third factors, but the use of genetic strategies to determine the extent to which risks are likely to be genetic or environmental may help to focus our search for mediators.

A potentially important but often unexplored reason for comorbidity is phenotypic causality, where the presence of one disorder alters the risk of another. This is most powerfully explored with longitudinal data, but genetically informative designs with cross-sectional data may be able to discriminate both between reciprocal versus unidirectional causality and also between the two alternative directions of causation (Heath et al., 1993; Neale & Cardon, 1992). In addition, it may be possible to distinguish between the models of shared risk factors and phenotypic causality. The differentiation is important in understanding mechanisms of risk and planning interventions. Under the shared genes/environment risk factors model, intervention at the level of the shared risk factor will have an impact on both disorders, while intervention at the phenotypic level for one disorder has no effect on the other. Under the unidirectional phenotypic causality model, however, intervention with respect to the "upstream" disorder will alter the "downstream" disorder. In the reciprocal causality model, influencing either disorder directly will affect the other.

Although behavior geneticists are aware of the range of explanations for comorbidity, the literature is replete with multivariate genetic analyses in which shared genetic/environmental factors are the only models explored. In some cases, this may be far and away the most plausible mechanism for comorbidity. In other cases, there may be difficulties in testing a variety of alternative models with the data available in any one study. However, the general failure to explore alternative mechanisms may have left many readers unaware of the potential im-

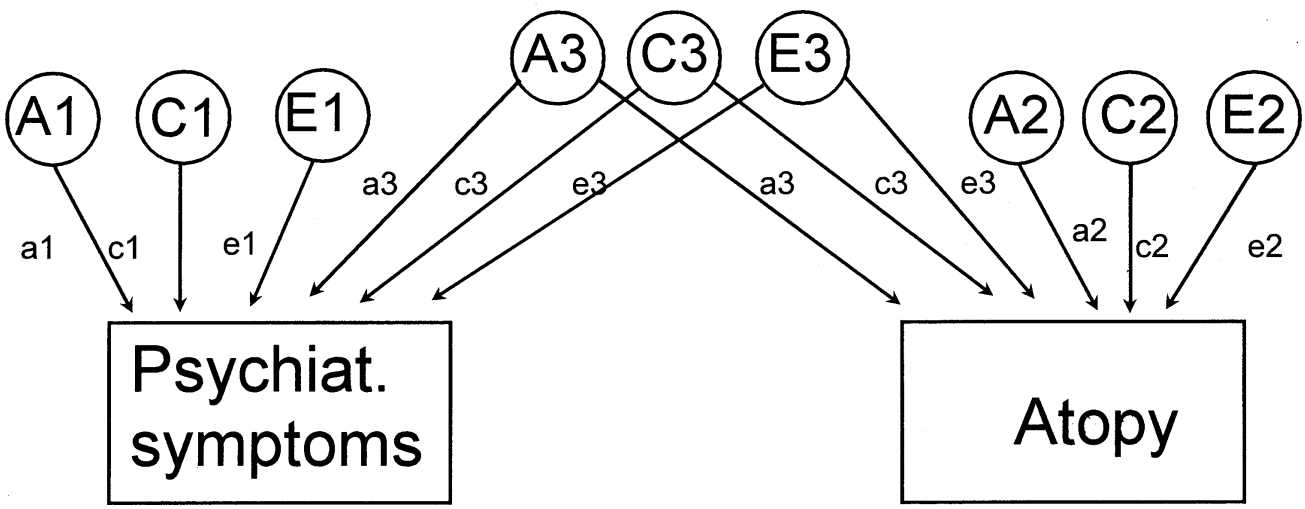

Figure 1. Bivariate genetic model. The figure shows the latent factors influencing one twin only. Psychiatric symptoms are affected both by additive genetic (A1), common environmental (C1), and unique environmental (E1) influences specific to psychiatric symptoms and also by additive genetic (A3), common environmental (C3), and unique environmental (E3) influences shared with atopy. Atopy is determined by specific additive genetic (A2), common environmental (C2), and unique environmental (E2) influences and ones (A3, C3, E3) shared with psychiatric symptoms. In the bivariate case, the magnitude of the A3, C3, and E3 parameters is constrained to be the same for both phenotypes. The bivariate genetic model is based on the classical twin assumptions; genetic influences for A1, A2, and A3 respectively are correlated 1.0 in MZ twins and .5 in DZ twins. Common environmental influences (C1, C2, and C3) are perfectly correlated in both twin types and unique environmental influences (E1, E2, and E3) are uncorrelated in both twin types. The association between the two phenotypes is estimated by the parameters a3, c3, and e3.

portance of the range of explanations for comorbidity discussed previously. Although some explanations cannot be examined without additional data, the shared genetic/environmental risk and phenotypic causality models can both be tested with a single data source in classical twin designs.

In a recent paper, Wamboldt, Schmitz, and Mrazek (1998) apply a shared genetic/environmental risk factors analysis to the relationship between atopy and emotional and behavioral problems in young twins. In support of the claim for comorbidity between the two phenotypes, they cite evidence showing an association between a range of atopic conditions such as hay fever (allergic rhinitis), eczema (atopic dermatitis), and asthma and psychiatric symptoms of depression in adults or of behavioral inhibition in children. They also refer to evidence supporting a familial component to this association, in which family members of children with hard-to-manage asthma have higher rates of a range of psychiatric symptoms. There are some methodological concerns with a number of the studies. However, in their own study, Wamboldt et al. demonstrate relatively small but significant correlations between atopy and different aspects of emotional and behavioural problems (.15–.21). Therefore, we shall take as given that there is a relationship, although it may not be strong, between psychiatric symptoms and atopy. In their discussion of the possible reasons for the association, the authors highlight some of the alternative causes of comorbidity, focusing on shared genetic and environmental risk, phenotypic causality, and correlated error variance. However, the analyses presented explore the shared genes/environment model only. Because the authors published their data, it was possible to explore the alternative model of phenotypic causality. Although consideration of correlated error variance is important because ratings of both behavior and atopy were made by mothers on both twins, this could not be examined, as there were no other raters.

The analyses conducted by Wamboldt et al. are comparable to many currently being published. The use of their data in this paper is not a criticism of their work but rather as an illustration of the importance more broadly of considering the likely causes for associations among disorders and to use an appropriate array of models to explore the alternative explanations.

Method

Data

Analyses were all based on the data made available by Wamboldt et al. (1998) on twins ranging in age from 3 years 9 months to 11 years with a mean age of 7 years 7 months. Detail of the sample is given in their article. There are minor deviations in the sample sizes in whom there were complete ratings from those reported by the authors: there were 58 monozygotic (MZ) pairs, 45 same-sex dizygotic (DZ-SS) and 88 dizygotic opposite-sex (DZ-OS) pairs.

Data Analysis

Data analysis was conducted using Mx (Neale, 1996), a structural equation modeling package designed for genetically informative studies. The individual fit of models was assessed using the χ^2 statistic, with appropriate degrees of freedom (*df*) and associated *p* value, and the Akaike Information Criterion (AIC, calculated as χ^2-2*df*). In both cases smaller (negative in the case of the AIC) values are associated with good fit. Comparison of nested models used the likelihood ratio χ^2 (LR χ^2; difference in the two χ^2 values) with the associated degrees of freedom being the difference in these in the compared models. Because the bivariate genetic model divides the paths associating the two phenotypes into three (genetic, environmental, and unique environmental), in comparison with the phenotypic causality model, the latter can be conceptualized as nested. The AIC is often used to compare non-nested models and the 95% confidence intervals for this statistic are presented.

The fit of two alternative genetic models of phenotypic association was compared. The full *bivariate genetic model* is shown in Fig. 1. Loehlin (1996) has helpfully discussed the alternative models of shared genetic and environmental influ-

https://doi.org/10.1017/S0021963099225826 Published online by Cambridge University Press

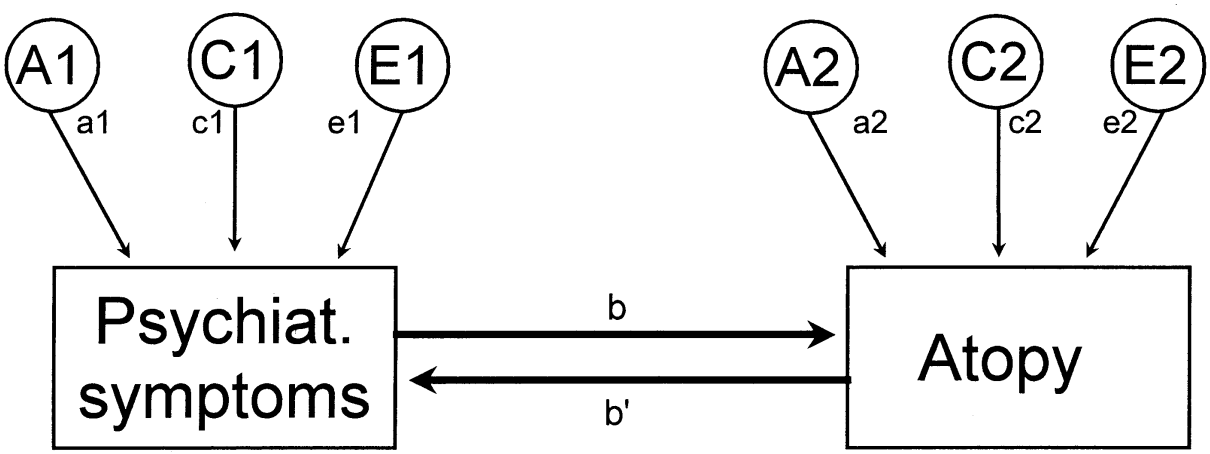

Figure 2. Phenotypic causality model. The figure shows the latent factors influencing one twin only. Psychiatric symptoms are affected by additive genetic (A1), common environmental (C1), and unique environmental (E1) influences and the phenotype of atopy whereas atopy is determined by specific additive genetic (A2), common environmental (C2), and unique environmental (E2) influences and the phenotype of psychiatric symptoms. The phenotypic causality model is based on the classical twin assumptions as indicated for Fig. 1. The association between the two phenotypes is estimated by the phenotypic paths b (from psychiatric symptoms to atopy) and b′ (from atopy to psychiatric symptoms).

ences for interested readers. Although the solutions give the same fit in the bivariate case and can be transformed to identify the same parameters, different models may be preferred in various situations because of the relative ease with which the output answers the research question. Model 1 is the parameterization used by Wamboldt et al. (1998). Each phenotype is comprised of its own set of genetic (A), common environmental (C), and unique environmental (E) factors and also of genetic, common environmental, and unique environmental factors shared with the other phenotype. In the bivariate case, to identify a solution, the value of parameters or the shared influences (A3, C3, and E3) must be constrained to be equal for the two phenotypes.

The *phenotypic causality model* is shown in Fig. 2. Here, both phenotypes have their individual genetic, common environmental, and unique environmental factors. The association between the two phenotypes comes from the effect of one phenotype on the other. In Fig. 2, this is depicted as a reciprocal pathway (b and b′), with each influencing the other, but the relationship may also be in one direction only. The usual test is to fit the reciprocal model and then the two alternative unidirectional models. If there is a significant reduction in fit with one unidirectional model but not the other, the one in which fit is not significantly reduced is accepted over both the other unidirectional and the reciprocal model. In practice, it is easier to distinguish direction of causality when the relative importance of genetic and environmental phenotypes on the two phenotypes are considerably different and also when the association between the phenotypes is strong. For interested readers, a detailed description of the phenotypic causality model, the underlying assumptions, and its limitations are given by Heath et al. (1993).

Results

For each of the three areas of psychiatric symptomatology analyzed by Wamboldt et al., results of the bivariate genetic and phenotypic causality models are given. The reciprocal and unidirectional phenotypic causality models have all been tested. For the bivariate genetic models, only those selected by Wamboldt et al. (see Corrigendum, p. 674) as the best fitting models are given. In all cases, this was the model in which there were genetic, common environmental, and unique environmental influences on psychiatric symptoms, genetic and unique environmental influences on atopy, and genetic and unique environmental influences shared between the two. For the phenotypic causality models, psychiatric symptoms were parameterized with genetic, common environmental, and unique environmental influences. Atopy was parameterized initially with genetic, common environmental, and unique environmental influences and subsequently common environmental influences were eliminated, with a test of the effect on model fit. The latter is conceptually most similar to the parameterization in the bivariate genetic models selected by Wamboldt et al. but the former provides an important test of the absence of common environmental influences on atopy under a different model. In all cases, the elimination of common environmental influences on atopy did not significantly worsen the fit and the results of the models including common environment are not shown. Because the aim was not to determine which model was most parsimonious (indeed with the relatively small sample sizes available, it could be argued that this is not appropriate), the impact of further removing genetic and environmental parameters from the phenotypic causality models was not explored. Only those models in which the direction of causation was altered were examined.

Table 1 shows the results for atopy symptoms and internalizing symptoms on the Child Behavior Checklist (CBCL). All four models demonstrate a good fit by χ^2 statistics. A good fit is usually easier to achieve with small sample sizes and therefore the absolute level of fit here is less important than the relative fit of different models and the ability to discriminate between models. For the 1 *df* comparison between nested models, an increase in the χ^2 of 3.84 is required to show that the more parsimonious model gives a worse fit. In relation to atopy and internalizing symptoms, neither of the unidirectional models was significantly worse than the reciprocal model (LR χ^2 1.29 and 0.19). Although one interpretation could be that both parsimonious models gave an equally good fit, another interpretation is that the power to discriminate the direction of causality is insufficient with the current data. The 95% confidence intervals on the AIC are given for all models, demonstrating a high level of overlap. This substantiates the inability to distinguish among different models of phenotypic causality. In

https://doi.org/10.1017/S0021963099225826 Published online by Cambridge University Press

Table 1
Atopy and Internalizing Behavior

Model	χ^2	*df*	*p*	95% CI on AIC
Bivariate genetic model[a]	19.62	23	.604	−46.000 to −32.189
Reciprocal causal model	19.51	23	.671	−46.000 to −32.357
Unidirectional causal model				
Behavior → atopy	20.80	24	.650	−48.000 to −33.580
Atopy → behavior	19.70	24	.714	−48.000 to −35.331

[a] Best fitting model in Wamboldt et al. (1998).

Table 2
Atopy and Externalizing Behavior

Model	χ^2	*df*	*p*	95% CI on AIC
Bivariate genetic model[a]	18.13	23	.751	−46.000 to −34.605
Reciprocal causal model	19.96	23	.644	−46.000 to −31.651
Unidirectional causal model				
Behavior → atopy	21.07	24	.635	−48.000 to −33.161
Atopy → behavior	20.46	24	.670	−48.000 to −34.121

[a] Best fitting model in Wamboldt et al. (1998).

Table 3
Atopy and Total Problems Score

Model	χ^2	*df*	*p*	95% CI on AIC
Bivariate genetic model[a]	21.97	23	.520	−46.000 to −28.518
Reciprocal causal model	23.21	23	.449	−46.000 to −26.625
Unidirectional causal model				
Behavior → atopy	23.58	24	.486	−48.000 to −29.289
Atopy → behavior	23.42	24	.495	−48.000 to −29.537

[a] Best fitting model in Wamboldt et al. (1998).

comparing the bivariate genetic model with the two unidirectional phenotypic causality models, the LR χ^2 were once again not significant for the 1 *df* test (1.19 and 0.12, respectively). The AIC 95% confidence intervals showed considerable overlap for both the unidirectional and the reciprocal phenotypic causality models when compared with the bivariate genetic model.

Tables 2 and 3 give results for the same models using the externalizing behavior and total problems score from the CBCL. Although there are minor differences in the overall levels of fit, the pattern of findings is very similar. In both cases, there is no significant worsening of the fit of either unidirectional phenotypic causality models over the reciprocal model (LR χ^2 0.29 and 0.59 for externalizing symptoms, 0.37 and 0.21 for total problems score). In comparing the 95% confidence intervals of the AIC for the bivariate genetic and different phenotypic causality models, they are very similar for models examining the same symptoms, suggesting little power to discriminate among models.

As a further demonstration of the difficulty in discriminating between bivariate genetic and phenotypic causality models, power calculations were carried out, based on the current data, to determine the sample sizes required to discriminate between two models. The comparison was made using this bivariate genetic model given in Tables 1–3 and the phenotypic causality model in which atopy → psychiatric symptoms, as this showed the lowest χ^2 in all three cases. Power calculations were performed using the χ^2 difference between the two models as a 1 *df* test. Sample sizes to obtain 80% power at a significance level of .05 were determined in Mx. For internalizing symptoms, a total sample size of 12,493 subjects was required to reject the phenotypic causality model; for externalizing and total problem score, the required sample sizes were 643 and 1027, respectively.

As a more general illustration of the sample sizes required to discriminate between phenotypic causality and bivariate genetic models, the results of a series of power calculations are given in Table 4. It is not possible to reject the bivariate genetic model when the phenotypic causality model has been specified, as the single causal path can always be partitioned in a bivariate genetic model into genetic and environmental paths. The proportions of the genetic and environmental paths shared between phenotypes will be determined by the genetic and environmental parameter estimates for the two traits, the strength of the causal association, and whether the phenotypic causal pathways are unidirectional or reciprocal. However, it is possible to exclude the phenotypic causality model when the true model is the bivariate genetic model. Table 4 gives the total sample sizes required for 80% power to reject the phenotypic causality model at a significance level of .05. Results are shown for various values of the bivariate genetic model and for the three possible phenotypic causality models: trait 1 → trait 2, trait 2 → trait 1, and reciprocal causality. The sample sizes given are for the total number of twin pairs. Wide variation in required sample sizes can be seen and several general principles emerge. First, the larger the parameters in the shared pathways (i.e. the stronger the phenotypic association), the greater the power (the smaller the sample

https://doi.org/10.1017/S0021963099225826 Published online by Cambridge University Press

Table 4
Power Calculations to Reject Phenotypic Causality Model under True Model Bivariate Genetic Model

True model parameters (bivariate genetic model)									Sample sizes to reject various phenotypic causality models		
Trait 1-specific			Trait 2-specific			Shared Trait 1 and Trait 2					
$a1^2$	$c1^2$	$e1^2$	$a2^2$	$c2^2$	$e2^2$	$a3^2$	$c3^2$	$e3^2$	Trait 1 → Trait 2	Trait 2 → Trait 1	Reciprocal causation
.2	.2	.2	.2	.2	.2	.2	.2	0	253	253	228
.2	.2	.2	.2	.2	.2	0	.2	.2	836	836	2.06×10^{13}
.2	.2	.2		.2	.2	.2	0	.2	2553	2553	3.89×10^{13}
.2	.2	.2	.2	.2	.2	.2	.1	.1	6890	6890	1.31×10^{14}
.2	.2	.2	.2	.2	.2	.1	.2	.1	2387	2387	1.81×10^{13}
.2	.2	.2	.2	.2	.2	.1	.1	.2	1745	1745	1430
.3	.3	.3	.3	.3		.1	0	0	3443	3443	1.46×10^{14}
.3	.3	.3	.3	.3	.3	0	.1	0	1216	1216	1.77×10^{14}
.3	.3	.3	.3	.3	.3	0	0	.1	927	927	741
.4	.2	.2	.2	.2	.4	.1	.1	0	1002	408	4061
.4	.2	.2	.2	.2	.4	.1		.1	1198	9322	13458
.4	.2	.2	.2	.2	.4	0	.1	.1	2808	10365	9480
.4	.1	.1	.4	.1	.1	.2	.1	.1	2264	2264	1876
.4	.1	.1	.4	.1	.1	.1	.2	.1	1247	1247	1049
.4	.1	.1	.4	.1	.1	.1	.1	.2		309	201
.4	.1	.1	.4	.1	.1	.4	0	0	195	195	172
.4	.1	.1	.4	.1	.1	0	.4	0	70	70	1.56×10^{12}
.4	.1	.1	.4	.1	.1	0	0	.4	92	92	2.65×10^{13}
.4	.1	.1	.1	.1	.4	.2	.1	.1	7169	326	
.4	.1	.1	.1	.1	.4	.1	.2	.1	2313	248	1866
.4	.1	.1	.1	.1	.4	.1	.1	.2	616	2304	498
.4	.1	.1	.1	.1	.4	.4	0	0	822	83	10007
.4	.1	.1	.1	.1	.4	0	.4	0	122	52	298
.4		.1	.1	.1	.4	0	0	.4	124	764	23870
.3	.2	.1	.2	.3	.1	.2	.1	.1	2341	1124	6283
.3	.2	.1	.2	.3	.1	.1	.2	.1	2173	2780	2834
.3	.2	.1	.2	.3	.1	.1	.1	.2	290	238	4102
.3	.2	.1		.2	.3	.2	.1	.1	3517	794	2561
.3	.2	.1	.1	.2	.3	.1	.2	.1	4485	580	3645
.3	.2	.1	.1	.2	.3	.1	.2	.2	435	36480	33781
.3	.2	.2	.3	.2	.2	.1	.1	.1	27808	27808	20540
.2	.1	.1	.2	.1		.2	.2	.2	4580	4580	2353
.2	.1	.1	.1	.2	.1	.2	.2	.2	4553	1798	4519
.2	.1	.1	.1	.1	.2	.2	.2	.2	6697	1289	17927

$a1^2 + c1^2 + e1^2 + a3^2 + c3^2 + e3^2 = 1.0$; $a2^2 + c2^2 + e2^2 + a3^2 + c3^2 + e3^2 = 1.0$.

sizes required to reject the phenotypic causality model). Second, the more the shared influences are unevenly distributed across the three possible parameters (genetic, common environmental, and unique environmental), the greater the power. It can be seen that there are some values for the bivariate genetic model where effectively the reciprocal causality model cannot be rejected. In those cases, the phenotypic causal paths have gone to values that are the same but differ in sign, e.g., −0.5 and 0.5. Although the possibility that such equal opposing effects may exist must be considered, such parameter estimates should alert the researcher to the possibility that the genetic and environmental parameters allow identification of a phenotypic model that cannot be resolved. As the usual method of testing models of phenotypic causality is first to test the reciprocal model, and then to compare its fit with that of the two unidirectional models, these are serious implications for distinguishing between bivariate genetic and phenotypic causality models. In instances where there is adequate power to reject unidirectional but not reciprocal models, it would not be possible to accept the true bivariate genetic model. This table complements that published by Heath et al. (1993), which examines the power to discriminate between the alternative phenotypic causality models.

Discussion

Understanding the reasons for comorbidity is a central question in child psychopathology research. A number of the reasons for comorbidity were outlined. Genetic strategies have the potential to contribute in important ways to testing alternative reasons for its existence. The designs and samples required to explore the alternatives properly do not often arise without careful planning, however. For example, to explore whether population stratification is a cause of comorbidity, population sampling in a manner that allows stratification could allow direct examination. Assessment of the role of correlated measurement error requires that multiple indicators of each trait be obtained. Therefore, those interested in using genetic analyses to answer substantive questions must plan ahead to consider what data should be collected to test alternative hypotheses.

In the last 10 years, a number of papers have been published using genetic strategies to explore associated phenotypes. In the overwhelming number of cases, the

https://doi.org/10.1017/S0021963099225826 Published online by Cambridge University Press

bivariate genetic model has been tested, with the aim of seeing whether shared risk could be described as due to either genetic or environmental factors that are shared across the phenotypes. In almost all instances, the possibility of phenotypic causality has been ignored or given short shrift. The current analyses show that the data of Wamboldt et al. (1998) have insufficient power to discriminate between the bivariate genetic and phenotypic causality models with the current sample sizes. Rather, sample sizes would seem to need to be increased some 3- to 65-fold before discrimination would be possible.

An important limitation of the phenotypic causality model relates to that of error variance, or unreliability of measurement, discussed by both Heath et al. (1993) and Neale and Cardon (1992). First, unreliability in the measurement of the traits will reduce the estimated casual paths, if it is not accounted for. Second, *differential* unreliability in the measurement of the two traits can lead to false conclusions about whether a reciprocal causality or a unidirectional model best fits the data and can even lead to acceptance of the wrong direction of causality (e.g., B → A when the correct model is A → B). Attempts to reduce unreliability, such as accounting for test–retest reliability and the use of multiple measures of each phenotype, are therefore very important. It should also be noted that unreliability of measurement will affect parameter estimates for any genetic model, and not just those of phenotype causality.

Although there may be many different reasons for comorbidity, as illustrated above, discrimination between the bivariate genetic and phenotypic causality models may be particularly important because there are potentially very different implications for intervention. Under the bivariate genetic model, the reduction of psychiatric symptoms occurring in the presence of atopy will only alter the rates of atopy if the intervention reduces risk factors shared between the two disorders. Under the phenotypic causality model, however, reduction of one phenotype can have a direct impact on the other. Under the atopy → behavior model, where the causal path coefficient is positive, reduction in atopic symptoms by whatever method would influence psychiatric symptoms. However, reduction in psychiatric symptoms would not affect atopy. Were the direction of causation reversed, then decreasing psychiatric symptoms would lower atopic symptoms. It is worth noting that, under the reciprocal model, it is possible to have a positive path coefficient in one direction; e.g., atopy → psychiatric symptoms, and a smaller negative path coefficient in the opposite direction, e.g., psychiatric symptoms → atopy, while retaining an overall positive phenotype correlation. In this case, a reduction in atopy would lead to a reduction in psychiatric symptoms but a reduction in psychiatric symptoms would cause (a smaller) increase in atopy.

We have seen that it is only possible to reject the phenotypic causality model when the true model is the bivariate genetic model, and not the reverse. Furthermore, under some sets of genetic and environmental parameters estimates, the sample sizes required are very large, often greater than those available in many of the current twin studies. Many readers may not be aware that the methods of structural equation modeling mean that smaller sample sizes, for the same variance-covariance matrices, will give better fit statistics, potentially luring the reader into the belief that the models explored give the best explanation. Rather, it may often be the case that there was insufficient power to detect the *lack* of fit. The implication is that analytic strategies must be selected to consider the plausible alternative explanations for the phenomenon under investigation.

In selecting between the bivariate genetic and phenotypic causality models, researchers should take into account a number of features in ranking the plausibility of the two models. Is there a consistent temporal relationship between the two traits? Where there is, a unidirectional phenotypic causality model may be the one preferred, although temporal links should not be taken to imply causal ones (Heath et al., 1993). It is entirely possible that shared genetic and environmental risk factors exert their effects on different traits at various points in development, so that one trait appears before another. Furthermore, this does not help in relation to reciprocal causality models, where a temporal relationship would not be expected. As information from other sources, e.g., biological and psychological, about risk factors becomes available, this can be brought into consideration in evaluating competing mechanisms of comorbidity. The role of third variables in mediating covariance often needs evaluation. Ideally, the possible causes of comorbidity should be identified at the stage of study design, so that sampling and measurement strategies can test more convincingly the causes of comorbidity. Genetic strategies will remain an important component of the investigation of comorbidity. However, additions to designs, including longitudinal data collection, stratified sampling, direct measures of third variables, and multiple raters, can greatly increase the robustness of interpretation. Because all designs have limitations, it will be important to demonstrate that similar results can be obtained from a variety of studies. In the meantime, the genetic designs considered here will be important, but need to be evaluated with knowledge of their limitations.

References

Berkson, J. (1946). Limitations of the application of four-fold table analysis to hospital data. *Biometrics*, *2*, 47–53.

Cardon, L. R., & Fulker, D. W. (1994). A model of developmental change in hierarchical phenotypes with application to specific cognitive abilities. *Behavior Genetics*, *24*, 1–16.

Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry*, *32*, 1063–1081.

Eaves, L. J., Silberg, J. L., Meyer, J. M., Maes, H. H., Simonoff, E., Pickles, A., Rutter, M., Neale, M. C., Reynolds, C. A., Erickson, M. T., Heath, A. C., Loeber, R., Truett, K. R., & Hewitt, J. K. (1997). Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. *Journal of Child Psychology and Psychiatry*, *38*, 965–980.

Faraone, S. V., Biederman, J., Jetton, J. G., & Tsuang, M. T. (1997). Attention deficit disorder and conduct disorder: Longitudinal evidence for a familial subtype. *Psychological Medicine*, *27*, 291–300.

Fergusson, D. M., & Horwood, L. J. (1987a). The trait and method components of ratings of conduct disorder—part I. Maternal and teacher evaluations of conduct disorder in young children. *Journal of Child Psychology and Psychiatry*, *28*, 249–260.

Fergusson, D. M., & Horwood, L. J. (1987b). The trait and method components of ratings of conduct disorder—part II. Factors related to the train component of conduct disorder scores. *Journal of Child Psychology and Psychiatry*, *28*, 261–272.

Gjone, H., & Stevenson, J. (1997). The association between internalizing and externalizing behavior in childhood and early adolescence: Genetic or environmental common influences? *Journal of Abnormal Child Psychology*, *25*, 277–286.
Goodman, R., Simonoff, E., & Stevenson, E. (1995). The impact of child IQ, parental IQ and sibling IQ on child behavioural deviance scores. *Journal of Child Psychology and Psychiatry*, *36*, 409–425.
Goodman, R., & Yude, C. (1996). IQ and its predictors in childhood hemiplegia. *Developmental Medicine and Child Neurology*, *38*, 881–890.
Heath, A. C., Kessler, R. C., Neale, M. C., Hewitt, J. K., Eaves, L. J., & Kendler, K. S. (1993). Testing hypotheses about direction of causation using cross-sectional family data. *Behavior Genetics*, *23*, 29–50.
Huziak, J., Heth, A., Madden, P., Reich, W., Bucholz, K., Slutske, W., Bierut, L., Neuman, R., & Todd, R. (1998). Latent class and factor analysis of DSM-IV ADHD: A twin study of female adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 848–857.
Kendler, K. S., Heath, A. C., Martin, N. G., & Eaves, L. J. (1987). Symptoms of anxiety and symptoms of depression. Same genes, different environments? *Archives of General Psychiatry*, *44*, 451–457.
Loehlin, J. C. (1996). The Cholesky approach: A cautionary note. *Behavior Genetics*, *26*, 65–69.
Nadder, T. S., Silberg, J. L., Eaves, L. J., Maes, H. H., & Meyer, J. M. (1998). Genetic effects on ADHD symptomatology in 7- to 13- year-old twins: Results from a telephone survey. *Behavior Genetics*, *28*, 83–99.
Neale, M. (1996). *Mx: Statistical modelling* (3rd ed.). Richmond, VA: Virginia Commonwealth University.
Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
Neale, M. C., & Kendler, K. S. (1995). Models for comorbidity of multifactorial disorders. *American Journal of Human Genetics*, *57*, 935–953.
O'Connor, T. G., Hetherington, E. M, Reiss, D., & Plomin, R. (1995). A twin-sibling study of observed parent-adolescent interactions. *Child Development*, *66*, 812–829.
Pauls, D. L., Leckman, J. F., & Cohen, D. J. (1993). Familial relationship between Gilles de la Tourette's syndrome, attention deficit disorder, learning disabilities, speech disorders, and stuttering. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 1044–1050.
Petrill, S. A., & Thompson, L. A. (1993). The phenotypic and genetic relationships among measures of cognitive ability, temperament, and scholastic achievement. *Behavior Genetics*, *23*, 511–518.
Rende, R. D., Plomin, R., Reiss, D., & Hetherington, E. M. (1993). Genetic and environmental influences on depressive symptomatology in adolescence: Individual differences and extreme scores. *Journal of Child Psychology and Psychiatry*, *34*, 1387–1398.
Silberg, J., Rutter, M., Meyer, J., Simonoff, E., Hewitt, J., Loeber, R., Pickles, A., Maes, H., & Eaves, L. J. (1996). Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *Journal of Child Psychology and Psychiatry*, *37*, 803–816.
Simonoff, E., Pickles, A., Hervas, A., Silberg, J. L., Rutter, M., & Eaves, L. J. (1998). Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. *Psychological Medicine*, *28*, 825–837.
Simonoff, E., Pickles, A., Hewitt, J., Silberg, J., Rutter, M., Loeber, R., Meyer, J., Neale, M., & Eaves, L. J. (1995). Multiple raters of disruptive child behavior: Using a genetic strategy to examine shared views and bias. *Behavior Genetics*, *25*, 311–326.
Simonoff, E., Pickles, E., Meyer, J., Silberg, J., & Maes, H. (1998). Genetic and environmental influences on subtypes of conduct disorder. *Journal of Abnormal Child Psychology*, *26*, 495–509.
Slustske, W. S., Heath, A. C., Dinwiddie, S. H., Madden, P. A. F., Bucholz, K. K., Dunne, M. P., Statham, D. J., & Martin, N. G. (1997). Modeling genetic and environmental influences in the etiology of conduct disorder: A study of 2682 adult twin pairs. *Journal of Abnormal Psychology*, *106*, 269–279.
Thapar, A., & McGuffin, P. (1996). Genetic influences on life events in childhood. *Psychological Medicine*, *26*, 813–820.
Wamboldt, M. Z., Schmitz, S., & Mrazek, D. (1998). Genetic association between atopy and behavioural symptoms in middle childhood. *Journal of Child Psychology of Psychiatry*, *39*, 1007–1016.

Manuscript accepted 23 December 1999

Corrigendum

To Wamboldt, Schmitz, & Mrazek (1998). *Genetic association between atopy and behavioural symptoms in middle childhood*. Volume *39* (7), pp. 1007–1016.

Through Dr Emily Simonoff's re-analysis of our data (see article above, pp. 667–674) we found a slight model specification change which has now been corrected and which actually provides stronger results. Revised Tables 4–6 are available from the authors by request.

Correct covariance matrix for the DZ opposite-sex pairs for atopy and total problem score

	1	2	3	4
1	0.710			
2	0.352	2.393		
3	0.213	0.044	0.825	
4	0.025	1.717	0.137	2.980

Corrected last paragraph of the Results section

Additive genetic effects accounted for most of the correlation between atopy and the three CBCL scores: 77% of the correlation between atopy and internalising symptoms, 100% of the correlation between atopy and externalising symptoms, and 93% of the correlation between atopy and total problem scores. The genetic correlation (rG) between atopy and INT was .34 and the nonshared environmental correlation (rE) was .16; for atopy and EXT rG was .29 and there was no rE; for total problems and atopy rG was estimated at .37 and rE at .06. The parameter estimates obtained from the bivariate factor model (full model, or A & E) are as follows: for atopy, h2 = .72 and e2 = .28; for INT, h2 = .35, c2 = .28 and e2 = .37; for EXT, h2 = .55, c2 = .23 and e2 = .22; and for Total Problems, h2 = .34, c2 = .45 and e2 = .21.