Testing Hypotheses Regarding the Causes of Comorbidity: Examining the Underlying Deficits of Comorbid Disorders

Soo Hyun Rhee, John K. Hewitt, Robin P. Corley, and Erik G. Willcutt
University of Colorado at Boulder

Bruce F. Pennington
University of Denver

The authors examined the validity of a method commonly used to test alternative hypotheses regarding the causes of comorbidity: the examination of underlying deficits of comorbid disorders. The authors simulated data in which the true causes of comorbidity were known, then compared the patterns of underlying deficits of the comorbid disorders found in the simulated data with the predicted results. The method of examining the underlying deficits of comorbid disorders could distinguish between several comorbidity models, including those that could not be distinguished well using other methods. The ability to distinguish the correct model decreased as the sample size and the correlation between the underlying deficits and the symptom scores decreased, suggesting that the issue of power should be considered carefully.

Keywords: comorbidity, comorbidity models, comorbid disorders, underlying deficits, simulations

Epidemiological evidence suggests that the rate of comorbidity among psychiatric disorders exceeds the rate expected by chance (e.g., Newman et al., 1996; Simonoff et al., 1997). Many researchers have considered alternative hypotheses for the causes of comorbidity and ways to test these hypotheses (Achenbach, 1990–1991; Angold, Costello, & Erkanli, 1999; Caron & Rutter, 1991; Klein & Riso, 1993; Neale & Kendler, 1995; Pennington, Groisser, & Welsh, 1993; Rutter, 1997; Simonoff, 2000). In the present study, we looked at the validity of one of the common approaches used to test alternative hypotheses regarding the causes of comorbidity, the examination of underlying deficits of the two comorbid disorders (e.g., Pennington et al., 1993; Schachar & Tannock, 1995; Willcutt et al., 2001), considering both an ideal situation in which the underlying deficit score and the symptom scores were perfectly correlated and more realistic situations in which the two scores were moderately correlated.

Improved knowledge regarding the causes of comorbidity between psychiatric disorders can have a significant impact on future research examining the two disorders, including research on classification, treatment, and etiology. For example, if comorbidity occurs because two disorders are alternate manifestations of a single liability distribution, this would suggest that there may be an artificial subdivision of one disorder into two disorders. In another example, if there is evidence that individuals with both disorders have a third, separate disorder with a separate etiology, it is possible that a treatment that is effective for one of the disorders when it occurs alone may not be effective for individuals with both disorders.

Models of Comorbidity

As mentioned above, alternative hypotheses for the causes of comorbidity between two psychiatric disorders have been addressed by many researchers. So far, Klein and Riso (1993) have provided the most comprehensive view on the subject. In 1995, Neale and Kendler presented the quantitative specifications of Klein and Riso’s models, describing 13 comorbidity models (see Figure 1). All of the 13 Neale and Kendler comorbidity models are versions of the continuous liability threshold model, which assumes that there is a continuous liability distribution of multifactorial causes (genetic and/or environmental causes) for a disorder and that a disorder occurs if an individual crosses a particular threshold in that liability distribution.

In Figure 1, the latent variable \( R \) refers to the multifactorial liability for each disorder (e.g., \( R_A \) = multifactorial liability for Disorder A, \( R_B \) = multifactorial liability for Disorder B). The liability distributions with the thresholds in the boxes are simply another way of representing the multifactorial liability for each disorder (note that the path coefficient from the latent variable \( R \) and the liability distributions with the thresholds is 1). The individuals who cross the threshold in the liability distribution manifest Disorder A or B. Below is a short description of the 13 comorbidity models. More detailed descriptions of the comorbidity models can be found in Klein and Riso (1993) and Neale and Kendler (1995).

Chance

The chance model hypothesizes that comorbidity between two disorders occurs by chance. If comorbidity occurs by chance and
the prevalence of one of the comorbid disorders is \( p \) and the prevalence of the other disorder \( r \), the prevalence of individuals with both disorders is \( p \times r \). A prevalence of comorbidity significantly higher than \( p \times r \) suggests that comorbidity is not due to chance.

Alternate Forms

The alternate forms model hypothesizes that comorbidity occurs because the two comorbid disorders are alternate manifestations of a single liability. For individuals who cross a particular threshold in that...
single liability distribution, the probability of having Disorder A is \( p \), and the probability of having Disorder B is \( r \). This means that both disorders share a single liability and that one person manifests Disorder A whereas another person manifests Disorder B because of chance or risk factors that vary across individuals. A Gene × Environment interaction in which the environmental risk is specific to an individual is an example of how comorbidity may occur through alternate forms. If two individuals have the same genetic liability but are exposed to different person-specific environmental risks, the first individual may manifest Disorder A whereas the second individual manifests Disorder B.

**Multiformity**

In multiformity models, an individual who has one disorder is at an increased risk for having the second disorder although he or she
does not have an elevated liability for the second disorder. Two of the multiformity models, the random multiformity of A and random multiformity of B models, are the phenocopy model often discussed in the literature. The phenocopy model hypothesizes that the first disorder produces a copy of the second disorder or that the first disorder is primary whereas the second disorder is secondary. For example, Pennington et al. (1993) suggested that reading disability (RD) might lead to the phenotypic manifestation of attention-deficit/hyperactivity disorder (ADHD) in the absence of etiological influences typically associated with ADHD in isolation. The phenocopy hypothesis suggests that a child might appear to be inattentive or hyperactive in the classroom because of frustration elicited by difficulties with reading rather than as a consequence of the neurocognitive difficulties that are typically associated with ADHD in the absence of RD.

In random multiformity models, if an individual has one disorder, the risk for having the second disorder is simply the probability \( p \), whereas in extreme multiformity models, one is affected by the second disorder if one crosses a higher, second threshold on the liability distribution. In random multiformity, being affected by Disorder A increases the probability of having Disorder B, and being affected by Disorder B increases the probability of having Disorder A. In random multiformity of A, only having Disorder A increases the probability of having Disorder B (although having Disorder B does not increase the probability of having Disorder A), whereas in random multiformity of B, only having Disorder B increases the probability of having Disorder A (although having Disorder A does not increase the probability of having Disorder B). Similarly, having either Disorder A or B increases the probability of having both disorders in extreme multiformity, whereas only having Disorder A increases the probability of comorbidity in extreme multiformity of A and only having Disorder B increases the probability of comorbidity in extreme multiformity of B.
**Three Independent Disorders**

According to the three independent disorders model, comorbidity occurs because the comorbid disorder is a disorder that is separate from either disorder occurring alone. It is sometimes referred to as the **subtype hypothesis** in the literature.

**Correlated Liabilities**

There are four types of correlated liabilities models: (a) correlated liabilities, (b) A causes B, (c) B causes A, and (d) reciprocal causation. The thing that the four models have in common is that there is a continuous relationship between the liability to one disorder and the liability to the second disorder. An increase in liability for one disorder is correlated with an increase in liability for the second disorder. (In contrast, in the multiformity models, a change in liability for one disorder has absolutely no effect on the second disorder unless an individual crosses the threshold for the first disorder and is actually affected by the disorder.)

In the correlated liabilities model, the relationship between the liability of the two disorders occurs because there is a significant correlation between the risk factors that influence A and the risk factors that influence B. In the other three models (A causes B, B causes A, and reciprocal causation), there is a direct causal relationship between the manifest phenotypes of the two disorders. In the A causes B model, Disorder A has a direct causal effect on Disorder B, and in the B causes A model, Disorder B has a direct causal effect on Disorder A, whereas in the reciprocal causation model, the two disorders cause each other in a feedback loop over time.

**Examining the Underlying Deficits of Disorders to Test Comorbidity Models**

Several researchers have been successful in examining personality correlates as common or unique components of comorbid disorders (e.g., Krueger, 2002). Taking this approach, Mineka, Watson, and Clark (1998) proposed an integrative hierarchical model for internalizing disorders, suggesting that negative affect is a shared component common to depression and anxiety disorders; that absence of positive affect is a specific, unique component of depression; and that there are unique factors associated with each of the anxiety disorders. Widiger and Clark (2000) noted that a parallel model may be helpful in understanding the causes of comorbidity among externalizing disorders and suggested that the personality dimension of disinhibition may be a major common factor.

Several researchers have also tested alternative, competing hypotheses regarding the causes of comorbidity by examining the underlying deficits, correlates, or outcomes of the two comorbid disorders (see Table 1). For example, Pennington et al. (1993) examined the causes of comorbidity between RD and ADHD by examining deficits in phonological processes (underlying cognitive deficits of RD) and deficits in executive functioning (underlying cognitive deficits of ADHD). Studies using this approach (e.g., Schachar & Tannock, 1995) have made specific predictions regarding the pattern of underlying deficits in patients with either or both of the comorbid disorders for 5 of the 13 Neale and Kendler (1995) models: the alternate forms model, the three independent disorders model, the correlated liabilities model, and the random multiformity of A/random multiformity of B models.

Researchers examining the underlying deficits of disorders to test comorbidity models base their conclusions regarding the causes of comorbidity on the answers to two questions: (a) Is there a significant double dissociation between Disorder A and Disorder B on measures of the underlying deficits (i.e., the deficits of Disorder A are not present in individuals with B only, and the deficits of Disorder B are not present in individuals with A only) and (b) what is the pattern of deficits in the comorbid group? The answer to the first question is important because the presence of double dissociation indicates that there are two separate sets of liabilities for Disorders A and B. The answer to the second question is important because the pattern of deficits in the comorbid group should depend on the causes of comorbidity between the two disorders. Therefore, the answer to the second question should discriminate the correct model among several comorbidity models hypothesizing two separate sets of liabilities for Disorders A and B.

The predictions regarding the alternate forms model, the correlated liabilities model, and the random multiformity of A/random multiformity of B are as follows:

**Table 1**

<table>
<thead>
<tr>
<th>Studies</th>
<th>Comorbidity model</th>
<th>Result expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schachar &amp; Tannock (1995)</td>
<td>Alternate forms</td>
<td>Comorbid group and the two pure groups will have equivalent levels of all deficits.</td>
</tr>
<tr>
<td>Leung et al. (1996)</td>
<td>Alternate forms</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
<tr>
<td>Pennington, Groisser, &amp; Welch (1993)</td>
<td>Three independent disorders</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
<tr>
<td>Purvis &amp; Tannock (2000)</td>
<td>Three independent disorders</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
<tr>
<td>Schachar &amp; Tannock (1995)</td>
<td>Three independent disorders</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
<tr>
<td>Leung et al. (1996)</td>
<td>Correlated liabilities</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
<tr>
<td>Lynam (1996)</td>
<td>Correlated liabilities</td>
<td>Comorbid group will exhibit deficits of one of the pure groups.</td>
</tr>
<tr>
<td>Manassis, Tannock, &amp; Barbosa (2000)</td>
<td>Random multiformity of A/random multiformity of B</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
<tr>
<td>Pennington, Groisser, &amp; Welch (1993)</td>
<td>Random multiformity of A/random multiformity of B</td>
<td>Comorbid group will exhibit deficits of both pure groups.</td>
</tr>
</tbody>
</table>
multiformity of B models have been consistent in the literature. If the alternate forms model is the correct comorbidity model, both of the pure groups (A only and B only) and the comorbid group (AB) should perform equivalently on each measure and differ from the control group. Therefore, there should be no double dissociation (e.g., Schachar & Tannock, 1995). Studies examining the underlying deficits of disorders to test alternative comorbidity models have made two assertions regarding the correlated liabilities model and the random multiformity of A/random multiformity of B models. First, there should be double dissociation (the authors of these studies [e.g., Pennington et al., 1993; Schachar & Tannock, 1995] have not specified whether the dissociation would be complete or incomplete). Second, the predictions regarding the pattern of deficits in the comorbid group should differ for the two models. If the correlated liabilities model is the correct comorbidity model, the two pure groups should have deficits on different variables, and the comorbid group should have the deficits of both disorders (e.g., Pennington et al., 1993). If the phenocopy model (random multiformity of A or random multiformity of B) is the correct comorbidity model, the comorbid group should have the deficits of the primary disorder (Disorder A in random multiformity of A and Disorder B in random multiformity of B) but not the deficits of the secondary disorder (Disorder B in random multiformity of A and Disorder A in random multiformity of B; e.g., Manassis, Tannock, & Barbosa, 2000).

In contrast, although researchers have agreed that there should be double dissociation if the three independent disorders model is the correct comorbidity model, they have made differing predictions regarding the pattern of deficits found in the comorbid group. Purvis and Tannock (2000) and Pennington et al. (1993) predicted that the comorbid group should exhibit deficits of both pure groups, whereas Leung et al. (1996), Manassis et al. (2000), Schachar and Tannock (1995), and Lynam (1996) predicted that the comorbid group should be different in the pattern of deficits shown from the two pure groups.

The discrepancies in the predicted results in these studies may simply reflect differences in the theoretical conceptualization of comorbidity despite the similarities in the terms used to describe the models (e.g., etiological subtype, subgroup, distinct condition). Pennington et al. (1993) hypothesized that the comorbid group’s profile should exhibit the deficits of both groups if the etiologic subtype hypothesis is the true model because the etiologic subtype hypothesis implies that a single etiology can sometimes produce both full syndromes in the same individual. Although similar language was used, this view is different from those of studies that have defined the etiological subtype as a third, distinct condition with underlying deficits that are separate from the underlying deficits of the two pure disorders (e.g., Schachar & Tannock, 1995).

The Present Study

The main goal of the present study was to test whether examining the underlying deficits of comorbid disorders would be a valid method of testing alternative hypotheses regarding the causes of comorbidity. First, we wanted to test whether the predictions frequently found in the literature would be valid. Second, we attempted to resolve the disagreement in the literature regarding the conceptualization of the three independent disorders model. Third, we wanted to test whether, using this method, one could discriminate between more comorbidity models than those typically discussed in the literature (see Table 1). Even if initial results were consistent with the predictions of a particular comorbidity model, they might also fit the predictions of alternative comorbidity models that have not been considered in previous studies. Also, there is a possibility that the method of examining underlying deficits could discriminate between comorbidity models that could not be distinguished well using other methods (e.g., Neale & Kendler, 1995; Rhee et al., 2004).

Simulations

We tested whether examining the underlying deficits of comorbid disorders would be a valid method of testing alternative hypotheses by conducting a simulation study. First, we simulated data in which the true cause of comorbidity was known. Simulated datasets were created for each of the 13 Neale and Kendler (1995) comorbidity models. In the dataset simulated for each model, we then examined the level of the underlying deficits associated with each disorder in four diagnostic groups (individuals with neither disorder, individuals with Disorder A only, individuals with Disorder B only, and individuals with both disorders). We compared the results found in data simulated for each of the comorbidity models with the predictions for that model found in the literature. There were two criteria for a valid prediction: (a) The predicted pattern of results should be found in the data simulated for that comorbidity model, and (b) to discriminate a particular model from alternative models, the predicted result should not be found in the data simulated for any other comorbidity model.

Four sets of simulations were conducted. Simulation 1 was a situation not likely to occur in nature, in which there was a perfect correlation between the symptom score and the underlying deficit score and a very large sample size (10,000 individuals); this simulation was conducted to test the method of examining underlying deficits in the most ideal situation possible. Simulation 2 was a more realistic situation in which the correlation between the symptom score and the underlying deficit was that reported for two disorders, ADHD and RD, and a very large sample size (10,000 individuals). Simulation 3 and Simulation 4 were similar to Simulation 2, but a smaller sample size was simulated.

Method


Data were simulated for each of the 13 Neale and Kendler (1995) comorbidity models with a simulation program that extended a simple simulation procedure for genetic models (Hewitt & Neale, 1989) using SAS (SAS Institute, 1985). The RANNOR function (a random number function that returns a random variate from a normal distribution) was used to simulate the genetic liability, shared environmental liability, and nonshared environmental liability for the two disorders, which were added together to create the overall liability score for each individual (i.e., the symptom score). Then, the simulation programs indicated which threshold an individual must cross to manifest either Disorder A or Disorder B, with the threshold chosen depending on the prevalence of Disorders A and B. The diagnosis score was simply 0 if an individual did not have a disorder and 1 if an individual had the disorder. Sample simulation programs that simulate data following the assumptions of the 13 Neale and Kendler comorbidity models are available on the following Web site: http://ibgww.colorado.edu/cedd/software/.

In Simulation 1, in which the correlation between the symptom score and the underlying deficit score was perfect, the underlying deficit score was also the symptom score. In Simulations 2–4, the correlation between the
symptom score and the underlying deficit score was .35 for Disorder A and .60 for Disorder B. To simulate the underlying deficit score, we again used the RANNOR function to simulate the genetic liability, shared environmental liability, and nonshared environmental liability for the two disorders, then added them together to create the overall underlying deficit score for each individual. Given that the underlying deficit score and the symptom score were significantly correlated, a proportion of the genetic liability, shared environmental liability, and nonshared environmental liability of the underlying deficit score was shared with that of the symptom score so that the phenotypic correlation between the symptom score and the underlying deficit score was .35 for Disorder A and .60 for Disorder B.

Rhee et al. (2004) showed that these simulations are valid. When the Neale and Kendler (1995) model-fitting analyses were conducted on the simulated datasets using the same simulation method, the correct model fit the data well, and the estimated model parameters were very close to the simulated model parameters. Four sets of simulations were conducted. For each set, simulations were conducted 50 times to determine how many times the correct comborbid model could be distinguished from alternative models.

In Simulation 1, the model parameters (i.e., the prevalences of the disorders, the rate of comborbidty, and the magnitude of genetic and environmental influences) were chosen so that they would approximate those for two common comorbid disorders, ADHD (Disorder A; e.g., Levy, Hay, McStephen, Wood, & Waldman, 1997) and RD (Disorder B; e.g., Willcutt et al., 2003). The correlation between the symptom score and the underlying deficits was perfect so that we could examine the ability of the examination of underlying deficits to discriminate the correct comborbid model in an ideal situation. In the simulations, the prevalence of both disorders was approximately 5%, with comborbidty of 30%, resulting in approximately 9,150 individuals with neither disorder, 350 individuals with ADHD only (Disorder A only), 350 individuals with RD only (Disorder B only), and 150 individuals with both ADHD and RD (Disorders A and B). For ADHD (Disorder A), heritability was .75, the magnitude of shared environmental influences was .00, and the magnitude of nonshared environmental influences was .25. For RD (Disorder B), the heritability was .60, the magnitude of shared environmental influences was .20, and the magnitude of nonshared environmental influences was .20. In Simulation 1, we chose to simulate data for a very large sample (10,000 individuals) to test whether the method of examining the underlying group deficits validly discriminated alternative comorbid models from each other when there was adequate power.

Simulation 2 was conducted to test the validity of examining underlying deficits of comorbid disorders to test alternative comborbid models when the underlying deficit score is not perfectly correlated with the symptom score. The model parameters for Simulation 2 were identical to Simulation 1 except that the underlying deficit score was not perfectly correlated with the symptom score. Again, we chose correlations between the underlying deficit score and the symptom score that approximated those found for ADHD and underlying deficits associated with ADHD (e.g., executive functioning; Nigg, Hinshaw, Carte, & Treuting, 1998) and for RD and underlying deficits associated with RD (e.g., phonological processes; Willcutt et al., 2001). The simulated correlations between the underlying deficit score and the symptom score approximated .35 for Disorder A (ADHD) and .60 for Disorder B (RD). Again, a very large sample size (10,000 individuals) was simulated to test whether the method of examining the underlying deficits of comorbid disorder validly discriminated alternative comborbid models from each other when there was adequate power.

Simulation 3 and Simulation 4 were conducted to examine whether there was enough power to discriminate between comborbid models with more reasonable sample sizes (i.e., a large but realistic sample size in Simulation 3 and a typical sample size in Simulation 4). The parameters simulated for Simulations 3 and 4 were identical to those of Simulation 2 except for the sample sizes. In studies examining the underlying deficits to test comborbid models, individuals with both disorders have often been oversampled to increase power. In Simulation 3, data for 20,000 individuals were simulated (using the same parameters as in Simulation 2); then, 350 individuals with neither disorder, 350 individuals with Disorder A only, 350 individuals with Disorder B only, and 150 individuals with both Disorders A and B were chosen randomly. In Simulation 4, we matched the number of individuals in each group to the sample sizes found in a recent study examining the underlying deficits of ADHD and RD (Willcutt et al., 2001). Data for 10,000 individuals were simulated (using the same parameters as in Simulation 2); then, 121 individuals with neither disorder, 52 individuals with Disorder A only, 93 individuals with Disorder B only, and 48 individuals with both Disorders A and B were chosen randomly.

**Analyses**

For each simulated dataset, we conducted t tests to determine whether the mean underlying deficit score of each pair of groups (e.g., individuals with neither disorder and individuals with A only) was significantly different. A group difference was considered to be present if there was a statistically significant group difference at the .05 level.

**Results**

**The Comparison of Level of Deficit in $R_A$ and $R_B$ in Individuals With A Only, B Only, and AB**

Several published studies have based their conclusions regarding the causes of comborbidty on two sets of results. First, the presence of double dissociation indicates that there are two separate sets of liabilities ($R_A$ and $R_B$) for Disorders A and B. Second, the pattern of deficits in the comborbid group depends on the causes of comborbidty between the two disorders and should help discriminate the correct model among several comborbid models. Therefore, the presence of double dissociation and the pattern of deficits in the comborbid group were considered when examining the simulation results (see Table 2).

Figure 1 shows a graphical representation of each comborbid model and lists the possible pathways for the diagnosis of A only, B only, and AB. It is important to consider the alternative pathways for the diagnosis of each disorder because these pathways should be directly related to the pattern of deficits in each group. For example, in the chance model, the diagnosis of A only is only possible if one is above the threshold on the liability distribution for $R_A$ and below the threshold on $R_B$ (see Figure 1). This is consistent with the simulation result that the individuals with A only had an elevated deficit in $R_A$ and did not have an elevated deficit in $R_B$ (see Table 2). Below are detailed descriptions of results for Simulation 1.

**Chance Model**

**Double dissociation.** There was clear double dissociation in the chance model. Individuals with A only did not differ from the controls on their $R_B$ scores, and individuals with B only did not differ from the controls on their $R_A$ scores. Individuals with A only had deficits in $R_A$ and did not have deficits in $R_B$. Individuals with B only had deficits in $R_B$ and did not have deficits in $R_A$. This result makes sense, given that the individuals with A only were above the threshold on $R_A$ and below the threshold on $R_B$ and that the individuals with B only were below the threshold on $R_A$ and above the threshold on $R_B$.

**Deficits in the comborbid group.** The comborbid group had deficits in both $R_A$ and $R_B$, given that this group had to be above the threshold on both $R_A$ and $R_B$ to manifest both disorders.
Testing Hypotheses Regarding Comorbidity

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results for Simulations 1 and 2: Mean Underlying Deficit Score for Each</td>
</tr>
<tr>
<td>Diagnostic Group and the Most Common Pattern of Results</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Comorbidity model</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Chance</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Alternate forms</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

**Note.** Italics indicate group comparisons that had effect sizes smaller than .20. C = individuals with neither Disorder A nor Disorder B; A only = individuals with Disorder A but not Disorder B; B only = individuals with Disorder B but not Disorder A; AB = individuals with both Disorders A and B.

Alternate Forms Model

**Double dissociation.** Given that there was only one liability distribution, $R$, there was no double dissociation.

Deficits in the comorbid group. The simulation results were similar to the predictions of Schachar and Tannock (1995), who asserted that both of the pure disorder groups and the comorbid group would perform equivalently on each measure of underlying deficit. Individuals with A only, B only, and AB had increased deficits in $R$ when compared with controls. Also, the score on $R$ for the comorbid group was not significantly greater than the score on $R$ for the A-only or B-only group.

Random Multiformity Model

**Double dissociation.** There was clear double dissociation in the random multiformity model. Individuals with A only did not differ from the controls on their $R_h$ scores, and individuals with B
only did not differ from the controls on their $R_A$ scores. This was because individuals with A only were above the threshold on $R_A$ and below the threshold on $R_B$, whereas individuals with B only were below the threshold on $R_A$ and above the threshold on $R_B$.

**Deficits in the comorbid group.** The comorbid group had the deficits of both the $R_A$ group and the $R_B$ group because the comorbid group included both members who were above the threshold on $R_A$ and members who were above the threshold on $R_B$. There are three separate pathways to comorbidity in the random multiformity model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance), (b) above the threshold on $R_A$ and below the threshold on $R_B$, and (c) below the threshold on $R_A$ and above the threshold on $R_B$. The level of deficit in $R_A$ was greater in the A-only group than in the AB group, given that the comorbid group included some members who were below the threshold on $R_B$ (i.e., the second pathway to comorbidity).

**Random Multiformity of A Model**

**Double dissociation.** There was clear double dissociation in the random multiformity of A model, given that the individuals with A only were above the threshold on $R_A$ and below the threshold on $R_B$ and that the individuals with B only were below the threshold on $R_A$ and above the threshold on $R_B$.

**Deficits in the comorbid group.** The comorbid group had the deficits of both the $R_A$ group and the $R_B$ group because the comorbid group included both members who were above the threshold on $R_A$ and members who were above the threshold on $R_B$. There are two separate pathways to comorbidity in the random multiformity of A model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance) and (b) above the threshold on $R_A$ and below the threshold on $R_B$. The level of deficit on $R_A$ was not significantly different between the A-only group and the AB group, given that all members in the AB group were above the threshold on $R_A$. In contrast, the level of deficit in $R_B$ was greater in the B-only group than in the AB group, given that the comorbid group included some members who were below the threshold on $R_B$ (i.e., the second pathway to comorbidity).

**Random Multiformity of B Model**

The results here were opposite those for the random multiformity of A model (i.e., the level of deficit in $R_B$ was greater in the A-only group than in the AB group, and the level of deficit in $R_B$ was not significantly different between the B-only group and the AB group).

**Extreme Multiformity Model**

**Double dissociation.** There was clear double dissociation in the extreme multiformity model, given that the individuals with A only were above the threshold on $R_A$ and below the threshold on $R_B$ (i.e., comorbid by chance) and that the individuals with B only were below the threshold on $R_A$ and above the threshold on $R_B$. There are three separate pathways to comorbidity in the extreme multiformity model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance), (b) above the threshold on $R_A$ and below the threshold on $R_B$, and (c) below the threshold on $R_A$ and above the threshold on $R_B$.

The level of deficit in $R_A$ was greater in the A-only group than in the AB group, given that the comorbid group included some members who were below the threshold on $R_A$ (i.e., the third pathway to comorbidity). Similarly, the level of deficit in $R_B$ was greater in the B-only group than in the AB group, given that the comorbid group included some members who were above the threshold on $R_B$ (i.e., the second pathway to comorbidity).

**Extreme Multiformity of A Model**

**Double dissociation.** There was clear double dissociation in the extreme multiformity of A model, given that the individuals with A only were above the threshold on $R_A$ and below the threshold on $R_B$ and the individuals with B only were below the threshold on $R_A$ and above the threshold on $R_B$.

**Deficits in the comorbid group.** The comorbid group had the deficits of both the $R_A$ group and the $R_B$ group because the comorbid group included both members who were above the threshold on $R_A$ and members who were above the threshold on $R_B$. There are two separate pathways to comorbidity in the extreme multiformity of A model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance) and (b) above the threshold on $R_A$ and below the threshold on $R_B$. The level of deficit in $R_A$ was greater in the AB group than in the A-only group. This was because the members of the AB group who were above the threshold on $R_A$ and below the threshold on $R_B$ (i.e., the second pathway to comorbidity) had to surpass a second, higher threshold on $R_A$ to manifest both disorders, AB, than to manifest A only. The level of deficit in $R_B$ was greater in the B-only group than in the AB group, given that the comorbid group included some members who were below the threshold on $R_B$ (i.e., the second pathway to comorbidity).

**Extreme Multiformity of B Model**

The results here were opposite those for the extreme multiformity of A model (i.e., the level of deficit in $R_B$ was greater in the AB group than in the B-only group, and the level of deficit in $R_A$ was greater in the A-only group than in the AB group).

**Three Independent Disorders Model**

**Double dissociation.** There was clear double dissociation in the three independent disorders model, given that the individuals with A only were above the threshold on $R_A$ and below the threshold on $R_B$ and that the individuals with B only were below the threshold on $R_A$ and above the threshold on $R_B$.

**Deficits in the comorbid group.** The comorbid group had the deficits of both the $R_A$ group and the $R_B$ group because the comorbid group included some members who were above the threshold on both $R_A$ and $R_B$. There are two separate pathways to comorbidity in the three independent disorders model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance) and (b) above the threshold on $R_{AB}$ (the liability distribution for the third, independent disorder). The level of deficit in $R_A$ was greater in the A-only group than in the AB group, given that the comorbid group
included some members who were below the threshold on $R_A$ (i.e., those who were above the threshold on $R_{AB}$, the second pathway to comorbidity). Similarly, the level of deficit in $R_B$ was greater in the B-only group than in the AB group, given that the comorbid group included some members who were below the threshold on $R_B$ (i.e., those who were above the threshold on $R_{AB}$, the second pathway to comorbidity).

**Correlated Liabilities Model**

**Double dissociation.** The double dissociation in the correlated liabilities model was not complete. Although the level of deficit in $R_A$ was greater in the A-only group than in the B-only group, it was greater in the B-only group than in the control group. Similarly, although the level of deficit in $R_B$ was greater in the B-only group than in the A-only group, it was greater in the A-only group than in the control group. This was because there was a significant correlation between $R_A$ and $R_B$. Some members in the B-only group might have had elevated scores on $R_A$, given the correlation between $R_A$ and $R_B$. In these individuals, the score on $R_B$ had to have crossed the threshold to manifest B, and although the score on $R_A$ was elevated, given the significant correlation between $R_A$ and $R_B$, the score on $R_A$ must not have been high enough to cross the threshold on $R_A$ to manifest A. Similarly, some members in the A-only group might have had elevated scores on $R_B$, given the correlation between $R_A$ and $R_B$. In these individuals, the score on $R_A$ had to have crossed the threshold to manifest A, and although the score on $R_B$ was elevated, given the significant correlation between $R_A$ and $R_B$, the score on $R_B$ must not have been high enough to cross the threshold on $R_B$ to manifest B.

**Deficits in the comorbid group.** The comorbid group had deficits in both $R_A$ and $R_B$, given that this group had to have been above the threshold on both $R_A$ and $R_B$ to manifest both disorders. The level of deficit in $R_A$ was higher in the AB group than in the A-only group. Given the significant correlation between $R_A$ and $R_B$, individuals with higher scores on $R_A$ were more likely to also have a higher score on $R_B$ and manifest both disorders (AB) than individuals with lower scores on $R_A$. Similarly, the level of deficits in $R_B$ was higher in the AB group than in the B-only group. Individuals with higher scores on $R_B$ were going to be more likely to also have a higher score on $R_B$ and manifest both disorders (AB) than individuals with lower scores on $R_B$.

**A Causes B Model**

**Double dissociation.** The double dissociation in the A causes B model was not complete. The level of deficit in $R_B$ was not elevated in the A-only group. In contrast, although the level of deficit in $R_A$ was greater in the A-only group than in the B-only group, it was greater in the B-only group than in the control group. The score on $R_A$ was elevated in the B-only group because Phenotype A directly affected Phenotype B. Some members in the B-only group were individuals whose high score on $R_A$ directly led to a high score on Phenotype B (and Disorder B) but whose high score on $R_A$ was not high enough to surpass the threshold to manifest Disorder A.

**Deficits in the comorbid group.** There are two separate pathways to comorbidity in the A causes B model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance) and (b) above the threshold on $R_A$ (with the high score on $R_A$ leading to above-threshold phenotypic score on B) and below the threshold on $R_B$.

The comorbid group had deficits in both $R_A$ and $R_B$ because the comorbid group that was comorbid via the first pathway had significant deficits in both $R_A$ and $R_B$.

The level of deficit in $R_A$ was greater in the AB group than in the A-only group. Given that Phenotype A directly caused Phenotype B, individuals who had a higher score on $R_A$ were more likely to also have a higher score on Phenotype B and manifest both disorders (AB) than individuals who had a lower score on $R_A$. The level of deficit in $R_B$ was greater in the B-only group than in the AB group because the AB group included some individuals who were below the threshold on $R_B$ (i.e., the second pathway to comorbidity).

**B Causes A Model**

The results here are opposite those for the A causes B model (i.e., the level of deficit in $R_B$ was greater in the AB group than in the B-only group, and the level of deficit in $R_A$ was greater in the A-only group than in the AB group).

**Reciprocal Causation Model**

**Double dissociation.** The double dissociation in the reciprocal causation model was not complete. The level of deficit in $R_A$ was greater in the A-only group than in the B-only group, but it was greater in the B-only group than in the control group. The score on $R_A$ was elevated for the B-only group because Phenotype A directly affected Phenotype B. Some members in the B-only group were individuals whose high score on $R_A$ directly led to a high score on Phenotype B (and Disorder B) but whose high score on $R_A$ was not high enough to surpass the threshold to manifest Disorder A. Similarly, the level of deficit in $R_B$ was greater in the B-only group than in the A-only group, but it was greater in the A-only group than in the control group. The score on $R_B$ was elevated for the A-only group because Phenotype B directly affected Phenotype A. Some members in the A-only group were individuals whose high score on $R_B$ directly led to a high score on Phenotype A (and Disorder A) but whose high score on $R_B$ was not high enough to surpass the threshold to manifest Disorder B.

**Deficits in the comorbid group.** The comorbid group had deficits in both $R_A$ and $R_B$ because the comorbid group included both members who were above the threshold on $R_A$ and members who were above the threshold on $R_B$. There are four separate pathways to comorbidity in the reciprocal causation model: (a) above the threshold on $R_A$ and $R_B$ (i.e., comorbid by chance), (b) above the threshold on $R_A$ (with the high score on $R_A$ leading to above-threshold score on B) and below the threshold on $R_B$, (c) above the threshold on $R_B$ (with the high score on $R_A$ leading to above-threshold score on A) and below the threshold on $R_A$, and (d) feedback loop starting below threshold on $R_A$ and $R_B$ but finding equilibrium with both A and B above threshold. The level of deficit in $R_A$ was greater in the A-only group than in the AB group because the AB group included some individuals who were below the threshold on $R_A$ (i.e., the third pathway to comorbidity). Similarly, the level of deficit in $R_B$ was greater in the B-only group than in the AB group because the AB group included some individuals who were below the threshold on $R_B$ (i.e., the second pathway to comorbidity).

**Effect Sizes**

Given that the $t$ test is sensitive to sample size, we examined the effect sizes of the group differences in the means to determine
whether a very small effect was statistically significant only because of the large sample size in the simulations. The effect size for group differences in the mean, $d$, is calculated by taking the difference between the means and dividing by the standard deviation (Cohen, 1988). Because the underlying deficit scores are standardized scores, the effect size is simply the difference between the two group means. For example, the effect size for the first comparison shown in the last column of Table 2 is the difference between the mean of the AB (2.11) group and the A-only group (2.11), or 0.00. Cohen (1988) described a small effect size as a $d$ of at least .20, a medium effect size as a $d$ of at least .50, and a large effect size as a $d$ of at least .80. For all group comparisons that did not meet statistical significance, the effect size did not surpass the criterion for a small effect size. Most of the group comparisons that were statistically significant surpassed the criterion for a small effect size. Those group comparisons that had effect sizes smaller than a $d$ of .20 are italicized in Table 2. For these group comparisons, the effect sizes ranged from .13 to .17.

**Discrimination Between Comorbidity Models**

Table 2 shows the results for Simulations 1 and 2. (A similar table for Simulations 3 and 4 is not shown because there were so many mistakes in discrimination of the correct comorbidity model in Simulations 3 and 4 and often there was no typical pattern of result for each comorbidity model.) For each comorbidity model, the mean underlying deficit score for each diagnostic group ($C = \text{individuals with neither disorder}$, $A = \text{individuals with Disorder A only}$, $B = \text{individuals with Disorder B only}$, and $AB = \text{individuals with both Disorders A and B}$) is presented. Figure 2 also shows the mean underlying deficit scores for Disorder A and Disorder B for each diagnostic group for all comorbidity models (except the alternate forms model, in which there is only one liability distribution). Given limited space and the large amount of information, the pattern of results for each simulation cannot be shown here but can be found on the following Web site: http://ibgwww.colorado.edu/cadd/software (the results for each simulation are shown for Simulations 3 and 4 as well). In Table 2, the most common pattern of results is shown. For example, for the chance model, the most common pattern of results for underlying deficit for Disorder A was $AB = A$ only $> C = B$ only. This means that when the chance model was simulated, individuals with $AB$ or $A$ only had a significantly higher level of the underlying deficit for Disorder A than individuals with neither disorder or $B$ only. Note that there are some differences in the most common pattern of results in Simulation 1 and Simulation 2.

An examination of Table 2 suggests that a unique pattern of results was found for nearly all comorbidity models. The only exception is that an identical pattern of results was found for three comorbidity models: random multifinality, extreme multifinality, and three independent disorders. Unfortunately, the most common pattern of results shown in Table 2 was not found in each of the 50 simulations. Therefore, we compared the pattern of results found in each of the 50 simulations for every comorbidity disorder with the pattern of results found in each of the 50 simulations for every other comorbidity disorder to determine whether the pattern of results in each simulation was unique to the particular comorbidity model being tested.

Table 3 shows the results of that comparison for Simulations 1–4. For each comorbidity model, the number of times a unique pattern of results was not found (i.e., the correct comorbidity model could not be distinguished from all 12 other comorbidity models) and type of mistakes in discrimination are listed. For example, the first line indicates that in Simulation 1, for the chance model, in 3 simulations out of 50 simulations, the pattern of results was not unique to the chance model but found in data simulated for another comorbidity model. In 2 simulations, a pattern of results also found in data simulated for the random multifinality of A model was found, and in 1 simulation, a pattern of results also found in data simulated for the random multifinality of B model was found. The pattern of results found in 47 out of 50 simulations was found only in data simulated for the chance model and no other comorbidity model. In Table 3, if more than one comorbidity model is mentioned in the column titled Mistakes in discrimination between comorbidity models, this means that the same pattern of results was found in data simulated for several comorbidity models.

As suggested in Table 2, the information in Table 3 shows that the random multifinality model, the extreme multifinality model, and the three independent disorders model cannot be distinguished from each other by examining the underlying deficits of comorbid disorders. There were few other mistakes in discrimination in Simulation 1, with data simulated for the chance model being mistaken for the random multifinality of A or random multifinality of B model and some of the multifinality models being mistaken for another type of multifinality model. There were more mistakes in discrimination in Simulations 2 and 3 than in Simulation 1. Most of the mistakes were discriminations between the six multifinality models. In addition, the discrimination between the B causes A model and the reciprocal causation model was much worse. In Simulation 4, where sample sizes identical to those found in a recently published study (Willcutt et al., 2001) were simulated, the discrimination between comorbidity models was poor.

**Discussion**

In the present study, the validity of examining underlying deficits in comorbid disorders as a way to test alternative hypotheses of causes of comorbidity was tested. We compared the mean underlying deficit scores in individuals with neither disorder, A only, B only, and both A and B in data simulated for each of the 13 Neale and Kendler (1995) comorbidity models.

Our first goal was to test the assumptions of previous researchers using this method to assess alternative comorbidity models. In general, our simulation results validate the past studies’ predictions, although the simulation results were slightly different from the predicted results (see *The Comparison of Level of Deficit in $R_A$ and $R_B$ in Individuals With A Only, B Only, and AB section in Results for detailed descriptions*). For example, several researchers (e.g., Pennington et al., 1993; Schachar & Tannock, 1995) predicted that if the correlated liabilities model is the correct comorbidity model, the comorbid group should have the deficits of both pure groups. Our simulation results indicate that actually, the comorbid group had a higher level of deficits of both Disorders A and B than any other diagnostic group.

Our second goal was to resolve the discrepancy in the literature regarding the predicted pattern of results for the three independent disorders model. Purvis and Tannock (2000) and Pennington et al. (1993) predicted that the comorbid group would exhibit deficits of

(text continues on p. 361)
Figure 2. Results for Simulations 1 and 2: Mean underlying deficit score. C = individuals with neither Disorder A nor B; A only = individuals with Disorder A only; B only = individuals with Disorder B only; AB = individuals with both Disorders A and B.
Figure 2. (continued)
Figure 2. (continued)
Table 3
Results for Simulations 1–4: Type of Discrimination Mistakes (Number of Mistakes)

<table>
<thead>
<tr>
<th>Comorbidity model</th>
<th>Simulation 1: Type (no. of mistakes)</th>
<th>Simulation 2: Type (no. of mistakes)</th>
<th>Simulation 3: Type (no. of mistakes)</th>
<th>Simulation 4: Type (no. of mistakes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>RMA (2) RMB (1) BA/RC (1)</td>
<td>RMA (3) RMB (2) EMA (1) BA/RC (2)</td>
<td>RMA (2) RMB (1) AB (1) EM/AB/RC (1)</td>
<td>CL/BA/RC (2)</td>
</tr>
<tr>
<td>AF</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>CH (1) RMA (19) RMB (1)</td>
</tr>
<tr>
<td>RM</td>
<td>EM/TD (50)</td>
<td>EM/TD (50)</td>
<td>EM/TD (48)</td>
<td>BA/RC (2) EM/AB/RC (1) RMB (1) EM/TD (2)</td>
</tr>
<tr>
<td>RMA</td>
<td>EMA (1) RM/EM/TD (2) TD (1)</td>
<td>EMA (32) RC/EM/RC (2)</td>
<td>EMA (1) RM/EMA/TD (36) EM/TD (5)</td>
<td>BA/RC (1) EM/AB/RC (35) RMA (3)</td>
</tr>
<tr>
<td>RMB</td>
<td>EMB (2) RM/EM/TD (1) BA/RC (1)</td>
<td>EMB (35) RM/EM/TD (2)</td>
<td>CH (1) RM/EM/RC (2)</td>
<td>BA/RC (1) EM/AB/RC (2)</td>
</tr>
<tr>
<td>EM</td>
<td>RM/TD (50)</td>
<td>CH (2) RMA (24) RMB (3) RC (1)</td>
<td>RMA (20) RMB (1) EM/AB/RC (2) RM/EM/TD (1)</td>
<td>CH/AB/RC (1) RM/RMB/EM/RC (7)</td>
</tr>
<tr>
<td>EMA</td>
<td>None</td>
<td>RMA (19)</td>
<td>RMA (40) RMA/RC (1)</td>
<td>RMA (28) RMA (1) RM/EM/TD (10)</td>
</tr>
<tr>
<td>EMB</td>
<td>RMB (4) RM/EM/TD (2) RC (1)</td>
<td>RMB (39) BA (1)</td>
<td>RMA (10) RMA/RC (1)</td>
<td>RMA/EM/RC (1) RMA/RMB/EM/RC (7) RMB/AB/RC (2)</td>
</tr>
<tr>
<td>TD</td>
<td>RM/EM (50)</td>
<td>RMA (2) RM/EM/TD (24) RC (3)</td>
<td>RMA (39) RM/EM (28) RC (1)</td>
<td>RM/RMB/EM/RC (10) RMA/EM/RC (1) RMB/EM/RC (1)</td>
</tr>
<tr>
<td>CL</td>
<td>None</td>
<td>RC (1)</td>
<td>RC (1)</td>
<td>BA (1) RMA (10) CH/EM/RM/RC (1) CH/CL/RC (20)</td>
</tr>
<tr>
<td>AB</td>
<td>None</td>
<td>RC (2)</td>
<td>EM/RC (7)</td>
<td>CH/BA/RC (7) RMA (12)</td>
</tr>
<tr>
<td>BA</td>
<td>None</td>
<td>CH/RB (1) RC/RM (8)</td>
<td>EMB (2) CH/RB (2) RC (21)</td>
<td>EMB (2) RMA (1) RM (1) CH/EM/RC (2)</td>
</tr>
<tr>
<td>RC</td>
<td>None</td>
<td>CH (1) RMA (2) EM (20) RM/EM/TD (1)</td>
<td>CH (2) RMA (10) RMA/EM/RC (9) RC (21)</td>
<td>CH (23)</td>
</tr>
</tbody>
</table>

Note. CH = chance; AF = alternate forms; RM = random multiformity; RMA = random multiformity of A; RMB = random multiformity of B; EM = extreme multiformity; EMA = extreme multiformity of A; EMB = extreme multiformity of B; TD = three independent disorders; CL = correlated liabilities; AB = A causes B; BA = B causes A; RC = reciprocal causation.
both pure groups, whereas Leung et al. (1996), Manassis et al. (2000), Schachar and Tannock (1995), and Lynam (1996) predicted that the comorbid group should be different in the pattern of deficits shown from the two pure groups. The results from the simulations show that the three independent disorders model described by Neale and Kendler (1995) is closer to the comorbidity model described by Leung et al. (1996), Schachar and Tannock (1995), and Lynam (1996) than the comorbidity model described by Purvis and Tannock (2000) and Pennington et al. (1993). In the data simulated for the three independent disorders model, the A-only group had the deficits of Disorder A and did not have the deficits of Disorder B, whereas the B-only group had the deficits of Disorder B and did not have the deficits of Disorder A. Although the comorbid group had the deficits of both disorders (in that the comorbid group had a higher level of deficits of both A and B than the control group), the level of deficits of the comorbid group was different than the level of deficits of the two pure groups in that the A-only group had a significantly higher level of deficit of Disorder A than the comorbid group and the B-only group had a significantly higher level of deficit of Disorder B than the comorbid group.

The third goal of our study was to test whether additional comorbidity models could be distinguished from each other using this method. Simulation results in which the correlation between the underlying deficit scores and the symptom scores was perfect and the power was adequate suggested that there was a unique pattern of results for 10 of the 13 comorbidity models. In contrast, the performance of results for 10 of the 13 comorbidity models. In contrast, the patterns of results found in data simulated for the random multiformity model, the extreme multiformity model, and the three independent disorders model were identical. Although it is not possible to distinguish the random multiformity/extreme multiformity models and the three independent disorders model using the criterion of a distinct pattern of statistically significant group differences, it may be possible to distinguish between these models by examining the mean scores (see Figure 2). Although the difference between the AB group and the control group was significantly different for all three models in Simulation 1, this difference was much smaller in the three independent disorders model ($d = .30 – .32$ in Simulation 1, and $d = .11 – .16$ in Simulation 2) than in the random multiformity ($d = 1.25$ in Simulation 1, and $d = .41 – .71$ in Simulation 2) or extreme multiformity models ($d = 1.54 – 1.58$ in Simulation 1, and $d = .56 – .89$ in Simulation 2).

The results from Simulation 1 suggest that this method appears to be useful in extremely large samples in which the symptoms of each disorder are perfectly correlated with the underlying deficit score. In contrast, the results from Simulations 2–4 suggest that discrimination between comorbidity models becomes worse as the correlation between the measured underlying deficit and the symptom score decreases and the sample size decreases. The results for Simulation 4, which was simulated to approximate the model parameters, correlation between the underlying deficit and the symptom score, and sample sizes found in a recent publication (Willcutt et al., 2001), show that discrimination between comorbidity models was poor. These results indicate the importance of the consideration of power and suggest the importance of conducting simulation studies and power calculations.

Limitations

We used data simulated under the assumptions of the comorbidity models described by Neale and Kendler (1995) because they have described the most comprehensive set of alternative comorbidity models to date. However, Neale and Kendler have not considered all possible alternatives for causes of comorbidity. These alternatives include more complex causal models (e.g., Carey & DiLalla, 1994; Klein & Schwartz, 2002) and more complex comorbidity models in which formal modeling may be very difficult, such as a multiple deficit model in which two or more cognitive deficits (each with its own liability distribution) interact to produce a particular disorder. Also, all of the Neale and Kendler models assume multifactorial causes and do not consider other possible etiological models (e.g., a single major gene).

Second, all methods examining the causes of comorbidity, including the examination of underlying deficits, can be influenced by errors in comorbidity rates. Factors affecting comorbidity rates include underascertainment due to younger participants not being through the risk period, measurement method, nonrandom refusal by probands and relatives, time frame, severity criteria, range of disorders considered, and diagnostic method used (e.g., Clark, Watson, & Reynolds, 1995; Klein, 2003).

Third, a limitation that applies to most of the methods of examining comorbidity models (e.g., Rhee, Hewitt, Corley, & Stallings, 2003; Rhee et al., 2004), including the examination of underlying deficits, is the fact that they examine comorbidity between two disorders although comorbidity among mental disorders usually exists among more disorders. New methods that examine comorbidity among three or more disorders need to be developed.

Conclusions and Implications for Future Studies

In summary, our results suggest that under ideal conditions (i.e., a situation in which the $N$ for each group is very large and the correlation between the underlying deficits and the symptom score is perfect), examination of underlying deficits of comorbid disorders can be very useful in discriminating the correct comorbidity model from alternative hypotheses. This method could be especially powerful in combination with the Neale and Kendler (1995) model fitting of family data, given that the two methods have different strengths. Rhee et al. (2004) showed that the Neale and Kendler model-fitting approach can be used to analyze family data and can reliably discriminate between the alternate forms model, the multiformity models (i.e., the three random multiformity models and the three extreme multiformity models), the three independent disorders model, and the correlated liabilities models (correlated liabilities, A causes B, B causes A, and reciprocal causation), whereas the discrimination within the six multiformity models and the four correlated liabilities models was much worse. In contrast, with the examination of the underlying group deficits, the discrimination between random multiformity, extreme multiformity, and three independent disorders is not possible, but the four correlated liabilities models can be distinguished from each other.

In contrast to the results in an optimal situation, however, results of simulations with more realistic estimates of sample size and correlations between the underlying deficit scores and the symptoms of each disorder are more sobering. In simulations in which the correlations between the underlying deficits and the psychiatric disorder were not perfect and the sample sizes were similar to those found in the literature, there was not enough power to distinguish several of the alternative comorbidity models from each other consistently.
Most previous studies have examined only a few models at a time and have not considered the possibility that although a certain pattern of results is consistent with a particular comorbidity model, the same pattern of results also may be consistent with a number of other comorbidity models. The simulation results suggest that this is a valid concern. For example, several previous studies specifically attempted to test the three independent disorders model (e.g., Leung et al., 1996; Schachar & Tannock, 1995), whereas none considered the multifactorial models. Yet, the simulation results show that the pattern of results found for the three independent disorders model is difficult to distinguish from that of the random multifactorial model or the extreme multifactorial model.

In conclusion, our results suggest that this approach may be a useful method to test different hypotheses regarding the etiology of comorbidity but also indicate that it is essential for future studies to consider carefully the implications of the magnitude of the correlation between the underlying deficits and the psychiatric disorder symptoms and the available sample size in each group. These results underscore the importance of adequate statistical power for this method to be useful and suggest that simulation studies provide a useful technique to obtain the necessary calculations.

References

Received April 14, 2004
Revision received November 1, 2004
Accepted November 29, 2004