Applications of Taxometric Methods to Problems of Comorbidity: Perspectives and Challenges
Irwin D. Waldman and Scott O. Lilienfeld
Emory University

We share Meehl's view that taxometric procedures hold considerable promise for elucidating questions regarding psychiatric comorbidity. Drawing on examples from the domain of childhood disruptive disorders, we discuss why the issues raised by Meehl are scientifically and pragmatically important and outline several profitable applications of taxometric methods to questions of comorbidity (e.g., estimating the statistical relations between latent taxa). We explain why taxometric methods and other sophisticated latent variable methods are needed to answer such questions and provide examples of how certain statistical methods have been used to make erroneous inferences regarding taxonicity. Several important unresolved issues bearing on the use of taxometric procedures and their application to questions of comorbidity are delineated, including (a) the distributional assumptions of taxometric methods, (b) the construct validation of provisional taxa identified by taxometric analyses, (c) the relation of taxometric methods to other latent variable techniques (e.g., latent class analysis), (d) the potential existence of spurious taxa, (e) the question of “fuzzy taxonicity,” and (f) “configural taxa.” We conclude with a discussion of analytic methods for characterizing and understanding the covariation between latent dimensions as opposed to taxa.

Key words: comorbidity, psychopathology, classification, diagnosis, taxon. [Clin Psychol Sci Pract 8:520–527, 2001]
considered the latter usage at times). Meehl did not, however, distinguish explicitly between these two uses of the term comorbidity. As we have pointed out elsewhere (Lilienfeld et al., 1994), this distinction is critical, as different conceptual questions typically arise for the use of the term comorbidity for individuals (e.g., does a patient present with two separable but overlapping disorders or with two constellations of symptoms that reflect a single diagnostic entity?) as opposed to populations (e.g., do two putatively different disorders truly reflect distinct latent taxa? If so, what are the boundaries and relations between them?).

WHY SHOULD WE CARE?

Some readers might wonder why the issues addressed by Meehl are scientifically and pragmatically important. They might believe that questions of nosology are largely passe—that many or most of the important nosological issues in descriptive psychopathology already have been resolved. This conclusion would represent a serious misconception. In our view, it is disconcerting that more than two decades since the publication of DSM-III (American Psychiatric Association, 1980), which many consider the first serious empirically based psychiatric classification system, we still do not know with any degree of certainty the answers to many of the most fundamental questions regarding psychiatric classification and diagnosis.

Some examples of unresolved taxonomic issues from the domain of childhood disruptive behavior disorders should help clarify this point. Despite a plenitude of studies on these conditions over the past two decades, we still do not know whether the underlying structure of attention-deficit hyperactivity disorder (ADHD), for example, is taxonic or dimensional. We do not know whether, in nature, there truly exist three ADHD subtypes (viz., the combined, hyperactive-impulsive, and inattentive) or the two underlying dimensions (viz., hyperactivity-impulsivity and inattention) described in DSM-IV (American Psychiatric Association, 1994). Are the ADHD subtypes truly distinct taxa (or, to use Meehl’s phrase, “subtaxa”) representing a collection of conditions each with its own distinct pathology and etiology, or do they instead represent surface variations of a unitary diagnostic entity? Such taxonomic conundrums in this domain are not limited to ADHD, but also extend to oppositional defiant disorder (formerly called oppositional disorder) and conduct disorder (ODD and CD, respectively). We do not know, for example, whether either or both ODD and CD are taxonic, and if so, whether these taxa are distinct from each other. Alternatively, is ODD merely a developmental precursor to CD that reflects the same latent taxon or symptom dimension as CD?

WHY TAXOMETRIC PROCEDURES ARE NEEDED: INCONCLUSIVE EFFORTS TO INFER TAXONICITY

The situation we have described would be troublesome enough if child psychopathology researchers had not attempted to address such issues. In the two decades following the publication of DSM-III, numerous studies purporting to resolve these taxonomic issues have been published. Nevertheless, most of these studies have failed in their primary mission due to methodological limitations. These studies have used a variety of statistical methods in fruitless efforts to resolve the dimensional versus taxonic nature of disruptive disorders. Perhaps the simplest of these approaches is the use of analysis of variance (ANOVA) designs to detect mean differences among diagnostic groups on relevant variables. For example, Rey et al. (1988) used this strategy to show that CD children exhibited worse outcomes than ODD children. But the authors were mixed in their interpretation of these findings, in some places arguing that ODD is merely a less severe form of CD, and in other places arguing that ODD is a unique diagnostic entity distinct from CD. Indeed, the very title of their article, “Oppositional Disorder: Fact or Fiction,” implies that ANOVA designs can be used to settle the question of whether a condition, in this case ODD, is qualitatively distinct from other conditions. Using a similar approach, Reeves, Werry, Elkind, and Zametkin (1987) and Werry, Reeves, and Elkind (1987) attempted to distinguish DSM-III ODD from CD, attention-deficit disorder with hyperactivity (ADDH), and anxiety disorders on a variety of background characteristics. These studies yielded mixed findings, with the groups differing on some measures and not on others, leading the authors to reach equally mixed conclusions regarding the distinctiveness of ODD, CD, and ADDH. The findings of these studies could be explained equally well by a qualitative difference model as by a dimensional model in which ODD and CD are expressions of the same underlying disposition but differ in severity. Simply put, ANOVA designs cannot be used to resolve questions of taxonicity, as even large mean differences may merely reflect substantial differences along one or more dimensions.
A second and more sophisticated approach to such questions involves the use of confirmatory factor analysis (CFA) in an effort to test alternative a priori models of the dimensional structure presumably underlying measures of a disorder. This method has been applied recently to ADHD (e.g., Beiser, Dion, & Gotowiec, 2000; Pillow, Pelham, Hoza, Molina, & Stultz, 1998), and most of these CFAs have demonstrated that a model positing two correlated dimensions corresponding to hyperactive-impulsive and inattentive symptoms fits better than a unidimensional model. Although these CFAs have been quite useful in elucidating the latent dimensional structure of ADHD symptoms, they in no way demonstrate that the latent structure of ADHD is dimensional. Just as cluster analyses will typically yield categories regardless of the true latent structure of a disorder, CFA will yield dimensions regardless of the true latent structure. A complementary procedure to CFA is latent class analysis (LCA), which is useful for testing alternative a priori models of the categorical structure underlying a disorder and (at least in theory) for testing whether the latent structure is taxonic or dimensional. A number of recent studies have applied LCA to ADHD in large data sets (e.g., Hudziak et al., 1998; Neumann et al., 1999), but their results do not, in our view, provide conclusive answers to the question of whether ADHD consists of a set of taxa or of latent dimensions.

A third approach to this taxonomic question involves behavior genetic analyses of data on monozygotic and dizygotic twins. Although the details of these methods are beyond the scope of our commentary, these analyses use multiple regression techniques to contrast the heritability of a continuous trait with the heritability of a particular symptom dimension (i.e., accounting for 91% vs. 75% of the variance in ADHD liability, respectively), this difference in heritabilities was not statistically significant, leading the authors to conclude that ADHD was best construed as a continuum rather than as a category (Levy, Hay, McLaughlin, Wood, & Waldman, 1997). Nevertheless, these analyses did not and cannot conclusively resolve the question of whether ADHD is a category or continuum.

Similar conclusions can be drawn about the application of this analytic method to other areas of childhood psychopathology, such as depression (Rende, Plomin, Reiss, & Hetherington, 1993) and internalizing and externalizing problems (Deater-Deckard, Reiss, Hetherington, & Plomin, 1997). This is because the analytic procedure in question contrasts the magnitude of the genetic influences, but does not address whether the same or different genes are involved in the categorical diagnosis as in continuous symptom levels. For example, even if the group heritability is significantly higher than the heritability for the continuous trait, this finding may only indicate higher heritability with increasing symptom severity. That is, the same genes may be involved in the etiology of the diagnosis as for continuous symptom levels, but they may exert a greater effect at higher symptom levels. Alternatively, the categorical diagnosis of ADHD and continuous ADHD symptom levels may be influenced by different sets of genes, but the magnitude of effects of these different sets of genes may be similar. Thus, this behavior genetic analytic method is no better at resolving whether conditions are taxonic or dimensional than the analytic methods discussed earlier. Other behavior genetic analytic techniques (see Neale & Kendler, 1995) may be better suited to resolving questions of taxonicity and comorbidity.

USEFUL APPLICATIONS OF TAXOMETRIC PROCEDURES TO QUESTIONS OF COMORBIDITY

If one grants the proposition that most of the important nosological questions in psychopathology research remain to be resolved, what are some useful applications of taxometric procedures to questions of comorbidity?

The first, described by Meehl in this issue and in many other taxometric articles, is resolving whether a diagnostic entity is better viewed as a latent taxon or a latent dimension and, in the case of more than one such condition, resolving this issue for each. A corollary of this issue is the question of whether two or more superficially distinct disorders are truly alternative manifestations of a single underlying condition (i.e., a unitary latent taxon), or whether they represent bonafide distinct latent taxa. For example, are antisocial personality disorder and somatization disorder merely sex-differentiated manifestations of
the same latent taxon, or do they represent qualitatively different taxa (e.g., see Cloninger, 1978; Lilienfeld, 1992)? Or could these two conditions merely represent somewhat different densification points produced by the confluence of several underlying dimensions (e.g., impulsivity, attention seeking)?

Second, once several distinct latent taxa have been established, it is important to estimate the extent of their co-occurrence in different populations (e.g., clinically referred vs. general population, males vs. females). It would be useful to know whether such taxa show minimal or substantial co-occurrence once they are properly distinguished via the use of taxometric methods, as such information would be helpful for estimating risk for additional disorders in individuals with a given disorder.

A third application of taxometric procedures is to clarify diagnostic conditions as a prerequisite to investigations of their course and prognosis (Robins & Guze, 1970), pathology and etiology (Meehl & Golden, 1982), and treatment response. It is easy to imagine how such investigations are handicapped by misclassification, as well as how the etiology, course, and treatment response of a diagnostic entity could be clarified considerably by an accurate representation of its latent status.

Fourth, there is the example that Meehl treats extensively: the application of taxometric analyses to understand better the specific disorders with which a patient presents, and the implications of comorbid disorders for the course, prognosis, and treatment of the “primary” disorder. Despite the promise of these potential applications, they represent ideals that have yet to be applied to most psychopathological conditions.

UNRESOLVED ISSUES IN TAXOMETRIC ANALYSIS

A number of unresolved issues in taxometric analysis are pertinent to questions regarding psychopathological comorbidity. Although we wholeheartedly agree with Meehl that taxometric analyses hold considerable promise for resolving questions of comorbidity, their proper application to such questions may rest on the resolution of a number of important methodological and conceptual issues.

Distributional Assumptions of Taxometric Methods

What are the distributional assumptions, if any, of the indicator variables to allow for a valid taxometric analysis? Most measures of psychopathology (e.g., symptom ratings, laboratory measures) possess undesirable distributional properties, often manifesting considerable skewness and kurtosis, and frequently exhibiting an inverse J-shaped or L-shaped distribution. In factor analyses, these distributional properties can result in method or difficulty factors (“artifacts”) if they characterize many of the indicator variables (McDonald, 1965), thus raising the question of whether they can similarly bias the results of taxometric analyses and lead to spurious inferences of taxonicity. Although there has been some work exploring the effects of the skewness of the indicator variables on taxometric results (N. G. Waller, personal communication, February 2, 2001), the sensitivity of the performance of taxometric analyses to such distributional properties merits further examination.

Construct Validation of Provisional Taxa Identified by Taxometric Analyses

How can one best validate the results of taxometric analyses? In our view, taxometric analyses are similar to other procedures (e.g., other latent variable models) for establishing the internal validity of constructs (see Skinner, 1981) in that the latent taxa uncovered by such methods possess a provisional status. To establish more firmly the validity of such constructs, the accumulation of evidence for external validity is also required. This evidence could constitute findings regarding pathology, etiology, course and prognosis, and treatment response. It is important to recognize that this process of validation with external variables does not proceed unidirectionally from putative diagnostic entities (viz., the latent taxa) to the external validators, as if the latter were criteria representing “gold standards” (see Faroane & Tsuang, 1994). Rather, this process of external validation is likely to proceed in an iterative fashion, with the initial relations between the latent taxa and the validators used to refine the indicator set for the latent taxa, and the refined latent taxa compared again with the external validators. It is through such a cyclical and self-correcting process, long ago outlined by Cronbach and Meehl (1955), that the construct validity of latent taxa can be most firmly established (see also Loevinger, 1957; Tellegen & Waller, 1994, for discussions of this strategy in the development of assessment instruments).

Another useful procedure for establishing the validity of taxometric procedures for revealing true latent taxa would be to select examples of known taxonic diseases
along with relevant indicators from the medical literature and determine how well they are recaptured by taxometric methods. Taxometric analyses of well-chosen indicators of such single-gene disorders as Huntington’s disease on established cases and noncases would provide a strong test case for the successful application of taxometric methods to real-world classification problems. Conversely, it would be useful to subject taxometric procedures to presumed dimensional disorders in the medical domain (e.g., type II diabetes) to verify that these procedures reject the hypothesis of latent taxa and correctly detect latent dimensions. Nevertheless, the latter situation may be more epistemically ambiguous, as some disorders previously assumed to be entirely dimensional might encompass certain subtypes that are due to a heretofore undetected dichotomous causal agent.

Relation of Taxometric Methods to Other Latent Variable Techniques
How do taxometric methods relate to other latent variable models that have been used in classification research to address problems of comorbidity? For example, what are the similarities and differences, or relations and boundaries, between taxometric methods and other latent variable models such as latent class analysis (Lazarsfeld & Henry, 1968), admixture analysis (e.g., Cloninger, Sigvardsson, von Knorring, & Bohman, 1984), and cluster analysis (e.g., Blashfield, 1984)? Systematic comparisons of these analytic procedures on both simulated and real psychiatric data would prove illuminating but, to our knowledge, have only very recently been carried out (Cleland, Rothschild, & Haslam, 2000). In an extension of the idea raised earlier, it also would be useful to compare the results of these procedures in their application to recovering Huntington’s disease and other known taxa in the medical domain. Such a latent variable olympiad would provide a better understanding of the properties of different latent variable methods, in the especially useful context of recapturing known real-world taxa.

Spurious Taxa
How effectively can taxometric methods distinguish true taxa from “institutional pseudo-taxa” (Grove, 1991; see also Meehl & Golden, 1982) or other spurious taxa (see Cattell, 1946)? Numerous situations might produce the false appearance of taxonicity among variables that are actually dimensional at an underlying level. Sample selection and referral factors are prime suspects in creating such scenarios. For example, being a psychiatric inpatient, an incarcerated felon, or an NBA basketball player all result partly from selection and referral processes that involve surpassing extreme thresholds on one or more underlying dimensions. Taxometric procedures will be most useful to the extent that they can distinguish genuine taxa from such “institutional pseudo-taxa.” The only investigation of which we are aware along these lines is a study by Grove (1991) that attempted to distinguish true taxa from institutional pseudo-taxa using cluster analyses with a variety of stopping rules. No one method clearly emerged as superior to the others or as capable of making this distinction reliably; rather, each method appeared to be generally liberal or conservative regardless of the presence or absence of genuine taxa. Further research is needed to understand the capacity of taxometric methods per se (as well as other latent variable methods) to distinguish true taxa from “pseudo-taxa” or other spurious taxa, as well as to determine the scenarios under which taxometric methods perform well in this context from those in which they perform poorly.

Fuzzy Taxonicity
A fifth unresolved issue involves the issue of “fuzzy taxonicity.” Medical genetic studies have revealed that there is considerable variability in the expression of diseases, even those that are due to the effects of a single gene. For example, in the case of both Huntington’s disease (as Meehl noted in this issue) and fragile X syndrome, the normal action of the respective genes is disrupted by mutations that involve an excess number of basepair repeats. The number of these basepair repeats varies among individuals in the population and results in the fully manifested disease only when the number surpasses a threshold. An active area of research on fragile X involves examining the effects of subthreshold numbers of basepair repeats on more subtle phenotypic manifestations. Some researchers have found, for example, that individuals with an excess number of repeats, but who are below the disease threshold, still show mild phenotypic impairment (e.g., Murray et al., 1996; but see Crawford et al., 1999). These findings may pose a problem for certain taxometric methods, in that although the disorder is truly taxonic, less severe manifestations of the etiological process may perturb the inference of taxonicity. Further study of taxometric methods as applied to these real-world cases of
“fuzzy taxonicity” is warranted to examine whether these methods can correctly detect disease taxa in the presence of such gray cases.

**Configural Taxonicity**

A final unresolved issue for taxometric analysis involves what might be termed “configural taxonicity,” the case in which a taxon may be viewed as emergent from the conjunction of extremity on a number of underlying dimensions. This is somewhat akin to the case of institutional pseudo-taxa mentioned earlier, except that here the taxonic structure is not an artifact of selection processes, but rather exists in nature. The central question is whether a taxon can arise from the presence of individuals whose extreme status on a number of dimensions creates an essentially qualitative difference between them and the rest of the population (e.g., see Cloninger’s model of personality disorders; Cloninger, 1987). For example, psychopathy may be a taxon (Harris, Rice, & Quinsey, 1994; but see Lilienfeld, 1998) arising from the conjunction of extremity on several traits such as superficial charm, guiltlessness, and failure to show normal avoidant conditioning to punishment cues (Lykken, 1995; Newman & Kosson, 1986). Although each of these traits may be distributed continuously in the population, a natural group may exist that is composed of individuals who are extreme on all of these dimensions. In this hypothetical scenario, individuals who are extreme on only a subset of these dimensions would not be perceived as psychopathic. This possibility raises two crucial taxometric questions: first, whether such a scenario could justifiably be called taxonic, and, if so, whether taxometric methods could accurately detect such a taxon (see Meehl & Golden, 1982, for a further discussion of different forms of taxonicity).

**Alternatives for Understanding the Covariation Among Latent Dimensions and Its Causes**

Up to this point, we have focused exclusively on issues involving the detection of taxonicity. It is worth noting, however, that many or most forms of psychopathology, even many Axis I conditions such as mood and anxiety disorders, may be dimensional rather than taxonic. In this case, as Meehl observed (this issue; see also Lilienfeld et al., 1994), the use of the term and concept of comorbidity will typically be inappropriate and misleading.

Fortunately, a number of well-developed analytic methods exist for identifying these latent dimensions, characterizing the covariation among them, and understanding their shared and unique etiologies. For example, CFA has begun to see widespread use in many domains of psychopathology as a means of clarifying the latent dimensional structure of both constructs and measures. These analyses permit the investigator to adopt a model testing framework to contrast alternative, a priori hypothesized models for their underlying structure, and to estimate the relations among the latent constructs free of measurement error. The etiological adjunct to this method is multivariate behavior genetic analysis, in which the covariation among such latent constructs is explained in terms of their shared etiology, namely common and unique genetic and environmental influences (see Carey & DiLalla, 1994; Waldman & Slutske, 2000, for further details and illustrations). In the years to come, these methods should greatly facilitate the efforts of psychopathology researchers to better understand the nature and extent of the covariation among diagnostic entities that are underpinned by latent dimensions rather than taxa.

**Acknowledgments**

We thank Will Grove and Niels Waller for their helpful input on several sections of this article. I.D.W. was supported in part by NIMH grant K01-MH01818.

**References**


Cloninger, C. R. (1978). The link between hysteria and socio-


Received February 8, 2001; accepted April 9, 2001.