Psychological Methods
2006, Vol. 11, No. 3, 287–305

1082-989X/06/$12.00 DOI: 10.1037/1082-989X.11.3.287

Measurement of Psychological Disorders Using Cognitive Diagnosis Models

Jonathan L. Templin
University of Kansas

Robert A. Henson
University of North Carolina at Greensboro

Cognitive diagnosis models are constrained (multiple classification) latent class models that characterize the relationship of questionnaire responses to a set of dichotomous latent variables. Having emanated from educational measurement, several aspects of such models seem well suited to use in psychological assessment and diagnosis. This article presents the development of a new cognitive diagnosis model for use in psychological assessment—the DINO (deterministic input; noisy "or" gate) model—which, as an illustrative example, is applied to evaluate and diagnose pathological gamblers. As part of this example, a demonstration of the estimates obtained by cognitive diagnosis models is provided. Such estimates include the probability an individual meets each of a set of dichotomous *Diagnostic and Statistical Manual of Mental Disorders* (text revision [*DSM–IV–TR*]; American Psychiatric Association, 2000) criteria, resulting in an estimate of the probability an individual meets the *DSM–IV–TR* definition for being a pathological gambler. Furthermore, a demonstration of how the hypothesized underlying factors contributing to pathological gambling can be measured with the DINO model is presented, through use of a covariance structure model for the tetrachoric correlation matrix of the dichotomous latent variables representing *DSM–IV–TR* criteria.

Keywords: diagnostic assessment, latent class analysis, psychological disorders, cognitive diagnosis

Cognitive diagnosis models are special cases of latent class models that characterize the relationship of observable data (typically in the form questionnaire responses) to a set of categorical latent variables (typically dichotomous or binary-valued [0/1]). These models, also known as *multiple classification latent class models*, draw their name from the cognitive nature of the traits they measure (typically a set of skills that an individual may possess). Most commonly, cognitive diagnosis models have been applied in educational testing situations, in which more detailed score reporting is desired. In this article, we show how cognitive diagnosis model applications can provide beneficial information in the diagnosis of psychological disorders. Specifically, cognitive diagnosis models use a hypothesized set of constraints to measure a set of preconceived dichotomous latent variables or dichotomous traits. These constraints are given in the form of specification of the specific traits required to positively respond to each item. As a result of such constraints, the general response characteristics of each group are known. Because of the nature of the dichotomous latent traits and the constraints, cognitive diagnosis models are different from factor analytic approaches (which do not directly model group membership) or traditional methods of clustering and classification (in which additional analyses are required to determine the basic characteristics of each group). In addition, cognitive diagnosis models allow for the study of the structural factors underlying such traits while simultaneously providing diagnostic information.

In this article, we first provide a brief review of various analyses that might be used for similar purposes as cognitive diagnosis models. Such analyses include factor analytic techniques, taxonometric analyses, and classification procedures such as cluster analyses and unconstrained latent class models. We then provide a basic overview of cognitive diagnosis models, giving some discussion to the types of models, their estimation methods, and current approaches to

Jonathan L. Templin, Department of Psychology, University of Kansas; Robert A. Henson, Department of Educational Research Methodology, School of Education, University of North Carolina at Greensboro.

Both authors contributed equally to this article. We thank Jeff Douglas and Sara Templin for their helpful comments on this research and on a draft of this article. Software developed to estimate the model featured in this article can be obtained by contacting J.L.T.

Correspondence concerning this article should be addressed to Jonathan L. Templin, Department of Psychology, University of Kansas, 1415 Jayhawk Boulevard, Room 426, Lawrence, KS 66045. E-mail: jtemplin@ku.edu

the assessment of model fit. Finally, we introduce a new cognitive diagnosis model in the context of a descriptive example, demonstrating how types of such a model can be used to diagnose and study the etiology of psychological disorders (in our example, we investigate pathological gambling).

Factor Analytic Techniques

Factor analytic techniques have played a major role in the development of theories of personality, yet their benefits are limited in clinical diagnostic settings or situations in which the goal is classification. For instance, the Big Five personality traits emanated from a factor-analysis-based body of research (see, e.g., John, 1990). The bulk of such analyses have traits represented by continuous latent variables, with individuals lying on a continuum from low levels to high levels of the latent factor(s). Although the factors that underlie many clinical disorders can be thought of as being on a continuum, clinical diagnoses have distinct classes, in which individuals are diagnosed as either having or not having a specific disorder. Determination of the point on a continuum that designates where a specific diagnosis would be needed cannot be directly accomplished in a single analysis with such an approach.

As in cognitive diagnosis models, some factor analytic approaches (e.g., confirmatory factor analysis) allow the user to define the basic set of underlying traits on the basis of preconceived hypotheses. If group membership is determined using the magnitude of an individual's estimated score on each trait, one can necessarily infer the basic underlying characteristics for each group. Unlike cognitive diagnosis models, however, the latent traits in a factor analytic approach are assumed to be continuous, so additional analyses must be conducted to determine the thresholds used to define group membership. Specifically, because the factor model does not directly assume discrete levels of each trait, if it is to be used for classification purposes, subsequent analyses must be performed to determine the values that categorize each factor. B. O. Muthén (1996) gave an example of such a process regarding the use of factor analysis for diagnosing alcohol abuse and dependency. His article provided a discussion about the possible methods that can be used to determine the thresholds used for diagnosis. Because each method does not necessarily provide the same diagnoses, a description of the usefulness and evaluation of each method follows. Use of the factor analytic approach is feasible if one has an interest in both the factor scores and a possible diagnosis. If the purpose of such an analysis is to attain diagnostic information, then cognitive diagnosis models, which assume discrete latent traits, are one way to determine group membership (established by grouping similarly responding individuals) without the need of additional methods to establish group cut points.

Classification Methods and Taxonometric Approaches to the Measurement of Psychological Traits

In contrast to the factor analytic approaches that assume underlying continuous traits, taxonometric approaches have been developed and applied in the study of psychological disorders. A set of multivariate taxonometric methods have been used for some time, largely stemming from the work of Paul Meehl (see, e.g., Meehl & Golden, 1982). Such methods as MAMBAC (mean above minus below a cut; Meehl & Yonce, 1994) and MAXCOV (maximum covariance; Meehl & Yonce, 1996) are useful for detecting whether individual differences are the result of latent classes or latent continua. Both methods use the observed covariance matrix to make inferences as to the extent of the distinction between continua and taxa. Methods such as these are typically used in the evaluation of whether the latent construct under study is a trait (i.e., continuous) or a type (i.e., discrete) rather than solely for the purpose of classification of individuals. These methods have been applied to varying degrees in personality assessment and clinical diagnostic settings (for summary, see Meehl, 2001; Waller & Meehl, 1997).

Latent class models constitute an additional set of statistical methodologies that provide taxonometric information (see, e.g., Lazarsfeld & Henry, 1968). The underlying premise of these models is that individuals belong to one of a set of latent classes representing groups of individuals with similar response patterns. Commonly, the observable indicator variables are responses to questionnaire items, although the methodology is not limited to survey research.

The generic latent class model is a stochastic model for the probability of a given response. By being stochastic, latent class models automatically incorporate measurement error through their parameterization. To demonstrate, let $P(X_{ij} = x_{ij}|c)$ indicate the probability of a response x_{ij} for individual i on item j ($j = 1, \ldots, J$), conditional on individual i being a member of class c ($c = 1, \ldots, C$). Let $P(c)$ indicate the probability that any given individual is a member of class c, with the constraint $\sum_{c=1}^{C} P(c) = 1$. By assumption, responses are independent conditional on class membership (local independence), so the probability of a response vector, $\mathbf{X}_i = \mathbf{x}_i$, given class c, is the product of the J-conditional item response probabilities given by the latent class model. The generic latent class model is given by

$$P(\mathbf{X}_i = \mathbf{x}_i) = \sum_{c=1}^{C} P(c) \prod_{j=1}^{J} P(X_{ij} = x_{ij}|c). \tag{1}$$

In personality assessment, latent class analysis has seldom been used. Eid, Langeheine, and Diener (2003) demonstrated the potential for using latent class models to detect cross-cultural response differences in the measurement of subjective well-being. In a similar demonstration of the use of classification models in personality assessment, Reise and Gomel (1995) used a mixed-type measurement model (Rost, 1990) to detect class differences in response patterns on a factor of the Multidimensional Personality Questionnaire (Tellegen & Waller, 1994). In an examination of self-monitoring, von Davier and Rost (1997) used latent classification techniques (with a hybrid classification–Rasch model) to study how to best measure the construct. In each of these studies, classification methods yielded information about the likelihood of class membership for each individual, information not directly obtainable by use of factor analytic techniques.

Latent class models have been used in many studies of clinical disorders. For instance, Fink et al. (2004) used a latent class analysis of self-report questionnaire data to group individuals into three classes of hypochondriasis (the classes represented hypochondriasis, other somatoform disorders, and nonsomatoform disorders). Studies investigating the dimensionality of attention-deficit/hyperactivity disorder (ADHD) have compared factor analytic and latent class techniques (see, e.g., Hudziak, Wadsworth, Heath, & Achenbach, 1999; Neuman et al., 1999; for a cross-cultural example, see Rohde et al., 2001). The latent class analyses used in these studies have helped to expand the understanding of ADHD from a single dimension to a multidimensional construct by demonstrating differing etiologies of the disorder across segments of the population. Similarly, a debate over the use of latent continua versus latent categories in diagnoses of depression has been investigated by using a combination of Meehl's taxonometric methods and latent class analyses (for a thorough review, see Solomon, Haaga, & Arnow, 2001).

In addition to Meehl's taxonometric approaches and latent class analysis techniques, a host of other clustering methods have been applied to psychological assessment. Techniques such as *K*-means clustering, hierarchical clustering, and, to a degree, discriminant analysis have been used to study and understand psychological phenomena (for a domain-specific example, see Fals-Stewart, Birchler, Schafer, & Lucente, 1994).

Most clustering–classification applications used in psychological assessment share several common features. Specifically, the number of classes represented in the observed data is not completely determined prior to the time of the analysis but is often chosen, typically by using model comparison techniques. A set of models (each with a different number of classes) is estimated, and the fit of each model is evaluated (e.g., Akaike's information criterion [Akaike, 1974]; the Bayesian information criterion; or Schwarz's Bayesian criterion [Schwarz, 1976]). The solution that "fits" best determines the final preferred number of classes. Often, the "best fitting model" (as chosen by some goodness-of-fit or model comparison criterion) can be difficult to interpret, with class characteristics at odds with what is known from theoretical postulates. For these reasons, the response characteristics of each class in the final selected model are not necessarily fixed prior to an analysis.

With the number of classes unknown prior to the analysis, taxonometric approaches are similar to exploratory factor analytic procedures in that the preferred number of factors is determined by the results of the subsequent analyses. Furthermore, the search for interpretation of the classes is similar to the search for interpretation of factors in an exploratory factor model. Once the best fitting model has been chosen, classes are often described by examining the profiles of conditional item response probabilities estimated by the model (for an example of this process, see von Davier & Rost, 1997, p. 300). In exploratory factor analysis, the extracted factors are described by the magnitudes of the loadings in the final solution. Although similar fit-based searches for the "best" model are used in confirmatory factor analysis and structural equation modeling, these are inherently different because of the specified nature of the latent traits (and their relations both with observable variables and with other latent traits) being defined prior to each analysis (and subsequent fit evaluation).

The process of determining class characteristics post hoc is perhaps a primary drawback to using classification methods in empirical psychological applications. As an alternative, cognitive diagnosis models are similar to methods of confirmatory factor analysis and structural equation modeling—rather than general latent class models (or other clustering methods)—in that the number of classes and the characteristics of each class (with respect to the observable variables and the other classes) are defined prior to each analysis. Just as with confirmatory factor analysis and structural equation modeling, however, nothing precludes the analyst from using cognitive diagnosis modeling techniques in a more exploratory manner, searching for the best fitting (interpretable) model.

Models for Cognitive Diagnosis

Cognitive diagnosis models are latent class models that use constraints to characterize the relation of a set of categorical latent variables to observable responses. Unlike the general unconstrained latent class model, cognitive diagnosis models are unique in the types of constraints they place on the set of latent class item response probabilities. By specification of the model constraints, cognitive diagnosis models fix the number of classes and define the meaning of class membership for the analysis, a process that presumes some semblance of a guiding theory. Furthermore, the item response constraints allow for a larger number of latent classes to be estimated than in the unconstrained latent class model. In this section, we review the cognitive diagnosis modeling framework in the context of educational measurement, the domain from which such models have emanated.

Cognitive diagnosis models link a set of categorical latent variables to a set of equality constraints placed on the model item response probabilities across all latent classes. The preponderance of these models use binary or dichotomously valued latent variables, commonly called *attributes*. It should be noted that not all cognitive diagnosis models require the attributes to be defined for only two levels (for polytomous category attribute models, see Karelitz, 2004; Templin, 2004; von Davier & Yamamoto, 2004). For an individual, the attributes (commonly denoted by $\boldsymbol{\alpha}_i = [\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{iK}]'$) are binary indicators of the possession of a set of K attributes, combinations of which are needed to produce a successful or positive response to an item. Cognitive diagnosis models describe such responses using the multiple dimensions represented by the latent attributes. These parameters are the multiple latent indicators described by the *multiple classification latent class model* synonym for cognitive diagnosis models. The finite set of possible combinations of these binary indicators, containing 2^K possible attribute patterns, represents the fixed number of classes for the analysis.

To signify the relationship between items and the attributes needed for positive responses, an item-by-attribute *Q-matrix* is constructed. For a given item, the entries in the Q-matrix denote that a positive response to the item is heavily influenced by the indicated attributes. Similar to a confirmatory factor model, these Q-matrix indicators are binary: Either the item response is influenced by an attribute or the item response is not influenced by an attribute. In the analysis of a questionnaire with a cognitive diagnosis model, poor specification of the Q-matrix can result in model parameters with little diagnostic or classificatory value (an important point, discussed in more detail in the remarks following the example application). Multiple types of cognitive diagnosis models exist (as explained below), and thus the type of cognitive diagnosis model and the entries of the Q-matrix define the set of probabilistic equality constraints for item response probabilities of the latent class model, of which the cognitive diagnosis model is a subset.

Primarily, cognitive diagnosis models have been used in educational measurement, in which the set of categorical latent variables is thought to represent the set of skills an individual may possess (or to make up the knowledge state of an individual; Falmagne, Koppen, Villano, Doignon, & Johannesen, 1990). A common example of an application of cognitive diagnosis models in educational measurement comes from the measurement of mathematics ability. Specifically, cognitive diagnosis models have been applied to estimate the elements of specific abilities that middle school students must have mastered to successfully subtract fractions (C. Tatsuoka, 2002; K. Tatsuoka, 1990). For example, to successfully find the answer to $\frac{3}{4} - \frac{3}{8}$, an individual must know how to (a) find a common denominator and (b) subtract numerators. Mastery of both of these skills will produce a higher probability of a correct answer, whereas the lack of one or both of these skills will dramatically reduce the chances of finding the correct solution. Cognitive diagnosis models can be valuable to educators because of the diagnostic information that can be generated about each individual test taker. A teacher's knowledge of the types of skills students possess can greatly help the course of instruction by ensuring that oft-mastered skills are focused on less than are skills that many students may lack. The fraction subtraction domain is a narrow example of the areas in which cognitive diagnosis models have been used (see, e.g., de la Torre & Douglas, 2004). We note that other domains (and even other tests of fraction operations) may not share the same qualities of this illustrative example, in that development of Q-matrices may be much more complicated (for an example of such, see Bouwmeester, Sijtsma, & Vermunt, 2004).

Cognitive Diagnosis Modeling Framework

To illustrate typical analyses using cognitive diagnosis models, in this section we detail the general modeling framework (the types of models incorporating the item response and latent variable structure, typical estimation methods, and model fit evaluation procedures).

Measurement Models

A common model used for cognitive diagnosis is the DINA model (Haertel, 1989; Junker & Sijtsma, 2001; Macready & Dayton, 1977). Although this type of model originated from the work of Macready and Dayton (1977), the name *DINA* (deterministic inputs; noisy "and" gate), coined by Junker and Sijtsma (2001), has become a popular acronym for this model. Defined for dichotomous response items, the DINA model imposes constraints on response probabilities such that the probability of a positive response is one of two values. Determining the probability of positive response for individual i and item j is the binary latent variable ξ_{ij}:

$$\xi_{ij} = \prod_{k=1}^{K} \alpha_{ik}^{q_{jk}}, \tag{2}$$

where q_{jk} is the binary Q-matrix entry defining possession of attribute k as influential in the response tendencies for item j. If individual i has mastered all Q-matrix-defined attributes for item j, then $\xi_{ij} = 1$. If individual i has not mastered one or more Q-matrix-defined attributes for item j, then $\xi_{ij} = 0$.

Using the ξ_{ij} parameter, the DINA model specifies the conditional probability of a successful response as

$$P(X_{ij} = 1|\xi_{ij}) = (1 - s_j)^{\xi_{ij}} g_j^{(1-\xi_{ij})}, \qquad (3)$$

where $1 - s_j > g_j$. The s_j parameter is commonly referred to as the *slip* parameter, where the value represents the probability that a master of all necessary attributes slips and answers the item incorrectly. Here $1 - s_j$ is the probability of a correct response to item j if all Q-matrix-defined attributes have been mastered ($\xi_{ij} = 1$). The g_j parameter represents the *guess* parameter, which is the probability that a nonmaster of at least one necessary attribute answers the item correctly (by guessing). By dividing individuals into groups on the basis of mastery of all attributes or not, the DINA model is a conjunctive or noncompensatory model, meaning that lacking one attribute cannot be compensated for by a preponderance of another attribute. To elaborate, we borrow from the fraction subtraction example presented above, with the item $3/4 - 3/8$. Note that to answer this item correctly, an individual must know how to (a) find a common denominator and (b) subtract numerators. In a conjunctive or noncompensatory model, an individual who lacks mastery of either of these attributes will most likely fail to answer this item correctly, irrespective of any other attributes this individual may possess. This is in contrast to compensatory models such as the common factor model, in which the additive parameterization allows for low levels on certain factors to be compensated for by high levels on other factors.

Through use of the binary-valued latent ξ parameter, for a given item, the DINA model maps the set of all 2^K attribute patterns to one of two item response probabilities per item (either $1 - s_j$ or g_j). The "and" portion of the DINA acronym is represented by the product in Equation 2, where mastery of all Q-matrix-defined attributes for item j places individuals into the high-probability response class (with response probability $1 - s_j$). The two-probability constraint restriction used by the DINA model can be unnecessarily restrictive, treating all nonmastered attributes equivalently. In practice, the combination of this restriction and poorly formed Q-matrices can lead to abnormally large values of slip and guess parameters (indicating that masters often slip and nonmasters often guess correctly). Other cognitive diagnosis models relax this restriction so that more item response equivalence classes are introduced. The NIDA (noisy inputs; deterministic "and" gate) model by Maris (1999; acronym by Junker & Sijtsma, 2001) allows for $2^{q_{j\cdot}}$ different response probability equivalence classes per item (where $q_{j\cdot} = \Sigma_{k=1}^{K} q_{ik}$ for an item, representing the sum of the number of Q-matrix indicators across all attributes), with the additional constraint that items with identical Q-matrix entries will have identical response functions. The RUM (reparameterized unified model; DiBello & Stout, 2003; Hartz, 2002) also allows for $2^{q_{j\cdot}}$ response probability constraints and for differing response functions for items with identical Q-matrix entries. Expanding beyond cognitive diagnosis models, similar latent class model approaches have also been developed (see, e.g., Formann, 1992; Yamamoto & Everson, 1997).

The conditional item response probability portion of a cognitive diagnosis model can be considered analogous to those found in factor analytic measurement models (although with categorical latent traits rather than continuous latent traits), defining the relationship between the first-order (categorical) latent variables and the data. The DINA, the NIDA, and the RUM all model the conditional response probability of a latent class model—the $P(X_{ij} = x_{ij}|c)$ term in Equation 1.

Structural Models

Also included in the latent class model is the class membership probability, $P(c)$. Because cognitive diagnosis models define the number of latent classes through the attributes defined by the Q-matrix, the total number of latent classes in cognitive diagnosis models is exponential in the number of attributes, $C = 2^K$. The exponential number of latent classes makes full specification of the class membership probability vector implausible because of the number of parameters needed for the saturated model ($2^K - 1$).

Because of the dichotomous nature of the attributes found in cognitive diagnosis models, one method for modeling the probability of class membership is to use the concept of tetrachoric correlations to place structure on the joint distribution of the attributes. A tetrachoric correlation is the correlation between two underlying normally distributed variables (with zero mean and unit variance) that have both been dichotomized by respective cut-point parameters. Extrapolating from the bivariate distribution of any pair of given attributes to the joint distribution of all attribute patterns, the tetrachoric model presumes underlying continuous multivariate normal variables with a zero mean vector and a tetrachoric correlation matrix, $\boldsymbol{\Xi}$. The set of parameters that dichotomize each of the continuous variables are referred to as the *cut-point parameters* (for a given attribute k, this is denoted by κ_k), from which information regarding the marginal proportion of individuals being considered masters of attribute k is found. Because of the tetrachoric model's assumption of an underlying cut point, these parameters are transformed from the normal scale by taking the function $\Phi(-\kappa)$, where $\Phi(\cdot)$ is the standard normal distribution function. Such information provided by the cut-point parameters is often useful in educational measurement to determine the extent to which a given attribute (or skill) has been mastered in the population of interest, providing information regarding the base rate for attribute mastery. Furthermore, if base rate information is known prior to analysis, these parameters can be fixed in estimation to incorporate the knowledge of the proportion of individuals in possession of each attribute.

Through use of the tetrachoric correlation parameterization, the total number of population parameters is decreased (from $2^K - 1$ parameters, representing class membership or attribute pattern probabilities in the saturated model, to

$$K + \frac{K[K - 1]}{2}$$

in the tetrachoric correlation model with K cut points and

$$\frac{K[K - 1]}{2}$$

tetrachoric correlations). From this reduced set of parameters, the probability of any given attribute pattern—the class membership probability $P(c)$ in Equation 1—is found by use of a multivariate integral across the multivariate normal distribution, with latent attribute tetrachoric correlation matrix denoted by $\boldsymbol{\Xi}$ and with limits of integration being set by the cut-point parameters (κ_k, $k = 1, \ldots, K$):

$$P(c) = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_k}^{b_k}\frac{1}{(2\pi)^{K/2}|\boldsymbol{\Xi}|^{1/2}} \exp\left[-\frac{1}{2}\tilde{\boldsymbol{\alpha}}_c\boldsymbol{\Xi}^{-1}\tilde{\boldsymbol{\alpha}}'_c\right]d\tilde{\alpha}_{c1}d\tilde{\alpha}_{c2}\cdots d\tilde{\alpha}_{cK}, \quad (4)$$

where, for each possible attribute pattern, c (denoted $\alpha_c = [\alpha_{c1}, \alpha_{c2}, \ldots, \alpha_{cK}]$), the limits of integration for attribute k are

$$a_k = \begin{cases} \kappa_k & \text{if} \quad \alpha_{ck} = 1 \\ -\infty & \text{if} \quad \alpha_{ck} = 0 \end{cases}, \quad b_k = \begin{cases} \infty & \text{if} \quad \alpha_{ck} = 1 \\ \kappa_k & \text{if} \quad \alpha_{ck} = 0 \end{cases}\cdot$$

As discussed in the next section, the structural model can be estimated simultaneously with the parameters of the item response model.

Algorithms have been developed to estimate the full tetrachoric correlation matrix (Hartz, 2002) using estimation procedures developed to ensure positive definiteness of the entire matrix (Barnard, McCulloch, & Meng, 2000). Another benefit of modeling the class membership probabilities by tetrachoric correlations is that further structures can be imposed on the matrix of tetrachoric correlations ($\boldsymbol{\Xi}$). For instance, a one-factor correlation structure has been developed, estimating attribute association through use of a higher order trait (de la Torre & Douglas, 2004; Templin, 2004). As we demonstrate below, using the structural parameterization of Templin (2004), the correlation structure model can be generalized from a single factor to multiple factors in a manner consistent with common structural equation modeling parameterizations (such as the reticular action model; McArdle & McDonald, 1984).

The tetrachoric structural modeling framework in cognitive diagnosis has evolved in a separate but parallel tract from the development of similar mixed-type latent variable models (see, e.g., B. O. Muthén, 1984). For general mixed-type latent variable models, the association between the classes (defined as the categorical latent variables) and the traits (defined as the continuous latent variables) has been modeled routinely, and such abilities have been incorporated into widely available software packages such as Mplus (L. K. Muthén & Muthén, 1998). The concepts of such structural model parameterizations from cognitive diagnosis share similar features with those of the general mixed-type latent variable models; however, the use of such structural models in cognitive diagnosis has been much more limited when compared with the general structural models for mixed-type latent variables. Within the set of cognitive diagnosis models, such modeling procedures have not been applied outside of the full-matrix and one-factor approaches.

Estimation Methods

Cognitive diagnosis models, like latent class models, have been estimated using a number of different algorithms. In the earlier applications of such models, the expectation maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977) was featured as the main estimation technique (see Haertel, 1989; Maris, 1999). Over the past 15 years, Markov chain Monte Carlo (MCMC) algorithms for estimation of statistical models have become more common. For instance, the Arpeggio software program provides an MCMC algorithm for estimation of the RUM and has been used by the Educational Testing Service in analyses with cognitive diagnosis models (DiBello & Stout, 2003; Hartz, 2002; Stout, 2002). Both the EM and MCMC algorithms can have difficulty with achieving convergence in estimation, with nonconvergence typically indicating that something is amiss with the overall analysis (e.g., a poorly specified Q-matrix). In cases in which convergence is achieved, with both types of algorithms, the model parameter estimates obtained from either method are, in general, statistically sound.

Goodness-of-Fit Evaluation Methods

Because cognitive diagnosis models are constrained latent class models, measures used to evaluate model fit in cognitive diagnosis are similar to those used to evaluate the fit of general latent class models, which are customarily based on statistical hypothesis tests. Such fit statistics are based on functions comparing observed response patterns with those expected under the latent class model. The probability of any given response pattern is found by inserting item and individual parameter estimates into Equation 1. For each possible response pattern, the expected number of occurrences is then compared with the actual number of occurrences. One of three commonly used statistics is computed: the Pearson statistic, the likelihood ratio test, and the Cressie–Read statistic (for a description, see Read & Cressie, 1988). As the sample size approaches infinity (and a set of regularity conditions are met), each of these three statistics is chi-square distributed, with degrees of freedom equal to the number of possible response patterns minus the number of parameters in the model minus the number of latent classes defined. All of these statistics are computed across all possible response patterns, making them applicable in situations in which the number of variables is small and the size of the sample is large.

Because the number of possible response patterns increases exponentially as the number of variables increases, sparse tables of observed response patterns result for most reasonable sample sizes. In cases of sparse tables, Monte Carlo (bootstrap) resampling techniques to produce "empirical" p values for the model fit statistics have been advocated (Langeheine, Pannekoek, & van de Pol, 1996). Note that the Monte Carlo model fit evaluation techniques are different than the MCMC algorithms used to estimate the model parameters. The Monte Carlo bootstrap procedure takes the parameter estimates found from an analysis and then simulates multiple data sets from the latent class (or cognitive diagnosis) model using the estimated parameters from the analysis. For each simulated data set, the latent class (or cognitive diagnosis) model is then estimated, and the values of the model fit statistics (Pearson, likelihood ratio, or Cressie–Read) are recorded. Because this process is repeated for multiple simulated data sets (approximately 300 for the Pearson or the Cressie–Read statistic, as suggested by Langeheine et al., 1996), the proportion of statistics greater than the observed value is used as the empirical p value. If this value is greater than .05, the model is not rejected," implying that the model adequately fits the data.

The aforementioned goodness-of-fit methods have been developed directly from applications of general latent class models. With respect to cognitive diagnosis models, however, several goodness-of-fit indices have been developed and implemented. De la Torre and Douglas (2004) have advocated comparing measures of item association (e.g., item correlation, log odds, or Cohen's κ) in the evaluation of model fit. In this context, and for each pair of items, the observed values of a given measure of item association are compared with values that are predicted by the parameters of the cognitive diagnosis model. For a pair of items j and k, the model-predicted item association is found by constructing the contingency table of model-predicted item response frequencies. The model-predicted probability for response a to item j and response b to item k (for dichotomous items, a and b can be either 0 or 1) is given by

$$\hat{P}(X_j = a, X_k = b) = \sum_{c=1}^{C}[\hat{P}(X_j = a|c)\hat{P}(X_k = b|c)]\hat{P}(c), \quad (5)$$

where all functions and variables with a circumflex are those quantities predicted by the estimated model (appearing in the innermost portion of the right-hand side). To summarize the overall discrepancy between observed and model-predicted item association, the root-mean-square error (RMSE) comparing these values is computed across all items. In assessing the model fit of a structural equation model, this fit statistic is also called the *standardized root-mean-square residual*. The RMSE of the discrepancy between observed and predicted item association should be small if the cognitive diagnosis model adequately fits the data. Such goodness-of-fit methods were developed to mimic those predominantly used in evaluating the fit of structural equation models.

In many instances, questionnaires or inventories of even moderate size can render the methods described above ineffectual. For 40 items, there are over 1 trillion possible response patterns, for each of which the model-predicted probability must be computed to calculate any of the general latent class goodness-of-fit indices (Pearson, likelihood ratio, or Cressie–Read). Additional computation of any of these three fit statistics under the Monte Carlo bootstrap approach compounds the number of computational steps needed to assess the model goodness of fit. Although the methods used by de la Torre and Douglas (2004) provide for a value indicating the goodness of fit of the model (on the scale of the measure of item association used), these methods do not provide any empirical p value for use in a test of the fit of the model. As an alternative, to judge the goodness of fit of the model, we suggest combining the Monte Carlo goodness-of-fit procedures of Langeheine et al. (1996) with the methods of de la Torre and Douglas (2004)—model evaluation indices based on item association. The methods of the resampling procedure are the same as detailed in Langeheine et al. (1996), with exception that the statistic of interest is the RMSE for item association. The result of this combination is a goodness-of-fit statistic that is easily interpretable (the RMSE for the discrepancy between model-predicted and observed item association) and an associated bootstrap p value indicating the probability that such a result is extreme when compared with data generated from a similar model.

The ideal of practical methods for model fit in cognitive diagnosis remains an open research question. A balance must be struck between computationally intensive fit-evaluation procedures (that end up taking much longer to run than the initial analysis) and simpler heuristics that indicate the fit of a model in a reasonable amount of time. The item association RMSE (using any measure of item association) evaluates the fit of the model to describe item pairs rather than the entire response pattern (as is computed for the general latent class fit measures), allowing for practical fit evaluations of the application of latent class or cognitive diagnosis models with conventional data sets in psychological applications.

The Use of Cognitive Diagnosis Models in Psychological Assessment

The diagnostic capabilities of cognitive diagnosis models seem well suited to applications in psychological measurement. Take, for instance, psychological disorders, as defined in the *Diagnostic and Statistical Manual of Mental Disorders* (text revision [*DSM–IV–TR*]; American Psychiatric Association, 2000). Many disorders are diagnosed as being pathological if a set of dichotomous criteria have been satisfied.[1] The set of criteria can be measured as the dichotomous latent attributes found in cognitive diagnosis models, and assessment can determine the probability that an individual meets a given criterion or the probability that an individual has a disorder (where a disorder is defined by satisfying a set of dichotomous criteria). Furthermore, cognitive diagnosis models allow for additional information to be culled from the analysis of the dichotomous attributes. Underlying structural models can be estimated simultaneously (e.g., models for the tetrachoric correlation of the attributes), aiding in the development of theories pertaining to the path that underlying psychological traits take in manifesting themselves as pathological disorders. The ability to combine specific classificatory methods with structural models makes use of cognitive diagnosis models a potentially valuable method for psychological assessment.

To demonstrate the use of cognitive diagnosis models in psychological assessment, we focus on the analysis of an instrument designed to assess the structure of underlying latent variables in pathological gambling. We provide an analysis of this instrument with a cognitive diagnosis model, and we show how this instrument—developed for structural analyses, not diagnostic purposes—can be expanded to diagnose probable pathological gamblers with the help of cognitive diagnosis models. The demonstration is primarily provided as an example and is not intended to espouse either the benefits of using this specific instrument over other instruments or the validity of the diagnostic criteria for which this instrument was created to measure (as defined in the *DSM–IV–TR*). Ultimately, we hope this example will provide a guide for the possible information that can be culled from applying cognitive diagnosis models in psychological assessment.

Diagnosis of Pathological Gambling

Pathological gambling is defined in the *DSM–IV–TR* as an *impulse-control disorder not elsewhere* (i.e., otherwise) *classified* (American Psychiatric Association, 2000). Impulse-control disorders are defined as behaviors in which individuals suffer from "the failure to resist an impulse, drive, or temptation to perform an act that is harmful to the person or others" (American Psychiatric Association, 2000, p. 663). For an individual to be classified as a pathological gambler, 5 of the 10 criteria displayed in Table 1 must be satisfied. The *DSM–IV–TR* also specifies that the gambling behavior cannot be a result of a manic episode.

Notice that the 10 criteria are dichotomously defined: An individual either does or does not satisfy a criterion. Again, we note that use of such criteria to define pathology assumes that the criteria are valid indicators of the pathology. In this context, cognitive diagnosis models estimate a *profile* for each individual, specifying the probability that an individual meets each criterion. The use of such profiles shows differences with respect to why an individual received a diagnosis as a probable pathological gambler (meeting at least 5 of the criteria). For instance, two individuals who may both be diagnosed as pathological gamblers may have different etiological paths for the disorder. In addition to meeting several other criteria, the first individual may be obsessed with thinking about gambling (Criterion 1) and become irritable when trying to cut back on gambling (Criterion 4). The second individual may differ from the first in that he or she may lie to loved ones to hide his or her gambling (Criterion 7) and has perhaps caused career damage by his or her gambling habits (Criterion 9). Both individuals may be pathological gamblers, but the manifestations of their disorders (and perhaps the potential avenues for treatment) are the results of different behaviors. The likelihood of any individual satisfying any of the defined criteria would be determined by the differential profiles that would be estimated by a cognitive diagnosis model.

Although the *DSM–IV–TR* defines criteria to be met for an individual to be diagnosed as a pathological gambler, in many cases information at the level of the criteria has not been the focus of psychological questionnaires. For example, one of the most commonly used instruments used to study pathological gambling is the South Oaks Gambling Screen (SOGS; Lesieur & Blume, 1987). Lesieur and Blume (1987) went to great lengths to develop an instrument that would aid in the diagnosis of pathological gambling by externally validating the results of the instrument with clinical diagnoses for a large number of individuals. The SOGS consists of 20 dichotomous (yes/no) items, whereby the sum of any 5 *yes* responses indicates that an individual is a probable pathological gambler. The distinction of 5 or more positive responses to SOGS items being indicative that the respondent is a probable pathological gambler was made from empirical research in calibrating the SOGS (Lesieur & Blume, 1987). Although the SOGS is known for its ability to identify pathological gamblers, it

[1] We must note, however, that the effectiveness of such models rests largely on the theory guiding the development of diagnostic instruments. If, for example, the definition of a disorder, as specified by the *DSM–IV–TR*, is misguided, model-based diagnoses will have poor validity.

Table 1
DSM–IV–TR Diagnostic Criteria for Pathological Gambling

Persistent and recurrent maladaptive gambling behavior as indicated by five (or more) of the following:
(1) is preoccupied with gambling (e.g., preoccupied with reliving past gambling experiences, handicapping or planning the next venture, or thinking of ways to get money with which to gamble)
(2) needs to gamble with increasing amounts of money in order to achieve the desired excitement
(3) has repeated unsuccessful efforts to control, cut back, or stop gambling
(4) is restless or irritable when attempting to cut down or stop gambling
(5) gambles as a way of escaping from problems or of relieving a dysphoric mood (e.g., feelings of helplessness, guilt, anxiety, depression)
(6) after losing money gambling, often returns another day to get even ("chasing" one's losses)
(7) lies to family members, therapist, or others to conceal the extent of involvement with gambling
(8) has committed illegal acts such as forgery, fraud, theft, or embezzlement to finance gambling
(9) has jeopardized or lost a significant relationship, job, or educational or career opportunity because of gambling
(10) relies on others to provide money to relieve a desperate financial situation caused by gambling

Note. From *Diagnostic and Statistical Manual of Mental Disorders* (text revision [*DSM–IV–TR*], p. 674), by American Psychiatric Association, 2000, Washington, DC: Author. Copyright 2000 by American Psychiatric Association. Reprinted with permission.

does not allow for estimation of the set of criteria an individual has met. In fact, investigation of the items of the SOGS in comparison with the *DSM–IV–TR* criteria definition reveals that not all of the criteria are measured by the SOGS.

To investigate the *DSM–IV–TR* criteria, Feasel, Henson, and Jones (2004) developed the Gambling Research Instrument (GRI). The GRI was developed to measure pathological gambling on a continuum, which provided a way to study the psychological traits that underlie pathological gambling and *not* to diagnose pathological gamblers. In the development of the GRI, 41 Likert-type scale items were created (25 items were used for analyses). Individuals respond to an item by indicating the extent to which they participate in particular activities or indicating the extent to which they feel an item adequately describes their behaviors. In addition, the GRI was written so that each criterion can be measured with a subscore (meaning the items being used for the subscore are psychometrically unidimensional). By using subscores for all 10 criteria, the underlying factor structure can be explored. Feasel et al. (2004) noted that the *DSM–IV–TR* implies a three-factor structure in the description of impulse-control disorders not elsewhere classified.

As it was developed, the GRI cannot be used directly as a diagnostic instrument. It is not clear what combination of subscores from the GRI would be indicative of an individual being a pathological gambler or, at the criterion level, what cut point should be used to identify an individual who has met a specific criterion. However, if the data were analyzed using a model that explicitly requires the responses to be functions of the 10 underlying dichotomous criteria (as specified in the *DSM–IV–TR*), one could use the GRI as a diagnostic instrument. Cognitive diagnosis models would allow for the GRI to be used both as a diagnostic instrument and as a means to study the underlying structure of pathological gambling.

As discussed above, cognitive diagnosis models constitute a set of constrained latent class models in which each class is defined by an individual's set of dichotomous latent attributes. In the case of pathological gambling, the attributes are now considered to represent the 10 *DSM–IV–TR* criteria, with mastery of an attribute referred to as having *satisfied a criterion* and nonmastery of an attribute referred to as having *not satisfied a criterion*. Just as in general latent class models, all individuals with the same criteria profile (satisfy or not for each criterion, the class into which an individual belongs) are expected to respond to all items in a similar manner. On the basis of this set of models, we can estimate the probability that each criterion has been satisfied and, therefore, the probability that an individual is a pathological gambler (i.e., has satisfied 5 or more criteria). In addition, a model for the association between each of the criteria can be provided. In the case of pathological gambling, these associations (tetrachoric correlations) can be modeled in terms of the factor structure discussed in Feasel et al. (2004) while providing diagnostic information about the criteria.

A Cognitive Diagnosis Model for Psychological Assessment

Although cognitive diagnosis models seem to be readily extended for psychological assessment, the structure of some psychological instruments (such as the GRI) prevents their direct application. Modification of the educational measurement versions of such models is needed because of assumptions about the nature of the latent variables (conjunctive vs. disjunctive). Specifically, item responses in personality instruments such as the GRI are not expected to function as conjunctive or noncompensatory expressions of the dichotomous criteria (as is assumed by the DINA cognitive diagnosis model). Furthermore, there is a need to explore the underlying structure of these criteria as expressed by the association between the dichotomous latent traits representing the 10 *DSM–IV–TR* criteria. The nature of this structure goes beyond single-factor higher order approaches already developed (de la Torre & Douglas, 2004; Templin, 2004), and yet it is not as general as the unstructured approach (Hartz, 2002). The need is to have an association method for defined correlation structures, analogous to the mixed-type latent variable models of B. O. Muthén (1984) but in the context of cognitive diagnosis models. Therefore, we define a new cognitive diagnosis model intended to incorporate the properties of psychological assessment instruments, and as a demonstration, we provide an example using the GRI.

Evolving from educational assessment, cognitive diagnosis models such as the DINA, the NIDA, and the RUM are defined as conjunctive across the categorical latent attribute profile. For studying psychological disorders, however, conjunctive models are not necessarily reasonable. The point is even more evident when one explores the items written for the GRI. Take, for example, Item 16:

> 16. I am ashamed of the things I've done to obtain money for gambling.

A positive response may be provided by an individual for several reasons. The individual may be ashamed because he or she has committed illegal acts (Criterion 8) *or* relied on others to finance his or her gambling (Criterion 10). In contrast to conjunctive models, which assume that both criteria must be met before an individual will respond positively to this item, a disjunctive model (one stating that having satisfied either of the two dichotomous attributes will lead to a positive response) might better suit the item. In general, disjunctive models allow for multiple paths to a common response. In the case of the GRI, a disjunctive model would state that a positive response is the result of meeting one or more of the criteria measured by the item.

In the next section, we define a disjunctive model based on the DINA. The advantage of a DINA-based model is that it requires a limited number of item parameters (two per item) and can easily be modified to be a disjunctive model. We call the disjunctive version of the DINA the *deterministic input; noisy "or" gate* (or DINO) model. Following the introduction of the DINO, we develop an extension of the one-factor correlational structure model of Templin (2004), modeling the underlying factor structure of the criteria that define pathological gambling in the context of a cognitive diagnosis model.

The DINO model. As parameterized in the DINA, each item has two different probabilistic equivalence classes. The DINA accomplished this by defining, for each item and individual, the latent variable ξ_{ij}, as in Equation 2. Individuals satisfying all necessary criteria (or attributes, as defined by the Q-matrix entries) for an item ($\xi_{ij} = 1$) have a probability of a positive response equal to $(1 - s_j)$, and those lacking at least one required criterion ($\xi_{ij} = 0$) have a probability of a positive response equal to g_j. We define the DINO in a similar manner, with the probabilistic equivalence classes originally determined by ξ_{ij} now based on a disjunctive model. Specifically, we redefine ξ_{ij} using the notation ω_{ij}, where

$$\omega_{ij} = 1 - \prod_{k=1}^{K}(1 - \alpha_{ik})^{q_{jk}}. \tag{6}$$

Notice that ω_{ij} divides individuals into two groups: a group of individuals who have satisfied *at least* one Q-matrix-necessitated criterion ($\omega_{ij} = 1$) and a group of individuals who have not satisfied any Q-matrix-necessitated criteria ($\omega_{ij} = 0$). Given the definition of ω_{ij}, the DINO model defines the probability of a positive response in a manner similar to that of the DINA model:

$$P(X_{ij} = 1|\omega_{ij}) = (1 - s_j)^{\omega_{ij}} g_j^{1-\omega_{ij}}, \tag{7}$$

where $1 - s_j > g_j$. As in the DINA, s_j and g_j are the slip and guess parameters, respectively.

Modeling the hierarchical structure of criteria using the DINO. Up to this point, we have defined a cognitive diagnosis model that can be used to analyze psychological data. However, diagnosis is not the only focus of psychological studies. Often, research involves investigating the underlying structure of an individual difference characteristic or trait. For example, the GRI was developed to measure each of the 10 *DSM–IV–TR*-defined criteria with large enough variability so that the underlying structure of pathological gambling could be explored. The *DSM–IV–TR* suggests that impulse-control disorders not elsewhere classified may have three underlying traits (dependency, loss of control, and disruption). Because cognitive diagnosis models estimate the presence or absence of a set of criteria, it is possible to model the association between these criteria and, therefore, explore the structure suggested by Feasel et al. (2004).

As alluded to above, estimation algorithms have been developed that model the underlying structure of attributes using a hierarchical factor (de la Torre & Douglas, 2004; Templin, 2004). We expand this concept to allow for multiple factors by providing a general factor model for the tetrachoric correlation matrix characterizing the attribute associations, $\boldsymbol{\Xi}$ (from Equation 4):

$$\boldsymbol{\Xi} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}. \tag{8}$$

The structural model shown in Equation 8 is the matrix form commonly used in confirmatory factor analytic models. This model relates the K dichotomous attributes to a set of F higher order factors. The elements of the $\boldsymbol{\Lambda}$ matrix (of size $K \times F$) are the loadings of the attributes onto the higher order factors, and they are estimated through use of a factor pattern matrix constructed to ensure identifiability of the model (for identifiability conditions, see, e.g., McDonald, 1999). The elements of the $\boldsymbol{\Phi}$ matrix (of size $F \times F$) are the factor correlations. Finally, as in confirmatory factor analysis, the $\boldsymbol{\Psi}$ matrix (of size $K \times K$) is a diagonal matrix containing the attribute uniqueness parameter estimates. The result of this parameterization is a reduction of the total number of parameters to be estimated under the general unstructured tetrachoric correlation matrix, at the cost of imposing an F-dimensional factor structure on the matrix $\boldsymbol{\Xi}$.

Although the factor model is parameterized as in Equation 8, it bears mention that this parameterization is not the only structure that can be imposed on the attribute correlation matrix. Depending on the application, other parameterizations can be used. For instance, a path model of the latent attributes can be parameterized. Any commonly parameterized structural equation model (such as the LISREL model) can be used for the elements of $\boldsymbol{\Xi}$. If precise structural hypotheses are not known, the full tetrachoric correlation matrix describing the association between the attributes (or criteria) can be estimated (Hartz, 2002). Furthermore, parameterizations abandoning tetrachoric correlations, such as log-linear or hierarchical conjunctive models, can be used.

The tetrachoric correlation matrix is only part of the structural model in cognitive diagnosis. Recall that a tetrachoric correlation is the correlation between two underlying continuous (normally distributed) variables that have been dichotomized by a cut point. Therefore, in each of the aforementioned estimation algorithms, a set of cut-point parameters is estimated. The cut-point parameters hold significant meaning in that they can be converted to a proportion (or percentage) representing the marginal level of mastery for each attribute (or, in our case, satisfaction of each criterion). Such parameters provide information regarding the base rate of each attribute in the population represented by the data set.

Example Application: Diagnosis With the GRI

We use pathological gambling as an example to demonstrate the application of cognitive diagnosis models in psychological assessment. Particularly, we intend to demonstrate that an instrument developed for psychological assessment of the factors underlying a clinical disorder can be used to provide diagnostic information via analysis with a cognitive diagnosis model. Note that this example is intended to be a demonstration of the potential information that a cognitive diagnosis model may provide; it is not intended to provide support for the instrument from which the estimates were obtained, an instrument that was not developed for diagnostic purposes.

Method

The data for our analysis are from a study of gambling tendencies of college students from a small midwestern state. The motivation for the study came from the recent advent of online gaming and nationally televised poker tournaments on various sports, travel, and entertainment channels. Such television programs seem to be somewhat popular among the college-age population, and online gaming has allowed college students access to gaming on a scale that was not possible until recently.

Data were collected from 593 individuals who were given course credit as payment for participation. The 41 GRI items were constructed on a 6-point Likert-type scale (scored 0–5). To use these items with the DINO model (which models dichotomous responses), we dichotomized each item by using the mean value as a cut point. For most items, the mean value was just under 1, something not unexpected for a college population.

The responses to the GRI were analyzed using the DINO model introduced in this article. Note that cognitive diagnosis model analyses require a Q-matrix, an indicator matrix specifying the attributes measured by each item. As noted above (and discussed below), the construction of the Q-matrix is central to the validity of the diagnostic results of this analysis. In the case of pathological gambling, the Q-matrix contains 10 columns, one for each of the 10 criteria. Two individuals experienced in measurement of pathological gambling discussed each item and agreed on a Q-matrix, which was constructed assuming the disjunctive structure parameterized by the DINO. The Q-matrix constructed had an average of 1.34 attributes assigned per item and an average of 5.5 items assigned per attribute. To illustrate how the Q-matrix looks, the Q-matrix entries for the first 15 items are presented in Table 2. It is important to note that estimation of the DINO model with a different Q-matrix could produce results substantially different from those of the current analysis.

The DINO model parameters were estimated using an MCMC algorithm that had uniform priors on all item and structural parameters. As a check of the accuracy of the MCMC algorithm, we ran a small simulation study prior to the analysis, with results indicating adequate recovery of all estimated parameters. The estimation algorithm was run with a chain length of 50,000 iterations, of which the first 40,000 were used as the burn-in period. Convergence was evaluated using the Geweke (1992) index and by

Table 2
Q-Matrix Entries for the First 15 Items in the Example Application

	Criterion									
Item	1	2	3	4	5	6	7	8	9	10
1	1	1	1	0	0	1	0	0	0	0
2	1	0	0	0	1	0	0	0	0	0
3	0	0	0	0	0	1	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	1	1	1
7	0	0	0	0	1	0	0	0	0	0
8	0	1	1	0	0	1	0	0	0	0
9	0	0	0	1	0	0	0	0	0	0
10	1	0	0	0	0	0	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0
12	0	0	0	0	0	1	0	0	0	0
13	0	0	1	1	0	0	0	0	0	0
14	0	1	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	1	0	0	0

visual inspection of the MCMC chain plots. Given the convergence of the MCMC chain to a steady state, posterior mean estimates of all of item and structural parameters were obtained, from which model fit was evaluated and parameter estimates were interpreted.

From the model, we estimated posterior probabilities of satisfying each criterion for each individual in addition to obtaining an estimate of the posterior probability of meeting the *DSM–IV–TR* definition of a pathological gambler (satisfying 5 or more criteria). If an individual had a posterior probability of meeting 5 or more of the criteria that was greater than or equal to .50, then that individual was classified as a pathological gambler. We note that the .50 distinction is made because of the estimates being in the form of probabilities. Using the .50 distinction, we are specifying that the estimate for an individual is the most likely estimate found under the model.[2]

Results

Because of the differing types of information obtained in our example analysis, we divide the results into four sections: model fit, individual diagnoses, item parameters, and structural model parameters. Each of these sections describes a set of results provided by the cognitive diagnosis model related to the analysis of the GRI.

Model fit results. With the number of dichotomized items (41), a total of $2^{41} = 2.2 \times 10^{12}$ possible response patterns exist. The large number of possible response patterns compared with our sample size (593 individuals used for analysis) calls the appropriateness of each of the three commonly used model fit statistics into question (as described above). Therefore, we used Monte Carlo procedures (suggested by Langeheine et al., 1996) using the several measures of item association (suggested by de la Torre & Douglas, 2004) to evaluate the fit of our model. We chose two different measures of item association: the Pearson correlation and Cohen's κ, both chosen because of the understandable scale for the measure. For each measure, we computed the model-estimated values of the Pearson correlation and Cohen's κ for each item pair. We then computed the RMSE by taking the square root of the average squared difference for each item pair. The Monte Carlo procedure of resampling and estimating Pearson correlation and Cohen's κ was repeated 800 times. For our analysis, the Pearson correlation RMSE was found to be .042, with a Monte Carlo p value of .486, and the Cohen's κ RMSE was found to be .038, with a Monte Carlo p value of .504. The RMSE values indicate that there was not a great deal of discrepancy between the model-predicted and observed correlation and κ, and the moderate-sized p values indicated that what we had observed was not atypical for data simulated from our estimates (the p values indicate that our RMSE values for both statistics were near the center of the distribution of RMSE values observed through Monte Carlo simulation).

Diagnostic results. Recall that the DINO model provides a posterior probability that each criterion has been satisfied. In addition, we can compute the posterior probability that each individual has met at least 5 of the criteria: the probability that an individual meets the *DSM–IV–TR* definition for being considered a pathological gambler. From these probabilities, diagnostic assessments can be made: Individuals with a posterior probability greater than .50 are classified as probable pathological gamblers.

Diagnosis of pathological gambling, however, is not the only information ascertained by the cognitive diagnosis analysis. To demonstrate the depth of information provided by the cognitive diagnosis analysis, the estimates for three example individuals are given in Table 3. Specifically, Table 3 gives the cognitive diagnosis criteria profile results (the posterior probabilities of satisfying each criterion) for 3 individuals in addition to their posterior probabilities of being a pathological gambler. Note that because of the structure of the cognitive diagnosis model, the estimated posterior probabilities are a function of an individual's response pattern in addition to the estimated base rates for each criterion (the cut-point parameters and correlations of the structural model). In educational measurement, the cognitive diagnosis profile estimates are used to diagram a path for successful remediation of lacked academic skills. By analogy, in clinical diagnosis, the profile estimates would provide a clinician with information that might be useful in the treatment of the disorder.

In the example, Individual A has a model-estimated probability of being a pathological gambler equal to .00. However, through use of the DINO model, estimates of the probability of meeting each criterion provide additional insight to the individual's criteria profile. Clearly, Individual A is most likely not a pathological gambler; however, it is likely that Individual A satisfies 1 of the criteria (Criterion 6). With the knowledge of which criteria are met, more information is available to aid in the development of preventative actions to reduce the chances of this individual becoming a pathological gambler.

A similar result can be seen for Individual B. Individual B has a model-estimated posterior probability of being a pathological gambler equal to .33. In addition, it is clear that responses provided by Individual B were more representative of pathological gambling than were the responses provided by Individual A, in that the estimates indicate that Individual B most likely satisfied 4 criteria (Criteria 1, 2, 4,

[2] Note that .50 is not the only way that individuals may be classified. Other distinctions can be made to reflect the uncertainty of an estimate near .50. For instance, any individual with a posterior probability estimated between .40 to .60 could be said to be in an "indifference" region, whereas individuals with probabilities greater than .60 would be considered pathological, and individuals with probabilities less than .40 would be considered nonpathological.

Table 3
Individual Example Estimates

Criterion	Individual A	Individual B	Individual C
1	.00	**.99**	**1.00**
2	.01	**.99**	**1.00**
3	.00	.29	.16
4	.00	**.99**	**.99**
5	.00	**1.00**	**1.00**
6	**.77**	.02	.05
7	.11	.06	**.73**
8	.00	.00	.00
9	.00	.01	.02
10	.00	.00	.00
PPG	.00	.33	**.75**

Note. Posterior probabilities of satisfaction greater than .50, indicating that an individual likely satisfies a particular criterion, are presented in boldface. PPG = probability of pathological gambling.

and 5). Without the criteria-specific information provided by the cognitive diagnosis model, diagnosis of these individuals with respect to pathological gambling might not be as specific regarding how Individual A and Individual B differ in terms of satisfied criteria.

Individual C is classified as a pathological gambler (with posterior probability of .75) using the model-estimated posterior probability of pathological gambling produced by the analysis of the GRI with the DINO model. In addition, the criteria-level information also suggests that Individual C most likely satisfies Criteria 1, 2, 4, 5, and 7, meaning that any treatment efforts should be focused on these gambling tendencies. Treatments focusing on the nonsatisfied criteria may not be helpful for this individual.

From these examples, it can be seen that cognitive diagnosis models provide both general diagnostic information (pathological vs. nonpathological gambler) and detailed information about why each individual meets the definition for diagnosis. The criterion-level information gives insight for tailoring individual-specific treatments for pathological gambling, potentially increasing these treatments' effectiveness.

Item parameter results. Useful information detailing the diagnostic quality of each item was also obtained. Table 4 provides a summary of the information that was obtained from the parameters of the model. Recall that the model had two parameters per item, s_j and g_j. For each item, the odds ratio between responding positively conditional on $\omega_{ij} = 1$ and responding positively conditional on $\omega_{ij} = 0$ was computed. The odds ratio is a function of the estimated model parameters, $(1 - s_j)/s_j$ / $g_j/(1 - g_j)$. Those items where this ratio was largest were the most diagnostic in that, according to the model, there was a wide separation of scores between the two latent classes. If this difference was small, it would imply that the Q-matrix-specified criteria were not useful at differentiating between the two classes.

For example, it can be seen that many of the items were useful in discriminating between individuals in the two item-equivalence classes defined by the model (satisfying at least one necessary criterion; not satisfying any necessary criteria). Many odds ratios between the s and g parameters were very high. For instance, examine the results for Item 22, measuring Criterion 8 ("has committed illegal acts . . . to

Table 4
Estimated Item Parameters

Item	s_j (*SE*)	g_j (*SE*)	$\frac{(1 - s_j)/s_j}{g_j/(1 - g_j)}$
1	.26 (.02)	.13 (.03)	18.63
2	.41 (.03)	.34 (.03)	2.80
3	.18 (.04)	.13 (.03)	29.37
4	.39 (.03)	.56 (.02)	1.21
5	.24 (.03)	.05 (.01)	64.19
6	.26 (.04)	.13 (.02)	19.20
7	.12 (.02)	.10 (.02)	63.56
8	.54 (.02)	.01 (.01)	68.03
9	.21 (.03)	.03 (.01)	108.20
10	.21 (.03)	.04 (.01)	102.07
11	.40 (.03)	.01 (.01)	101.67
12	.45 (.02)	.53 (.02)	1.10
13	.35 (.03)	.01 (.01)	122.49
14	.36 (.03)	.04 (.01)	46.96
15	.29 (.04)	.03 (.02)	81.15
16	.49 (.03)	.02 (.01)	54.15
17	.33 (.03)	.03 (.01)	58.06
18	.08 (.03)	.24 (.02)	35.44
19	.11 (.03)	.43 (.03)	10.29
20	.64 (.02)	.34 (.02)	1.11
21	.42 (.03)	.01 (.01)	197.64
22	.10 (.03)	.24 (.02)	29.41
23	.25 (.03)	.04 (.01)	80.55
24	.30 (.03)	.09 (.02)	21.81
25	.11 (.04)	.31 (.02)	18.92
26	.17 (.02)	.35 (.03)	8.71
27	.23 (.03)	.05 (.02)	59.09
28	.44 (.05)	.01 (.01)	242.86
29	.34 (.03)	.03 (.02)	59.87
30	.39 (.06)	.04 (.01)	39.27
31	.65 (.03)	.01 (.01)	59.67
32	.20 (.04)	.16 (.03)	20.52
33	.38 (.06)	.01 (.01)	125.12
34	.54 (.03)	.03 (.01)	31.00
35	.66 (.03)	.01 (.01)	55.01
36	.57 (.04)	.01 (.01)	109.70
37	.36 (.05)	.01 (.01)	136.73
38	.51 (.05)	.01 (.01)	198.63
39	.33 (.04)	.11 (.02)	15.64
40	.19 (.05)	.11 (.01)	34.44
41	.28 (.03)	.13 (.02)	18.22

finance gambling") and Criterion 10 ("relies on others to provide money to relieve a desperate financial situation caused by gambling"):

22. Gambling has hurt my financial situation.

The item parameter estimates revealed that this item had a class odds ratio of 29.41, indicating that individuals who satisfy either Criterion 8 or Criterion 10 (or both) are approximately 29 times more likely to respond positively to this item.

Alternatively, three items for which little difference was found between the two class odd ratios were Items 4 (measuring Criterion 1: "is preoccupied with gambling"), 12 (measuring Criterion 6: "after losing money gambling, often returns another day to get even"), and 20 (measuring Criterion 6):

4. I enjoy talking with my family and friends about my past gambling experiences.
12. When I lose money gambling, it is a long time before I gamble again.
20. When gambling, I have an amount of money in mind that I am willing to lose, and I stop if I reach that point.

For these items, the class odds ratios all were slightly above a value of 1, meaning that little difference existed between the responses from those individuals who had satisfied the necessary criterion and those who had not satisfied the necessary criterion (class odds ratios of 1.21 for Item 4, 1.10 for Item 12, and 1.11 for Item 10). The results suggest that these items were poorly written, that the entries in the Q-matrix were incorrect, or that it did not make sense to ask these items of a college-aged population.

Structural Parameter Results

The estimated structural model was based on the model used by Feasel et al. (2004). In their development of the GRI, the authors used unidimensional sum scores of items to represent latent factor scores for each of the 10 *DSM–IV–TR* criteria. A structural equation model was then fit to the correlations of these sum scores, relating the 10 *DSM–IV–TR* criteria to a set of three latent factors. The hypothesized correlational structure mapped the latent criteria onto dependence (Criteria 1, 2, 4, and 5), lack of control (Criteria 3, 6, and 7), and disruption (Criteria 8, 9, and 10).

In our analysis, estimation of the three-factor correlational structure required a total of 14 correlational parameters (11 for the loadings of the criteria onto the three latent continuous factors and 3 for the continuous factor correlations). Additionally, the cognitive diagnosis structural model added a cut-point parameter (κ) for each of the 10 dichotomous criteria, for a total of 24 structural model parameters, which characterized the $P(c)$ term in Equation 1. The results of our analysis can be found in Figure 1. Note that the path diagram shown in the figure contains 10 circular nodes, each bisected by a line, representing the 10 latent dichotomous criteria. This diagram convention signifies the tetrachoric correlation concept of underlying continuous latent variables that are dichotomized. Aside from

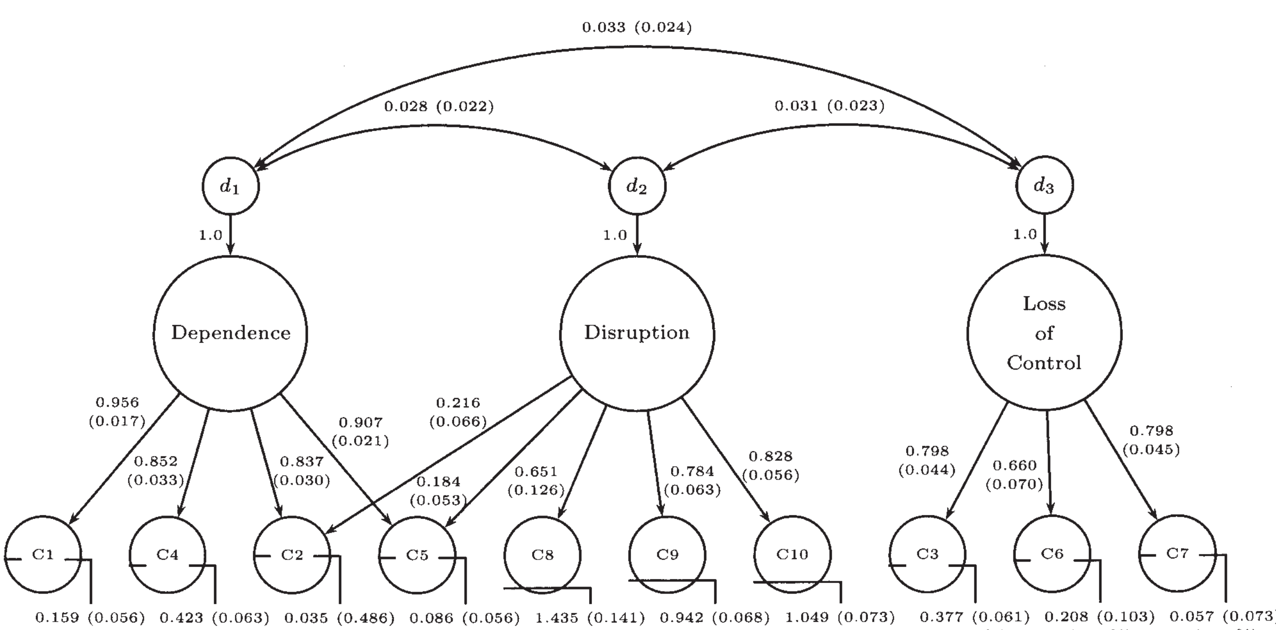

Figure 1. Gambling Research Instrument analysis structural model estimates. C = criterion.

this convention, the rest of the diagram is produced according to standard structural equation modeling diagram conventions (see McDonald & Ho, 2002).

Examination of the cut-point parameter estimates provided some interesting insights into the percentages of individuals satisfying each *DSM–IV–TR* criterion in our sample. These parameters reflect the estimated base rates for each of the criteria for the population of interest. As expected (because of the low percentage of pathological gamblers in our sample), the transformed cut-point estimates showed that the marginal percentage of individuals satisfying any given criterion was very low, with all values under 50%. The most frequently met criterion was Criterion 2 ("needs to gamble with increasing amounts of money in order to achieve desired excitement"), with an estimated 48.6% of our sample satisfying this criterion. To contrast this large value with a low estimate, the percentage of our sample that satisfied Criterion 8 ("has committed illegal acts . . . to finance gambling") was 7.8%. These estimates indicate that individuals often gamble as a way of escaping from problems but seldom commit illegal acts to finance their gambling.

We expect that these parameters may be helpful in several areas of diagnostic research: refinement of criteria used for diagnosis, development of scales to diagnose pathologies, and understanding the overall base rate of individuals satisfying each criterion in the population. Estimates that are extremely large (near 100%) or small (near 0%) may be indicators of a poor measure of the criterion. Furthermore, for disorders that are defined by summation across a set of satisfied criteria (such as pathological gambling, requiring 5 or more satisfied criteria), extreme results indicate that a criterion either plays a significant role in diagnosis (values near 100%) or plays no role at all (values near 0%). Additionally, if the base rates for the percentage of individuals satisfying each criteria are known, the cut-point parameters can be fixed to reflect these quantities.

The results for the structural model of the tetrachoric correlations yielded estimates similar in value to those found by Feasel et al. (2004). All factor loading parameters had estimates significantly different from 0, with most greater than 0.60 in magnitude. The moderate-sized loadings, however, occurred on parameters that had moderate-sized loadings reported by Feasel et al. (2004). The biggest difference from the previous structural analysis was for the nonsignificant loading of Criterion 2 onto the disruption factor (which was reported as not significantly different from 0 in Feasel et al., 2004). The large standard error of the loading of Criterion 8 onto the disruption factor may have been a result in part of the limited number of individuals in our sample or of the relatively low percentage of these individuals satisfying this criterion. Finally, the factor correlation parameters had low estimates, indicating the presence of continuous factors that are distinct in nature, having little overlap.

The ability to obtain factor structure estimates with cognitive diagnosis models provides avenues for research investigating both the diagnostic and etiological manifestations of pathology. The model used in our application provided a method of determining the magnitude of the correlational loading parameters and (although not performed) could provide guidance for adjusting the model on the basis of the agreement of empirical findings with theoretical expectations.

Concluding Remarks

The analysis presented in this article demonstrated the usefulness of cognitive diagnosis models as a method for jointly performing psychological assessment and clinical diagnosis while providing the ability to investigate structural facets of psychological traits and their relation to *DSM–IV–TR*-defined clinical criteria. The use of a cognitive diagnosis model, the DINO, allowed for an instrument that was created to investigate the structure of underlying personality factors in pathological gambling to provide diagnostic information for each criterion. Through understanding of the set of criteria an individual may have satisfied, in addition to the likelihood that an individual has satisfied more than 5 criteria (meeting the *DSM–IV–TR* definition of a pathological gambler), a path for treatment may be defined. Furthermore, this gain in information is coupled with the ability to estimate structural parameters indicating not only the percentage of individuals satisfying any given criterion but also the underlying factor structure of the traits fueling pathological gambling.

The models and example analysis presented in this article are not without limitations. To distinguish the limitations of cognitive diagnosis models in general from the limitations of our example analysis, we discuss each in the following separated sections.

Cognitive Diagnosis Model Limitations

Several important issues and potential limitations in the use of cognitive diagnosis models are in need of discussion. Primarily, the validity of any result using a cognitive diagnosis model rests on the theory that is driving the analysis. Specifically, if the definitions of the criteria are not carefully specified, or if any criteria are irrelevant to the diagnosis, the result will contain diagnostic information with little validity. For example, the definition of pathological gambling used in the demonstration analysis came from the *DSM–IV–TR*. If this definition is inaccurate or does not work empirically, the results are virtually meaningless. This concern, however, obtains not only for cognitive diagnosis models but, rather, for latent variable models in general. The use of diagnostic information provided by cognitive diagnosis models must be tempered with the knowledge that the result reflects the extent to which the driving theory is supported by reality.

Coupled with concerns regarding the validity of a cognitive diagnosis analysis result is the construction of the Q-matrix that ultimately sets the definition for the measured criteria. In practice, misguided Q-matrices lead to nonsensical results, reflected in DINA or DINO item parameters with inflated slip and guess parameters. Furthermore, poorly defined criteria or attributes can be detected through cutpoint parameters that indicate that almost every individual has met or (conversely) has not met a given criterion. Such a result has the effect of removing the latent attribute from the analysis entirely. Additionally, poorly created Q-matrices are also indicated by model fit indices with high RMSE values coupled with low Monte Carlo p values.

In effect, the quality of the cognitive diagnosis model fit rests on the quality of the Q-matrix used in the analysis. To this end, Q-matrix definition must blend the theoretical underpinnings of an instrument's domain with the empirical results of an analysis. It is acknowledged that often the exact specification of the Q-matrix is unknown. For this reason, empirically based Q-matrix discovery techniques are being pursued. For instance, both Templin and Henson (2006) and Henson and Templin (2006) have developed a Q-matrix discovery method that allows for uncertainty in inventory development by using subjective information of experts in a Bayesian algorithm. Ultimately, item and inventory construction should be conducted with a diagnostic model in mind (although we note that this was not the case for the GRI from our descriptive example). Indices conveying the diagnostic value of an item have been constructed and can be used in the inventory development process (Henson & Douglas, 2005). Through the use of such methods, construction of inventories providing quality diagnostic information can be achieved.

Perhaps most important to a cognitive diagnosis model result (and inventory) is the external verification of the diagnostic estimates. The development of an instrument that is to be used for diagnostic purposes must be conducted with the aid of measures of external validation. To exemplify this point, we cite Kikumi Tatsuoka, whose experience with diagnostic modeling comes from (among other procedures) her development of the rule space approach (K. Tatsuoka, 1983) in educational measurement: "In real estate, the three most important words are location, location, location. In diagnostic modeling, the three most important words are validation, validation, validation" (K. Tatsuoka, 2005). In inventory development, however, one would prefer that external validation come not only from pencil-and-paper-based instruments but also from diagnoses elicited from trained clinical professionals. External validation will give a degree of confidence to the interpretation of diagnostic estimates produced from a cognitive diagnosis model analysis.

Example Application Limitations

The demonstration example analysis of the GRI has several limiting factors. Our presentation of the example was included to demonstrate the types of information attainable from the analysis of an instrument with a cognitive diagnosis model. Our foremost concern was with the sample used in the analysis, 593 college students. Primarily, our sample was taken from an extremely narrow population, mainly freshman and sophomore students. Mechanisms of gambling pathology in the college-age population are not fully understood (and seem to be changing as a result of the prevalence of gambling opportunities now available to college students). Furthermore, the etiology of pathological gambling may be different for young adults (e.g., a different structural model or a differing set of diagnostic criteria may be needed). These concerns, however, are not necessarily detrimental to the analysis presented in this article; rather, they are part of the ongoing process of understanding the gambling tendencies of college students.

Further confounding our analysis was the use of data from an instrument developed to measure the latent criteria along a continuum rather than a taxonomy. The original reason for the development of the GRI was for purposes of studying the underlying factors in pathological gambling. Items were created with the intention of using a confirmatory factor model (with a simple structure) in the analysis. In cases of applications of cognitive diagnostic models to instruments created with continua in mind, care must be taken in evaluating the result. An analogous problem occurs frequently in educational measurement when diagnostic models are applied to tests that were developed for analysis by a unidimensional scale. Such tests are often created by culling items that give the maximum information with respect to a single latent continuum. Attempting to obtain classification results from a single latent continuum can result in estimates that a large number of individuals possess either all attributes or none of the attributes. In this case, all attribute correlations (from the structural model) would approach unity, an indication that a single continuum is truly underlying the data. In our example, however, the latent factor structure of the GRI allowed for 10 different latent factors to be used. The number of underlying factors in combination with the responses given to the items provided sufficient dimensionality for our analysis not to have produced a similar result.

It can be stated that the field of cognitive diagnosis models is relatively young, with large-scale assessments (in educational measurement) only recently investigated. Therefore, future directions for research into cognitive diagnosis models must be mentioned. Clearly, for cognitive diagnosis models to become widely used in psychology, models that incorporate differing types of data must be developed and implemented. Polytomous response models for categorical data have been developed (see von Davier & Rost, 1997), but further model development is needed to incorporate the multiple types of data commonly found in psychological assessment. Additionally, practical goodness-of-fit indices that sync with data typically collected in psychological assessment must be developed. Furthermore, estimation methods for cognitive diagnosis models have widely been based on MCMC algorithms that have not been typically expanded beyond the specific problem for which they were developed. It is our opinion that for cognitive diagnosis models to be a practical research tool, easily accessible estimation algorithms must be developed and offered to data analysts. As briefly mentioned, another topic in need of investigation is the development of empirical techniques for determining the entries of the Q-matrix. Often, Q-matrix entries are not obvious or do not show clear differences between classes (as with Items 4, 12, and 20 in our analysis). Techniques that allow the empirical data to mold the entries of the Q-matrix would provide helpful feedback for the construction of reliable instruments developed for use with cognitive diagnosis models.

This article has presented a review of the common cognitive diagnosis modeling framework, along with the development of a cognitive diagnosis model for application in psychological assessment. The type of information attainable by this type of model was demonstrated by the example application, simultaneously diagnosing individuals and studying the underlying factor structure of the specific disorder in question. The use of cognitive diagnosis models in psychology, however, is not limited only to the assessment and analysis of pathological gambling. Other psychological disorders have definitions in which diagnosis is implicated by a set of satisfied dichotomous criteria. Furthermore, we believe that cognitive diagnosis models may be able to augment research in fields such as personality, in which the dichotomous attribute nature of cognitive diagnosis models may help researchers to understand the fine-grained components of personality traits. Additionally, cognitive diagnosis models can play a role in assessment of individual's job-related abilities, which may be useful in fields such as industrial and organizational psychology. Finally, cognitive diagnosis models that allow for measurement of differences in attribute patterns cross-culturally can aide in each of these fields of psychology by detailing how cultural differences are manifested in the patterns of attributes individuals have mastered or the set of diagnostic criteria individuals have met. We see cognitive diagnosis models as a potentially useful statistical method for use in psychological research. It is our hope that the illustrations provided in this article will provide the groundwork for the expansion of the use of cognitive diagnosis models for inventory development in psychological assessment, research, and diagnostic applications.

References

Akaike, H. (1974). A new look at the statistical identification model. *IEEE Transactions on Automated Control, 19,* 716–723.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (text revision). Washington, DC: Author.

Barnard, J., McCulloch, R., & Meng, X.-L. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with applications to shrinkage. *Statistica Sinica, 10,* 1281–1311.

Bouwmeester, S., Sijtsma, K., & Vermunt, J. K. (2004). Latent class regression analysis for describing cognitive developmental phenomena: An application to transitive reasoning. *European Journal of Developmental Psychology, 1,* 67–86.

de la Torre, J., & Douglas, J. A. (2004). Higher order latent trait models for cognitive diagnosis. *Psychometrika, 69,* 333–353.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Statistical Methodology), 39,* 1–38.

DiBello, L., & Stout, W. (2003). Student profile scoring methods for information assessment. In H. Yanai, A. Okada, K. Shigemasu, & J. Meulmann (Eds.), *New developments in psychometrics* (pp. 81–92). Tokyo: Springer.

Eid, M., Langeheine, R., & Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis: A primer. *Journal of Cross-Cultural Psychology, 34,* 195–210.

Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P., & Johannesen, L. (1990). Introduction to knowledge spaces: How to build, test, and search them. *Psychological Review, 97,* 201–224.

Fals-Stewart, W., Birchler, G. R., Schafer, J., & Lucente, S. (1994). The personality of marital distress: An empirical typology. *Journal of Personality Assessment, 62,* 223–241.

Feasel, K., Henson, R., & Jones, L. (2004). *Analysis of the Gambling Research Instrument (GRI).* Unpublished manuscript.

Fink, P., Ørnbøl, E., Toft, T., Sparle, K. C., Frostholm, L., & Olesen, F. (2004). A new, empirically established hypochondriasis diagnosis. *American Journal of Psychiatry, 161,* 1680–1691.

Formann, A. K. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association, 87,* 476–486.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics 4: Proceedings of the Fourth Valencia International Meeting* (pp. 169–194). Oxford, England: Clarendon Press.

Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement, 26,* 333–352.

Hartz, S. M. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practice.* Unpublished doctoral dissertation, University of Illinois at Urbana–Champaign.

Henson, R., & Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement, 29,* 262–277.

Henson, R., & Templin, J. (2006, April). *Bayesian Q-matrix inference using the reparameterized unified model.* Paper presented at the annual meeting of the National Council for Measurement in Education, San Francisco, CA.

Hudziak, J. J., Wadsworth, M. E., Heath, A. C., & Achenbach, T. M. (1999). Latent class analysis of child behavior checklist attention problems. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 985–991.

John, O. P. (1990). The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25,* 258–272.

Karelitz, T. M. (2004). *Ordered-category attribute coding framework for cognitive assessments.* Unpublished doctoral dissertation, University of Illinois at Urbana–Champaign.

Langeheine, R., Pannekoek, J., & van de Pol, F. (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological Methods and Research, 24,* 492–516.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Boston: Houghton Mifflin.

Lesieur, H. R., & Blume, S. B. (1987). The South Oaks Gambling Screen (SOGS): A new instrument for the identification of pathological gamblers. *American Journal of Psychiatry, 144,* 1184–1188.

Macready, G. B., & Dayton, C. M. (1977). The use of probabilistic models in the assessment of mastery. *Journal of Educational Statistics, 2,* 99–120.

Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika, 64,* 197–212.

McArdle, J., & McDonald, R. (1984). Some algebraic properties of the reticular action model. *British Journal of Mathematical and Statistical Psychology, 37,* 234–251.

McDonald, R. P. (1999). *Test theory: A unified treatment.* Mahwah, NJ: Erlbaum.

McDonald, R. P., & Ho, M.-H. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7,* 64–82.

Meehl, P. E. (2001). Comorbidity and taxometrics. *Clinical Psychology: Science and Practice, 8,* 507–519.

Meehl, P. E., & Golden, R. (1982). Taxometric methods. In P. Kendall & J. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127–181). New York: Wiley.

Meehl, P. E., & Yonce, L. J. (1994). Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychological Reports, 74,* 1059–1274.

Meehl, P. E., & Yonce, L. J. (1996). Taxometric analysis: II. Detecting taxonicity with using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV procedure). *Psychological Reports, 78,* 1091–1227.

Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49,* 115–132.

Muthén, B. O. (1996). Psychometric evaluation of diagnostic criteria: Application to a two-dimensional model of alcohol abuse and dependence. *Drug and Alcohol Dependence, 41,* 101–112.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user’s guide.* Los Angeles: Muthén & Muthén.

Neuman, R. J., Todd, R. D., Heath, A. C., Reich, W., Hudziak, J. J., Bucholz, K. K., et al. (1999). Evaluation of ADHD typology in three contrasting samples: A latent class approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 25–33.

Read, T. R. C., & Cressie, N. A. C. (1988). *Goodness-of-fit statistics for discrete multivariate data.* New York: Springer-Verlag.

Reise, S. P., & Gomel, J. N. (1995). Modeling qualitative variation within latent trait dimensions: Application of mixed-measurement to personality assessment. *Multivariate Behavioral Research, 30,* 341–358.

Rohde, L. A., Barbosa, G., Polanczyk, G., Eizirik, M., Rasmussen, E. R., Neuman, R. J., & Todd, R. D. (2001). Factor and latent class analysis of *DSM–IV* ADHD symptoms in a school sample of Brazilian adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 711–718.

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement, 44,* 75–92.

Schwarz, G. (1976). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Solomon, A., Haaga, D. A. F., & Arnow, B. A. (2001). Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research. *Journal of Nervous and Mental Disease, 189,* 498–506.

Stout, W. (2002). Psychometrics: From practice to theory and back. *Psychometrika, 67,* 485–518.

Tatsuoka, C. (2002). Data analytic methods for latent partially ordered classification models. *Journal of the Royal Statistical Society, Series C (Applied Statistics), 51,* 337–350.

Tatsuoka, K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement, 20,* 345–354.

Tatsuoka, K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), *Diagnostic monitoring of skill* *and knowledge acquisition* (p. 453–488). Mahwah, NJ: Erlbaum.

Tatsuoka, K. (2005, April). *Rule space.* Training session presented at the annual meeting of the National Council on Measurement in Education, Montreal, Quebec, Canada.

Tellegen, A., & Waller, N. G. (1994). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In S. R. Briggs & J. M. Cheek (Eds.), *Personality measures: Development and evaluation* (Vol. 1, pp. 133–161). Greenwich, CT: JAI Press.

Templin, J. (2004). *Generalized linear mixed proficiency models for cognitive diagnosis.* Unpublished doctoral dissertation, University of Illinois at Urbana–Champaign.

Templin, J., & Henson, R. (2006, April). *A Bayesian method for incorporating uncertainty into Q-matrix estimation in skills assessment.* Paper presented at the annual meeting of the National Council for Measurement in Education, San Francisco, CA.

von Davier, M., & Rost, J. (1997). Self monitoring—A class variable? In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 296–305). Münster, Germany: Waxmann.

von Davier, M., & Yamamoto, K. (2004). Partially observed mixtures of IRT models: An extension of the generalized partial-credit model. *Applied Psychological Measurement, 28,* 389–406.

Waller, N. G., & Meehl, P. E. (1997). *Multivariate taxometric procedures: Distinguishing types from continua.* Thousand Oaks, CA: Sage.

Yamamoto, K., & Everson, H. (1997). Modeling the effects of test length and test time on parameter estimation using the HYBRID model. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 89–98). Münster, Germany: Waxmann.

Received November 16, 2004
Revision received December 5, 2005
Accepted April 13, 2006 ■

Correction to Hipp and Bauer (2006)

In the article “Local Solutions in the Estimation of Growth Mixture Models,” by John R. Hipp and Daniel J. Bauer (*Psychological Methods*, 2006, Vol. 11, No. 1, pp. 36–53), when discussing the start value algorithm in Mplus 3 (p. 50), the authors indicated that the start values for the latent class probabilities are unperturbed. In fact, except for variance or covariance parameters, start values for all parameters are perturbed, including the class probabilities. Additionally, the | syntax option calls a different start value algorithm than the one obtained with the BY syntax option that was described. This alternate algorithm perturbs parameters differently as a function of their standard deviations but was undocumented at the time the article was published. It is now described at the Mplus Web site (www.statmodel.com). The authors thank Bengt Muthén for calling this additional algorithm to their attention.