PSYCHOMETRIKA—VOL. 74, NO. 2, 191–210
JUNE 2009
DOI: 10.1007/S11336-008-9089-5

DEFINING A FAMILY OF COGNITIVE DIAGNOSIS MODELS USING LOG-LINEAR MODELS WITH LATENT VARIABLES

ROBERT A. HENSON

THE UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

JONATHAN L. TEMPLIN

UNIVERSITY OF GEORGIA

JOHN T. WILLSE

THE UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

This paper uses log-linear models with latent variables (Hagenaars, in Loglinear Models with Latent Variables, 1993) to define a family of cognitive diagnosis models. In doing so, the relationship between many common models is explicitly defined and discussed. In addition, because the log-linear model with latent variables is a general model for cognitive diagnosis, new alternatives to modeling the functional relationship between attribute mastery and the probability of a correct response are discussed.

Key words: cognitive diagnosis models, log-linear latent class models, latent class models.

1. Introduction

Over the past several decades, research in educational assessment has led to expanded models that provide a profile defining mastery or nonmastery of a set of predefined skills or attributes. These models, commonly called cognitive diagnosis models (CDMs) define an individual's ability based on the attributes that have or have not been mastered. Given this mastery profile, the probability of a correct response is defined by mastery/nonmastery of the skills that are required by an item. As opposed to more common models such as unidimensional item response theory (IRT) and classical test theory (CTT), CDMs promise an explanation of why an individual is not performing well based on those skills that have not been mastered. IRT and CTT provide a continuous measure of ability and, therefore, only provide a rank ordering of examinees. Any diagnostic information (mastery/nonmastery) must be obtained using additional analyses. This paper seeks to define a family of models, additionally describing the relationship between many common models that have been developed for cognitive diagnosis.

As they are currently defined, CDMs focus primarily on determining the set of skills that an individual has or has not mastered. This focus on educational achievement testing has led some to employ the label "skills assessment." However, stronger links to cognitive theory are possible. Where we define cognitive theory as describing the process of how the skills combine to produce item response behavior. These links include evaluations of theories based on the "fit" of specific CDM models. Where appropriate, connections to theory, evaluation, and development are addressed.

Within the CDM literature, assumptions about how skills (or cognitive/psychological processes) influence test performance are operationalized through selection of a specific CDM.

Requests for reprints should be sent to Robert A. Henson, The University of North Carolina at Greensboro, Greensboro, USA. E-mail: rahenson@uncg.edu

One of the largest differences is between those CDMs that are described as noncompensatory models (including conjunctive and disjunctive models) versus those models that are described as compensatory models. Of the noncompensatory models, conjunctive models are typically thought to be models for which one cannot "make up" for nonmastery of attributes by mastery of other attributes. To perform reasonably well on any item under the assumption of a conjunctive model, one must know all required attributes. Lacking even a single required attribute can dramatically decrease the probability of a correct response. Math tests are perhaps the best examples where a set of skills are all required to perform well on an item (e.g., mixed fraction subtraction, Tatsuoka, 1990), and if even one has not been mastered the item will most likely be missed, but other examples include verbal tasks such as verbal classification and synonym items (Embretson, 1985; Janssen & DeBoeck, 1997) in addition to the Raven's Progressive Matrices (Fischer & Forman, 1982).

Disjunctive models are an alternative set of noncompensatory models, which can be considered as the "opposite" of conjunctive models. Specifically, disjunctive models define the probability of a correct response such that mastering a subset (in some cases only one) of the attributes is sufficient to have a high probability of a correct response. Therefore, an examinee mastering all of the required attributes is expected to perform in a similar way as an individual mastering only a subset of the required attributes. Such models have been shown to be useful in psychology (Templin & Henson, 2006), but could also be useful in education when items have more then one strategy that could be used to obtain a correct answer. In that case, each strategy may require a different skill and thus mastery of one *or* the other skill could still result in a high probability of a correct response.

Unlike noncompensatory models, compensatory models have been described (using continuous variables) as allowing an individual to "make up" for what is lacked in one skill by having mastered another. However, we note that this terminology can be somewhat confusing given that a similar statement could be said of disjunctive models, therefore, a different terminology is used to discuss the differences between compensatory models and noncompensatory models in CMDs. Specifically, we note that another way of discussing the difference between compensatory and noncompensatory models is based on the relationship between any set of attributes (e.g., the two attributes A_1 and A_2) and the item (X). A compensatory model can be defined as a model such that the conditional relationship between any attribute and the item response does not depend on mastery or nonmastery of the remaining required attributes of that item (using notation of a model based on log-linear models this can be described as the model $\{A_1X, A_2X, A_1A_2\}$, and thus only two interactions of each attribute with an item are needed). In contrast, the conditional relationship between any attribute and the item response does depend on mastery or nonmastery of the remaining attributes for noncompensatory models (again using notational short hand of the log-linear model, the model $\{A_1A_2X\}$ must be true, and thus a three-way interaction is needed). Later in this manuscript a more detailed comparison of the distinctions between these models will be discussed by using the log-linear model with latent variables.

Even within the breadth of noncompensatory and compensatory models, a number of separate models has been developed. For example, common conjunctive models include the Deterministic Input; Noisy "And" gate (DINA; Junker & Sijtsma, 2001), Noisy Input; Deterministic "And" gate model (NIDA, Junker & Sijtsma, 2001), and the Reparameterized Unified Model (RUM, Hartz, 2002), whereas examples of disjunctive models include the Deterministic Input; Noisy "Or" gate model (DINO, Templin & Henson, 2006). In addition, two of the most common models that have been used as compensatory models are the General Diagnostic Model (GDM, von Davier, 2005) and a special case of the GDM called the compensatory RUM (Hartz, 2002).

Because each model has different assumptions and levels of complexity (e.g., number of parameters), one must first determine the appropriate model before completing an analysis. Little has been done to explicitly express the differences of each model and directly parameterize these differences. Currently, one of the most general models was developed by von Davier (2005). In his discussion of the general diagnostic model, he provides a general parameterization that because of its link to log-linear models with latent variables, allows for most cognitive diagnosis models. However, in his discussion, he provides no clear link between popular models in the literature. Instead, von Davier (2005) focuses on even more general situations with little to no mention of the models that are the focus of this paper.

The purpose of this paper is to first discuss several common models for cognitive diagnosis and then briefly discuss log-linear models with latent variables (Haberman, 1979; Hagenaars, 1993). Then we will show that by using a specific form of these models where both attributes and items are assumed to be dichotomous, the log-linear model with latent variables for cognitive diagnosis (called LCDM) defines a full continuum of models that can be expressed easily ranging from fully disjunctive models (i.e., the DINO) to fully conjunctive models (i.e., the DINA). With this parameterization, a better understanding of these models and their differences is provided. In addition, analyses will not be constrained to the models that have currently been developed, but instead can be determined based on underlying theory of the construct or through exploratory analyses. Finally, an example is provided where the full LCDM is estimated, which provides empirical information as to the appropriate model for *each* item.

1.1. Current Cognitive Diagnosis Models

As was briefly mentioned, CDMs define the probability of a correct response based on an attribute mastery profile. Because there are a finite set of attribute mastery profiles, CDMs are similar to latent class models. Specifically, CDMs are special cases of constrained latent class models where classes are defined by the attribute mastery profiles (for a general progression to CDMs, see Macready & Dayton, 1977; Rindskopf, 1983; Haertel, 1989). Thus, CDMs assume that all individuals mastering the same set of required attributes for any given item will have the same probability of a correct response for an item (i.e., probabilities are constrained to be equal across all examinees who have the same profile for attributes required by an item).

Although this assumption is somewhat restrictive, CDMs have shown promise in educational and psychological assessment and it is believed that their applications will continue to expand. For example, Tatsuoka (1990) and others (e.g., de la Torre & Douglas, 2004) have discussed the use of CDMs to assess mixed number subtraction (a data set that is later used as an example in this paper). In this case, it was argued that there are certain "skills" such as finding a common denominator that a student either knows or does not know. Another use of cognitive diagnosis models in education has been in language assessment among which one of the most complete studies was completed by Jang (Jang, 2005). Using two prototype tests of the Next Generation Test of English as a Foreign Language (NG TOEFL), developed by the Educational Testing Service, Jang seeks to first determine a reasonable set of dichotomous attributes and then validate the use of a CDM by providing teachers with diagnostic information about their students. Jang (Jang, 2005) was able to determine a set of nine attributes such as "Analyze and evaluate relative importance of information in the text by distinguishing major ideas from supporting details." In addition, Jang showed a measurable increase in the skills when teachers are informed of students' specific deficits. Finally, CDMs can be useful in the assessment of psychological disorders. Templin and Henson (2006) discussed the evaluation of pathological gambling based on identifying the set of 10 dichotomous criteria that are met by an individual. Currently, many disorders are identified by a set of dichotomous criteria are either met or not met. Although these are a few examples of applications where CDMs have been shown to be useful, many more situations in the social sciences could benefit from these models because of the more detailed diagnostic information that is provided.

As was the case in the examples previously discussed and in all models discussed in this paper, a Q-matrix, $\mathbf{Q}$ is required by the analysis. The Q-matrix is an item by attribute indicator matrix that defines which attributes must be mastered to have a high probability of a correct response. Specifically, for each element, $q_{jk} = 1$ if the jth item measures (and thus requires) the kth attribute and $q_{jk} = 0$ if the kth attribute is not required for the jth item. Notice that in specifying the q-matrix, it is implicitly assumed that the set of skills used for correctly answering each item can be determined. In addition, another assumption made by the q-matrix is that one and only one strategy is being used by students, which is one limitation suffered by many CDMs (for an example of a multiple strategy CDM, see de la Torre & Douglas, 2004).

By defining which items measure which attributes, the q-matrix explicitly defines all attributes (in the context of the test and the model). Any change in the q-matrix or model redefines, at least slightly, the interpretations of the set of user-specified attributes, *even if* we keep their substantive labels the same. Therefore, it is important that the construction of the q-matrix for a specific model be done with care. Ironically, while the development of the q-matrix and the identification of the model are among the most critical, as well the most challenging steps in a CDM analysis, most CDMs assume that the q-matrix is given and favor a specific model based partially on availability of software. Because of this, great effort must go in to the development of the Q-matrix based on a well-developed theory.

In addition to the Q-matrix, all models in this paper assume that an examinee's ability is characterized by a mastery profile $\boldsymbol{\alpha}_i$. Where the ith examinee's mastery profile, $\boldsymbol{\alpha}_i$, is a vector of length K (where K is the total number of attributes) indicating which attributes have been mastered. Specifically, $\alpha_{ik} = 1$ if the kth attribute has been mastered by the ith examinee and $\alpha_{ik} = 0$ if the kth attribute has not been mastered.

Given the Q-matrix and an examinee's attribute pattern, CDMs define the probability of a correct response. Cognitive diagnosis models differ in how the probability of a correct response is defined. Therefore, the following subsections describe models classified as noncompensatory and compensatory models.

1.2. Noncompensatory Models

Recall that noncompensatory models are defined as models where the conditional relationship between any attribute and the item responses *depend* on the remaining required attributes that have been mastered or not. Because of the nature of this dependency, noncompensatory models can be divided into conjunctive and disjunctive models.

1.2.1. Conjunctive Models Among some of the simplest conjunctive models is the DINA (Deterministic Input; Noisy "And" Gate, Haertel, 1989; Junker & Sijtsma, 2001) model. For each item, the DINA model divides examinees into two classes, those who have mastered all required attributes and those who have not that are indicated by the latent variable, ξ_{ij}. The value ξ_{ij} equals one for those examinees who have mastered all required attributes and is zero otherwise. Specifically, for examinee i for item j,

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

Once the value of ξ_{ij} is known, the probability of a correct response for examinee i for correctly responding to item j is defined by s_j and g_j, where s_j is the probability of incorrectly answering an item when in fact all required attributes have been mastered (or "slipping") and g_j is the probability of correctly responding to an item when all required attributes for that item have not been mastered (a "guessing" parameter).

$$s_j = P\big(X_{ij} = 0 | \xi_{ij} = 1\big), \tag{2}$$

$$g_j = P\big(X_{ij} = 1 | \xi_{ij} = 0\big). \tag{3}$$

Given the jth item's parameters and ξ_{ij}, the probability of a correct response can be written as

$$P(X_{ij}=1|\xi_{ij}) = (1-s_j)^{\xi_{ij}} g_j^{(1-\xi_{ij})}. \tag{4}$$

The DINA makes one additional constraint that $(1-s_j) > g_j$ and, therefore, an examinee mastering all required attributes, $\xi_{ij}=1$, has a higher probability of a correct response than an examinee who has not mastered all required attributes, $\xi_{ij}=0$.

Notice that by defining the DINA model in this way, the probability of a correct response is only high when an examinee has mastered all required attributes. If only a subset of the required attributes have been mastered, the probability of a correct response is expected to be low, meaning that a positive relationship between any required attribute and the item exists only if all other required attributes have been mastered, otherwise mastery of any required attribute is independent of the item response.

The DINA model has been considered too restrictive in that all examinees who lack at least one required attribute are assumed to have the same probability of a correct response, g_j. One model that addresses this concern is a reduced[1] version of the Reparameterized Unified Model (RUM; Hartz, 2002). Given an examinee's $\boldsymbol{\alpha}_i$, the reduced RUM defines the probability of a correct response as:

$$P(X_{ij}=1|\boldsymbol{\alpha}_i) = \pi_j^* \prod_{k=1}^{K} r_{jk}^{*\,q_{jk}(1-\alpha_{ik})}. \tag{5}$$

Here, the item parameter π_j^* is defined as the probability of a correct response to item j assuming that all required attributes have been mastered. The r_{jk}^* parameters are constrained such that $0 \le r_{jk}^* \le 1$, and indicate the proportional amount that the probability of a correct response to item j is reduced if the kth required attribute ($q_{jk}=1$) has not been mastered. Unlike the DINA, when using the reduced RUM, the probability of a correct response decreases for each attribute that has not been mastered.

As was the case when using the DINA, the conditional relationship between any attribute and an item response is smallest when all other required attributes are not mastered. This conditional relationship (i.e., which defines the difference of the log-odds comparing masters of a specific attribute to nonmasters) is largest when all other attributes have been mastered.

1.2.2. Disjunctive Models As an alternative to conjunctive models, disjunctive models assume that mastery of additional attributes provide little to no gain once a subset of the required attributes have been mastered. Templin and Henson (2006) describe a disjunctive model called the DINO (Deterministic Input; Noisy "Or" Gate) model. The DINO, much like the DINA, models the probability of a correct response as a function of a slipping parameter, s_j, and a guessing parameter, g_j. However, instead of defining ξ_{ij} they use the parameters ω_{ij}. The latent variable ω_{ij} is defined as

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

which is an indicator of whether the ith examinee has mastered *at least* one of the required attributes for the jth item. Therefore, $\omega_{ij}=1$ for any examinee having mastered one or more of the item's required attributes and $\omega_{ij}=0$ for an examinee who has not mastered any of the

[1]The full model includes a continuous latent variable in addition to the dichotomous latent variables. However, to define a family of cognitive diagnosis models, we currently only address those models that are constrained latent class models without any additional continuous variables, and thus use the reduced version of the RUM.

required attributes. Given ω_{ij} the probability of a correct response is defined as:

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

As in the DINA, examinees are divided into two groups, however, once one of the required attributes has been mastered the probability of a correct response increases from g_j to $(1 - s_j)$, where $g_j < (1 - s_j)$. The probability of a correct response does not increase for examinees who have mastered more than one of the required attributes, and thus the conditional relationship between any required attribute and the item response is zero given that at least one other required attribute has been mastered.

1.3. Compensatory Models

As was mentioned previously, compensatory models are defined such that the conditional association of any required attribute and an item does not depend on mastery of any other required attributes and, therefore, the increase of the log-odds of a correct response when comparing a nonmaster to a master is constant across all other levels mastery and nonmastery of the other required attributes.

One of the simplest examples of a compensatory model is the compensatory version of the RUM (Compensatory RUM; Hartz, 2002), which is a special case of the GDM (von Davier, 2005). The compensatory RUM is defined as:

$$P(X_{ij} = 1|\boldsymbol{\alpha}) = \frac{\exp(\sum_{k=1}^{K} r_{jk}^* \alpha_{ik} q_{jk} - \pi_j^*)}{1 + \exp(\sum_{k=1}^{K} r_{jk}^* \alpha_{ik} q_{jk} - \pi_j^*)}, \tag{8}$$

where all $r_{jk}^* > 0$. Notice that in the compensatory RUM, the lowest probability is defined as a function of only $-\pi_j^*$ (similar to a guessing parameter). The probability of a correct response is then increased as a function of each required attribute that is mastered (defined by r_{jk}^*). Therefore, the relationship between any attribute and item performance is not conditional on the remaining required attributes.

1.4. Summary of Cognitive Diagnosis Models

A number of cognitive diagnosis models have been defined, each with their own benefits and limitations and although we have described the differences based on the conditional relationships of each attribute with each item, little is understood as to the true differences between each of these models. In defining a model that unifies many of the previously defined models, a general approach to cognitive diagnosis is given, which will also quantify the true differences between these models. In the following section, a general description of the model is provided and its relationship to each of the previously defined models is discussed.

2. Log-Linear Models with Latent Variables

We begin by briefly discussing the basic log-linear model and its extension to include latent variables which was initially discussed and estimated by Haberman (1974, 1979) and later discussed by Hagenaars (1993). To better discuss the log-linear model, we use a basic example using three discrete variables X, Y, and Z. The basic log-linear model is used to evaluate relationships between X, Y, and Z, where X has categories $i = \{0, 1\}$, Y has categories $j = \{0, 1\}$, and Z has categories $k = \{0, 1\}$. We assume that n independent observations have been collected and for each observation, $X = i$, $Y = j$, and $Z = k$ are observed. One way to summarize the information of the n observations is using a 3-way contingency table. The log-linear model is used to model the frequency (F_{ijk}) of any cell. Specifically, using the full saturated model to predict the frequency of all $2 \times 2 \times 2 = 8$ cells F_{ijk} is defined as

$$\log F_{ijk} = \eta + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ij}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}. \tag{9}$$

Using the full saturated model, all cells are perfectly predicted in that there are also eight parameters. In addition, each λ represents a change in the log F_{ijk}. For example, λ_i^X represents the change in log F_{ijk} of being at category i for the variable X when compared to a defined reference group. Whereas λ_{ij}^{XY} represents the added change in the log F_{ijk} of being in the ij category of variables X and Y. One way of identifying this model is based on a reference coding such that one determine category of any variable is set to zero (this method is assumed throughout the rest of the paper, where the reference group contains those individuals who have not mastered any of the required attributes). For a more detailed description of the log-linear model, see Agresti (1990) and Hagenaars (1993).

The log-linear model has also been extended such that latent variables can be included in the model. Specifically, using the previous example with X, Y, and Z it may be hypothesized that the only relationships between X, Y, and Z are due to a set of latent variables. Hagenaars (1993) discusses a model such that these latent variables are discrete (see also Haberman, 1979). For the purpose of this paper, a discrete dichotomous latent variable will be indicated as α_u, where α_u is a 0/1 latent variable. In this case, the following model could be used.

$$\log F_{uxyz} = \eta + \lambda_u^\alpha + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ui}^{\alpha X} + \lambda_{uj}^{\alpha Y} + \lambda_{uk}^{\alpha Z}. \tag{10}$$

In this model, the variables X, Y, and Z are assumed to be conditional independent given α, however, α is related to each of X, Y, and Z. It should be noted that while this example only contains a single latent variable, the models easily extend to multiple discrete latent variables.

3. The Log-Linear Cognitive Diagnosis Model (LCDM)

The log-linear model with latent variables is a flexible model that allows the relationships between categorical variables to be modeled using a latent class model. For this reason, it easily generalizes to applications for cognitive diagnosis models (Fu, 2005; von Davier, 2005). In addition, because most cognitive diagnosis models are typically parameterized to define the probability of a correct response (i.e., each item is either correct $X_{ij} = 1$ or incorrect $X_{ij} = 0$) the log-linear model is re-expressed in terms of the log-odds of a correct response for each item (as a function of the latent variables). We note that this is not necessary, but defining the probability of a correct response is consistent with the literature for cognitive diagnosis models.

Currently, two different models have been discussed using the log-linear model with latent variables as models for cognitive diagnosis. Fu (2005) provides a complex model that includes the log-linear model as a submodel. In addition, von Davier (2005) discusses the General Diagnostic Model (GDM) as a general approach to log-linear models with latent variables, where the latent variables are both continuous and discrete in addition to focusing on ordered responses for items. Currently, the GDM has only been applied as a compensatory cognitive diagnosis model and discussed as a mixed Rasch model. However, as a special case, the GDMs general definition easily incorporates log-linear models with dichotomous latent variables for dichotomous responses (LCDM), and for that reason we use the notation of the GDM (von Davier, 2005) to define the LCDM. Specifically, the probability of a correct response is defined as:

$$P(X_{ij} = 1|\boldsymbol{\alpha}_i) = \frac{\exp(\boldsymbol{\lambda}_j^T \mathbf{h}(\boldsymbol{\alpha}_i, \mathbf{q}_j) - \eta_j)}{1 + \exp(\boldsymbol{\lambda}_j^T \mathbf{h}(\boldsymbol{\alpha}_i, \mathbf{q}_j) - \eta_j)}, \tag{11}$$

where the vector $\boldsymbol{\lambda}_j$ represents a $1 \times (2^K - 1)$ vector of weights for the jth item and $\mathbf{h}(\boldsymbol{\alpha}_i, \boldsymbol{q}_j)$ represents a set of linear combinations of the $\boldsymbol{\alpha}_i$ and the Q-matrix entries for the jth item, q_j.

The value η_j is used to define the probability of a correct response for the reference group, which is defined as those individuals who have not mastered any of the attributes. For the LCDM, $\mathbf{h}(\boldsymbol{\alpha}_i, \boldsymbol{q}_j)$ is defined as the set of all weights included in the full log-linear model with K latent dichotomous attributes. Using the LCDM, $\boldsymbol{\lambda}_j^T \mathbf{h}(\boldsymbol{\alpha}_i, \boldsymbol{q}_j)$ can be written as:

$$\boldsymbol{\lambda}_j^T \mathbf{h}(\boldsymbol{\alpha}_i, \boldsymbol{q}_j) = \sum_{u=1}^{K} \lambda_{ju}(\alpha_u q_{ju}) + \sum_{u=1}^{K} \sum_{v>u} \lambda_{juv}(\alpha_u \alpha_v q_{ju} q_{jv}) + \cdots . \tag{12}$$

Therefore, the conditional relationship between mastery or nonmastery of the uth attribute for the jth item is related to λ_{ju}, where $e^{\lambda_{ju}}$ describes the factor by which the odds of a correct response changes when comparing a nonmaster to a master given that all other attributes have not been mastered (i.e., the reference group represents those individuals who have not mastered any of the attributes). The extent to which the conditional relationship of attribute u and the item depends on a second attribute v for the jth item is defined by λ_{juv}. Thus, given that attribute v has been mastered, the odds of a correct response increases by a factor of $e^{\lambda_{ju}+\lambda_{juv}}$ when comparing nonmasters to masters, which is different from $e^{\lambda_{ju}}$ given attribute v has not been mastered. Such a model can be expanded to include all possible conditional relationships. In addition, we note that the notation has slightly changed from the log-linear model. For example, $\lambda_{ij}^{\alpha_1\alpha_2}$ would have been used to indicate the weight for level i of a discrete latent variable α_1 and level j for a discrete latent variable α_2. However, because all latent variables for the cognitive diagnosis models discussed in this paper have only two levels, λ_{j12} is used to indicate the weight that is added when $\alpha_1 = 1$ and $\alpha_2 = 1$. In all other instances, this weight would be zero.

A number of constraints are included to ensure identifiability of the LCDM. One set of constraints is determine by specification of the Q-matrix. Notice that in identifying a Q-matrix the analysis is comparable to a confirmatory analysis in that the definition of attributes is identified by those items that require each attribute. Without the Q-matrix, attributes could alternate in their definition, much like the rotational indeterminacy that can occur in exploratory factor analysis. A second set of constraints must be defined to ensure "monotonicity." For cognitive diagnosis models, monotonicity is defined as the property such that for any examinee that masters additional skills his or her probability of a correct response must be equal to or greater than the probability of a correct response prior to learning the additional skills. Specifically, monotonicity is defined as

$$P\big(X_{ij} = 1 | \boldsymbol{\alpha}_i^w\big) \geq p\big(X_{ij} = 1 | \boldsymbol{\alpha}_i\big) \quad \text{for all } w \tag{13}$$

where

$$\alpha_{ik}^w = \begin{cases} \alpha_{ik}, & \text{where } w \neq k, \\ 1 & \text{otherwise.} \end{cases}$$

One last constraint is based on the fact that attributes and Q-matrix entries are defined as 0/1 (although this is not necessarily required by the model). By requiring this constraint, a reference group is identified as those individuals who have not mastered any of required attributes for an item. Thus, identifying the probability of a correct response for those individuals who have not mastered any of the required attributes as the logit$(-\eta)$.

To aid in the clarity of the LCDM, we define the parameters for an item that requires two attributes (attribute one and attribute two), thus the probability of a correct response for this item using the LCDM is defined as

$$P\big(X_{ij} = 1 | \boldsymbol{\alpha}\big) = \frac{e^{\lambda_{j1}\alpha_1 + \lambda_{j2}\alpha_2 + \lambda_{j12}\alpha_1\alpha_2 - \eta_j}}{1 + e^{\lambda_{j1}\alpha_1 + \lambda_{j2}\alpha_2 + \lambda_{j12}\alpha_1\alpha_2 - \eta_j}}. \tag{14}$$

In doing so, we will show that any model can be fit by constraining a set of the LCDM parameters, which provides a method for model comparison. In addition, a theoretical "family" of cognitive diagnosis models is defined allowing for descriptions of additional models that have not yet been defined while providing a natural definition of compensatory, conjunctive, and disjunctive models.

3.1. Compensatory RUM

In many ways, the simplest relationship between the LCDM and any of the models is how a compensatory RUM can be defined as a reduced version of the LCDM. Specifically, the compensatory RUM is defined such that the conditional relationship between any required attribute and the item is fixed across all levels of the remaining required attributes. Using the example of the item which only requires attribute one and attribute two (defined in (14)), the compensatory RUM is defined by simply setting $\lambda_{j12} = 0$. Thus, the odds of a correct response will increase by a factor of $e^{\lambda_{j1}}$ when comparing a nonmaster to an individual who has mastered attribute one, which does not depend on mastery or nonmastery of attribute two. Notice that a similar strategy can be used if an item where to measure three or more attributes.

3.2. DINA

The LCDM can also fit many of the models that would typically be referred to as conjunctive models and, therefore, the difference when selecting a conjunctive model over a compensatory model can be described. One of the CDMs with the fewest number of parameters is the DINA. Recall that the DINA only defines two parameters for each item: a guessing parameter (g_j) and a slipping parameter (s_j). In this case, the probability of a correct response is equal to g_j unless all required attributes have been mastered in which case the probability of a correct response increases to $1 - s_j$. In the context of the LCDM, the same functional relationship can be described as only having a positive conditional relationship between a required attribute and the item when all other attributes have been mastered. All conditional relationships between an attribute and the item, given at least one attribute has not been mastered are set to zero. Specifically, there is no gain in the probability of a correct response (or odds of a correct response) for only knowing a subset of the attributes, but only when all required attributes have been mastered. Using the previous example for a two attribute item, the reduced LCDM is defined as

$$P\left(X_{ij} = 1|\boldsymbol{\alpha}\right) = \frac{e^{(0)\alpha_1 + (0)\alpha_2 + \lambda_{j12}\alpha_1\alpha_2 - \eta_j}}{1 + e^{(0)\alpha_1 + (0)\alpha_2 + \lambda_{j12}\alpha_1\alpha_2 - \eta_j}} \tag{15}$$

where the zeros have been added to emphasize the constraint that $\lambda_{j1} = 0$ and $\lambda_{j2} = 0$ and $\lambda_{i12} > 0$. Notice that if attribute two has not been mastered then there is no relationship between attribute one and the item (the odds of a correct response increases by a factor of $e^{\lambda_1} = e^0 = 1$ when comparing a nonmaster to a master). However, if attribute two has been mastered, the conditional relationship between attribute one and the item increases (the odds of a correct response *increases* by a factor of $e^{\lambda_1 + \lambda_{12}} = e^{0 + \lambda_{12}} > 1$ when comparing a nonmaster to a master). Because $\lambda_{i12} > 0$, we would characterize this model as a conjunctive model.

Defining the LCDM parameters for the example as a function of the DINA parameters:

$$\eta_j = -\ln\left(\frac{g_j}{1 - g_j}\right) \tag{16}$$

and

$$\lambda_{j12} = \eta + \ln\left(\frac{s_j - 1}{s_j}\right). \tag{17}$$

Although the parameters are defined for only the example, this result can generalize to an item measuring any number of attributes by changing λ_{j12} used in defining s_j to the weight that is only associated with mastery of all required attributes. It is important to notice that as in the DINA, the LCDM only requires the estimation of two parameters when these constraints are imposed.

3.3. *DINO*

As was mentioned previously, the DINO is similar to the DINA in that it only uses two parameters (s_j and g_j) to model the probability of a correct response. However, conceptually it is different in that only one of an item's required attributes must be mastered to have a high probability of a correct response. Mastery of any additional attributes will not change the probability of a correct response beyond the probability of a correct response when only a single attribute has been mastered. In terms of the LCDM, the DINO can be expressed as a model such that the conditional relationship is only positive when all other required attributes have not be mastered. Given that any other required attributes have been mastered, this conditional relationship between an attribute and the item is reduced to zero (i.e., the probability of a correct response does not change as a function of that attribute).

Specifically, the LCDM for the example item requiring only two attributes would be expressed as:

$$P\big(X = 1|\boldsymbol{\alpha}_i\big) = \frac{e^{\lambda_j \alpha_1 + \lambda_j \alpha_2 + (-\lambda_j)\alpha_1 \alpha_2 - \eta_j}}{1 + e^{\lambda_j \alpha_1 + \lambda_j \alpha_2 + (-\lambda_j)\alpha_1 \alpha_2 - \eta_j}}. \tag{18}$$

Here, λ_j is used to indicate a single value that is estimated for each item (in addition to η_j). Therefore, the odds of a correct response when comparing masters to nonmasters, given that all other attributes have not been mastered, increase by a factor of e^{λ_j}. Whereas the factor that the lod-odds increases is reduced to $e^{\lambda_j - \lambda_j} = e^0 = 1$ given that the second required attribute has been mastered ($\alpha_2 = 1$).

In the simple case with two attributes, it can be shown that

$$\eta_j = -\ln\left(\frac{g_j}{1 - g_j}\right) \tag{19}$$

and

$$\lambda_j = \eta_j + \ln\left(\frac{s_j - 1}{s_j}\right) \tag{20}$$

when using the original parameterization of the DINO.

A similar strategy can be applied to items requiring more than two attributes. Where all item parameters are defined such that the conditional distribution given mastery of any additional required attributes equal zero. In general, the sign in front of the weight, λ_j is determined as:

$$(-1)^{(c-1)} = \text{sign of the } \lambda_j \text{ weight}, \tag{21}$$

where c indicates the number of α_{ik} that are in the conditional relationship. For example, to determine the sign for the relationship of four attributes and the item (e.g., λ_{j1234}), $c = 4$ and, therefore,

$$(-1)^{(4-1)} = -1. \tag{22}$$

3.4. *Reduced RUM*

In many ways, the most complicated of the models to be derived from the LCDM is the reduced RUM. One cause of the complication between relating the reduced RUM to the LCDM is that the reduced RUM functions by penalizing an individual for not mastering any required attributes. Specifically, the probability of a correct response is defined as π_j^* unless someone has not mastered a set of the required attributes, in which this probability is reduced by each r_{jk}^* of the attributes not mastered. In contrast, the LCDM increases the odds, and thus the probability of a correct response for each attribute that has been mastered is increased. Therefore, before showing the relationship between the probability space defined by the reduced RUM and the LCDM, the "inverse" RUM is defined. Here, we use "inverse" to imply that mastery increases the probability of a correct response, a terminology that will be useful in defining the relationship between the LCDM and the RUM later.

The inverse RUM is mathematically equivalent to the reduced RUM and, therefore, the only difference is the definition of each item parameter and its respective space. The item response function for the inverse RUM is defined as:

$$P\left(X_{ij}=1|\boldsymbol{\alpha}_i\right)=\pi_j^{*\prime}\prod_{k=1}^{K}\frac{1}{{r_{jk}^*}^{q_{jk}\alpha_{ik}}}, \tag{23}$$

where $\pi^{*\prime}$ is equal to the probability of a correct response given that no required attributes have been mastered. Using the reduced RUM,

$$\pi_j^{*\prime}=\pi_j^*\prod_{k=1}^{K}{r_{jk}^*}^{q_{jk}}. \tag{24}$$

The remaining parameters are defined as the same r_{jk}^* parameters that were previously defined by the reduced RUM, where $0<r_{jk}^*<1$. However, because the inverse is used, for every attribute that is required and has been mastered by the ith examinee for the jth item (i.e., $q_{jk}\alpha_{ik}=1$), the probability of a correct response increases.

In doing this, the result of mastering an attribute is comparable to the functional form of the LCDM and so the LCDM can now be defined in terms of the parameters of the reduced RUM. We begin by defining

$$\eta_j=-\ln\left(\frac{\pi_j^{*\prime}}{1-\pi_j^{*\prime}}\right), \tag{25}$$

which is the intercept and is therefore used to compute the probability of guessing the correct answer when all required attributes *have not* been mastered. Next, the conditional relationships (λ_{ju}) are defined as

$$\lambda_{ju}=\ln\left(\frac{\pi_j^{*\prime}/r_{jk}^*}{e^{-\eta}+e^{-\eta}\pi_j^{*\prime}/r_{jk}^*}\right). \tag{26}$$

Recall, that if only the λ_{ju} values are nonzero, the LCDM is equivalent to the compensatory RUM. However, given that the true model is the reduced RUM, the remaining parameters of the LCDM have a specific functional form defined only by the λ_{ju} values. Specifically, the effect of mastery or nonmastery of the additional required attributes on the conditional relationship of any attribute and the item are defined based on the conjunctive nature of the attributes. Again, using the example item defined in (14) (without loss of generality), λ_{j12} is defined as

$$\lambda_{j12}=\ln\left(\frac{1+e^{-\eta}}{1+e^{\lambda_{j1}-\eta_j}+e^{\lambda_{j2}-\eta_j}-e^{\lambda_{j1}+\lambda_{j2}-\eta_j}}\right). \tag{27}$$

In defining the λ_{j12} based only on λ_{j1} and λ_{j2}, no additional parameters are required when compared to the total number of estimated parameters of the reduced RUM. In addition, the true nature of what is meant by a conjunctive model is defined. Specifically, based on (27), it can be shown that all $\lambda_{j12}>0$, which defines a conjunctive model in terms of the LCDM.

4. A Family of Cognitive Diagnosis Models

The LCDM is a general model that also includes many of the models that have been previously defined. However, even more importantly, the LCDM provides a parameterization that allows for a description of the differences between each model, while also providing for more complex data structures. In fact, through using the constraints described above, a natural ordering of the models can be provided ranging from disjunctive models to conjunctive models. To illustrate the ordering of the models previously discussed based on the their parameterizations using the LCDM, a two attribute item is used, but this ordering will be consistent for items requiring more than two attributes.

Because our goal is to display the differences between each model using the LCDM parameterizations, a simple case is used where each attribute has an equal degree of discrimination[2] (i.e., $\lambda_j 1 = \lambda_j 2$). In addition, to aid in the interpretation, the minimum probability of a correct response (i.e., relating to an examinee who has not mastered either of the required attributes) is fixed to 0.20 and the maximum probability for this item (for an examinee mastering both of the required attributes) is fixed to 0.90. Figure 1 plots the item parameters of the LCDM when fitting the DINO, compensatory RUM, reduced RUM, and DINA, respectively. Each plot within this figure is a bar graph (to simplify visual inspection of the parameters) indicating the values of $\lambda_{j1}, \lambda_{j2}$, and λ_{j12} for each model. Note that η_j is not included because for all models $\eta_j = 1.39$, which defines the minimum probability of a correct response for the illustrative item j at 0.20. In addition, because the maximum probability is 0.90, $\lambda_{j1} + \lambda_{j2} + \lambda_{j12} = 3.58$, in all models.

The first of the plots provides the parameters for the DINO model, which is the disjunctive model that was discussed earlier. In this example, $\lambda_{j1} = \lambda_{j2} = 3.58$ and $\lambda_{j12} = -3.58$, indicating the disjunctive nature of this model. In this "or" model, mastery of either attribute or both attributes results in the same sum of 3.58. The second plot, labeled compensatory RUM, shows the parameter space assuming a compensatory RUM were fit to the illustrative item. Notice that λ_{j1} and λ_{j2} have reduced to 1.79, and $\lambda_{j12} = 0$. Finally, the bottom two plots show the parameter space for two conjunctive models, the reduced RUM and DINA, respectively. The parameters for the reduced RUM show that λ_{j1} and λ_{j2} are even smaller and λ_{j12} is now positive, whereas for the DINA, $\lambda_{j1,}$ and λ_{j2} equal zero and λ_{j12} is positive and equal to 3.58, which was the initial value of λ_{j1} and λ_{j2} for the DINO.

The LCDM defines a continuum of models where disjunctive models will always have a negative λ_{j12} (characterizing the change on the conditional relationship given that an attribute is mastered) and large λ_{j1} and λ_{j2}. As the models move through to conjunctive models, the λ_{j12} increases and λ_{j1} and λ_{j2} decrease. By defining such a family of models, it is now conceivable to modify the original assumptions made by each model. For example, imagine a model that has a large λ_{j12}, much like the DINA, but λ_{j1} and λ_{j2} are slightly nonzero. Such a model would allow an examinee to have a slightly better chance at guessing the correct answer when at least one of the required attributes has been mastered, while still modeling the importance of knowing all required attributes. An alternative model could function much like the DINO, where only one of the specified attributes is required to have a high probability of correctly responding to the item. However, if the magnitude of λ_{j12} were slightly reduced, then there could be some benefit, although minor to mastering additional attributes.

Based on these results, any possible set of constraints can be used to define a model that fits the cognitive theory of item responses. In addition, a better understanding of the relationships between any model (including new models) can be described in a general parametric form.

[2]Here, we use discrimination to indicate that the increase in the probability of a correct response is the same when comparing an examinee knowing only attribute one to an examinee knowing only attribute two.

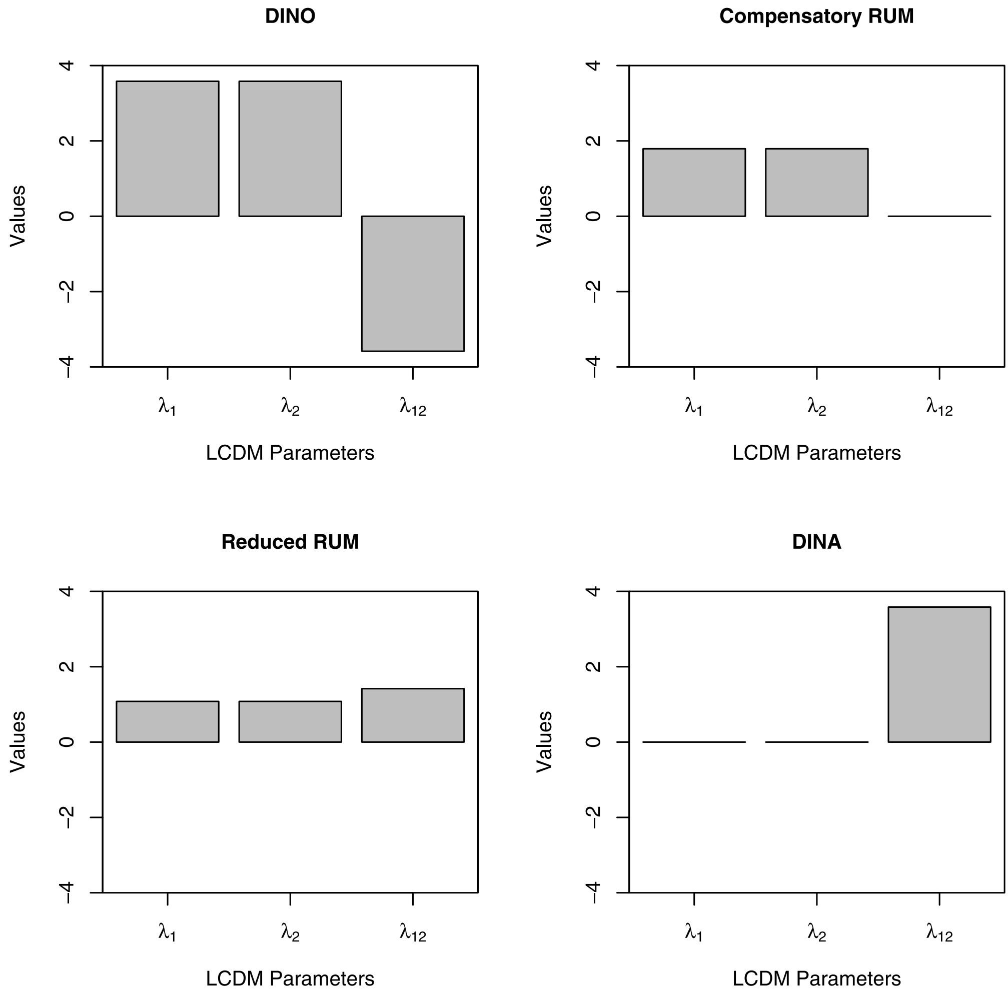

FIGURE 1.
Four graphs to provide graphical representation of model differences in terms the LCDM parameterization.

5. Estimation

The primary intent of this paper was to introduce the LCDM and provide its relationship to other models with comparable assumptions. However, defining a model is not enough, because its estimation and application must be feasible. In this paper, LCDM estimation is accomplished using a Markov chain Monte Carlo algorithm (MCMC). MCMC estimation is has been growing in popularity and is a common choice in the literature to estimate cognitive diagnosis models (e.g., Hartz, 2002; Templin, Henson, Templin, & Roussos, in press; Templin & Henson, 2006). However, it should be noted that alternative estimation algorithms that allow estimation of constrained latent class analysis could be used such as M-Plus (e.g., M-plus; Muthén & Muthén, 1998–2006), as is discussed by Templin, Henson, and Douglas (2008), or Latent Gold (Vermunt & Magidson, 2005).

In the MCMC estimation algorithm, all priors for item parameters ($\boldsymbol{\lambda}_j$ and η_j for all $j = 1, \ldots, J$) are set to have uniform distributions $U(-10, 10)$. However, uniform priors are not assumed for examinee attributes because of the possibility that attributes are related. That is, an examinee that has mastered one attribute will tend to have also mastered additional attributes. Hartz (2002) was one of the first to describe an algorithm using an empirical Bayes structure for CDMs as a way to estimate attribute associations. In Hartz (2002), tetrachoric correlations were used to describe the relationship between any attribute-pair association, and thus an underlying multivariate distribution is implemented (also see DiBello, Stout & Roussos, 2007). Although this method has been shown to perform well, as the number of attributes increases, the number of estimated item parameters increases exponentially. As an alternative, de la Torre and Douglas (2004) (and also Templin, 2004) describe an alternative method to reduce the number of estimated parameters. Specifically, the assumption is made that the attribute associations can be modeled using a nonlinear factor model. Therefore, the prior distribution of attribute mastery vectors is assumed to have a dichotomized multivariate normal distribution with "cutoffs" equal to κ_k for $k = 1, \ldots, K$ and factor weights, γ_k, for $k = 1, \ldots, K$. Because attributes are typically based on a common content (e.g., math) such an assumption may not be too restrictive. In addition, this method has been shown to be robust to minor violations (Templin et al., in press). Because both κ_k and γ_k are also estimated the method is described as "empirical Bayes" (the prior of examinee attribute patterns is estimated from the sample). In addition, to ensure a unique solution, the constraints previously specified are incorporated. Specifically, "monotonicity" and specification of the Q-matrix is assumed and attribute mastery and nonmastery is defined as a 0/1 variable, as in all discussed CDMs.

6. Mixed Number Subtraction Example

To provide an example of the benefits of the LCDM, we use an example that was originally presented by Templin, Henson, and Douglas (2006). They use a subset of items developed to measure mixed number fraction subtraction from a data set using 2,144 examinees that was originally developed and collected by Tatsuoka (1990). The value of using such a data set is that it has been established as a realistic example for the implementation of cognitive diagnosis (Tatsuoka, 1990; de la Torre & Douglas, 2004; Templin et al., 2006). In particular, Templin et al. (2006) use 13 of the items that are determined to measure a total of four attributes. The four attributes (or skills) are defined as (1) convert a whole number to a fraction, (2) separate a whole number from a fraction, (3) simplify before subtracting, and (4) find a common denominator. Although these attributes are used as an initial analysis to recover their original parameters, a new set of three attributes were defined later (Templin et al., 2008) which will be used for our example. The new attributes were determined based on the fact that a different set of models were being used. Notice that this again emphasizes the importance of the Q-matrix. In many ways, the usefulness of the Q-matrix is defined by the theory underlying the model and its application. Therefore, effort must always go into establishing a cognitive theory as to what skills are necessary and the functional form that these skills define the probability of a correct response. The skills defined for this example are (1) borrowing from a whole number, (2) separating a whole number from a fraction and (3) determining a common denominator. Table 1 provides the 13 items and the their required skills. In addition, we note that the item number 7 originally discussed by Templin et al. (2006) was $3\frac{4}{5} - 3\frac{2}{5} =$ was removed because none of the three newly defined skills were required.

6.1. Method

In their paper, Templin et al. use the DINA model and provide estimates of the s_j and the g_j parameters for each item determining a reasonable fit. One advantage of using this data set as an example is that the algorithm developed to estimate the LCDM can be used with the previously defined constraints for the DINA and the estimated parameters can be compared to

TABLE 1.
Mixed number fraction subtraction items.

Item number	Item	Q-matrix		
1	$3\frac{1}{2}-2\frac{3}{2}$	1	1	0
2	$3-2\frac{1}{5}$	1	0	1
3	$3\frac{7}{8}-2$	1	0	1
4	$4\frac{4}{12}-2\frac{7}{12}$	1	0	0
5	$4\frac{1}{3}-2\frac{4}{3}$	1	1	0
6	$\frac{11}{8}-\frac{1}{8}$	1	1	0
8	$2-\frac{1}{3}$	1	0	1
9	$4\frac{5}{7}-1\frac{7}{4}$	1	1	1
10	$7\frac{3}{5}-\frac{4}{5}$	1	0	0
11	$4\frac{1}{10}-2\frac{8}{10}$	1	0	0
12	$7-1\frac{4}{3}$	1	1	1
13	$4\frac{1}{3}-1\frac{5}{3}$	1	1	0

the parameters estimated by Templin et al. (2006) as a preliminary evaluation of the estimation software. Therefore, initially the agreement of LCDM estimation is reported after transforming the LCDM parameters to the DINA parameterizations based on the relationships defined in (16) and (17). The mean absolute deviation (MAD) is used as measure of the consistency of the LCDM estimated s_j and g_j when compared with the original parameter estimates.

As a second analysis, the same data set is used for an LCDM calibration where all parameters are estimated using the *newly defined* Q-matrix (as opposed to using constraints consistent with the DINA model). The chains are visually explored for convergence and the item parameter estimates are obtained. Here, we note that in its current form, this analysis was used as an exploratory analysis to suggest possible alternative models for particular items as opposed to the DINA model. However, prior to determining the reasonableness of an alternative model, a basic measure of model fit was used based on a posterior predictive check. Specifically, the observed item associations were compared to what is predicted by the model (for a general description, see Gelman, Carlin, Stern, & Rubin, 1995; Gelman, Meng, & Stern, 1996; Lynch & Western, 2004). The statistic used for this comparison is the mean absolute deviation (or MAD). In general, as shown in (28), the MAD can be used to compare the observed item associations (ρ_{ij}) to the model predicted item associations ($\hat{\rho}_{ij}$) across all possible pairs of items where $i \neq j$. However, when a Bayesian model is used, all item parameters have an estimated posterior distribution and, therefore, this discrepancy measure (the MAD) also has a distribution. The distribution of the MAD comparing the observed item correlations to the "predicted" item correlations is computed as the measure of discrepancies (Templin & Henson, 2006). The mean and credibility interval is reported as measure to indicate "fit" (i.e., the ability of the estimated model to recover characteristics of the observed data)

$$\mathrm{MAD} = \frac{\sum_{j=1}^{J} \sum_{i \neq j} |\rho_{ij} - \hat{\rho}_{ij}|}{J(J-1)}. \tag{28}$$

After evaluating the fit of the full model, the LCDM is used to explore alternatives to the full model, which is an advantage of defining a family of CDMs based on the LCDM. Therefore, after fitting the full model, the item parameters are explored for a simplification. Having determined possible constraints to the full model, a new reduced model is estimated and the relative fit of each model is compared using the AIC (Akaike, 1974) and BIC (Schwarz, 1978).

6.2. *Results*

As a preliminary analysis, the data is fit using the LCDM with DINA constraints and, therefore, all $\lambda_j = 0$ with the exception of the weight associated with all attributes required by the item. A chain of length 10,000 is used with a burnin of 5,000. The chains were visually inspected and it was determined that convergence had been reached. Given convergence, the posterior means were used as estimates of LCDM parameters and these values were transformed to DINA parameters. In comparing the estimated parameters with the values reported by Templin et al. (2006), the MAD of the slip parameters was 0.03 and the MAD for the guess parameters was 0.01, providing evidence that the algorithm can be effective in estimating the DINA.

As a second analysis, the LCDM was estimated without any additional constraints using the new Q-matrix. All chains were visually inspected for convergence and it was determined that a chain length of 10,000 with a burnin of 5,000 was sufficient. The posterior means are used to serve as parameter estimates and standard deviation of the posteriors are used to estimate parameters standard errors.

The posterior distribution of the MAD comparing the observed item associations to the model predicted item associations had a mean of 0.064 where the 95% credibility limit is 0.052 to 0.077 suggesting an acceptable recovery of the item associations and model fit. Therefore, the estimated parameters for each of the items were visually inspected for patterns as a method to determine item specific models, and thus suggestions as to the underlying cognitive theory. Note that such a distinction between multiple models is only relevant for those items that require more than a single attribute. Recall that the distinctive feature that differentiates each of the models is based on the extent to which the conditional distribution of any attribute and the item response depends on mastery or nonmastery of other attributes. In the event that an item only measures a single attribute, all models reduce to the same model and for this reason items 4, 6, 10, and 11 are not included in our discussion. The remaining item parameters for each item are summarized in Fig. 2.

Figure 2 contains a set of graphs (one for each of the included items) similar to the figures presented previously in that the X-axis represents the estimated parameters other than the intercept values (these are not used to differentiate between models) for each of the eight items. The Y-axis represents the estimated value of each of the parameters. In addition, because these values are estimated, a slightly different presentation is used. Specifically, instead of a bar graph, the estimate (as indicated as a point) and its 95% credibility limits (computed as the 2.5th percentile and the 97.5th percentile of each posterior distribution) are provided. In general, these bounds can provide an approximate guide as to the significance of each parameter although we acknowledge that there are certain parameters for which the true distribution is skewed. That is, monotonicity constrains the association of any attribute with any given item to be positive and, therefore, the corresponding weights (e.g., λ_1 and λ_2 for Item 1) must be positive.

Because these graphs are similar to Fig. 1, it is possible to identify items that may be associated with specific models. For example, recall that the LCDM reduces to the DINA model when there is no conditional association between any given attribute and the item unless all other required attributes for that item have been mastered. Using the full parameterization, it can seen that Item 3 is consistent with this pattern. Item 3 ($3\frac{7}{8} - 2$) was defined as requiring the ability of borrowing and finding a common denominator. In this case, it could be argued that lacking either of these skills would dramatically impede an examinee's chances of correctly responding to this item. However, the DINA model is not appropriate for many of these items. For example, Item 8 ($2 - \frac{1}{3}$) would appear to be closely related to the assumptions made by the DINO model. That is, an examinee is expected to perform well if they know how to borrow from a whole number or find a common denominator, which suggests two possible strategies to obtaining the correct answer. In the first strategy, examinees have mastered the attribute of borrowing, and for that reason, they recognize that $2 = 1\frac{3}{3}$ which allows the solution of this item to be obtained. However, if

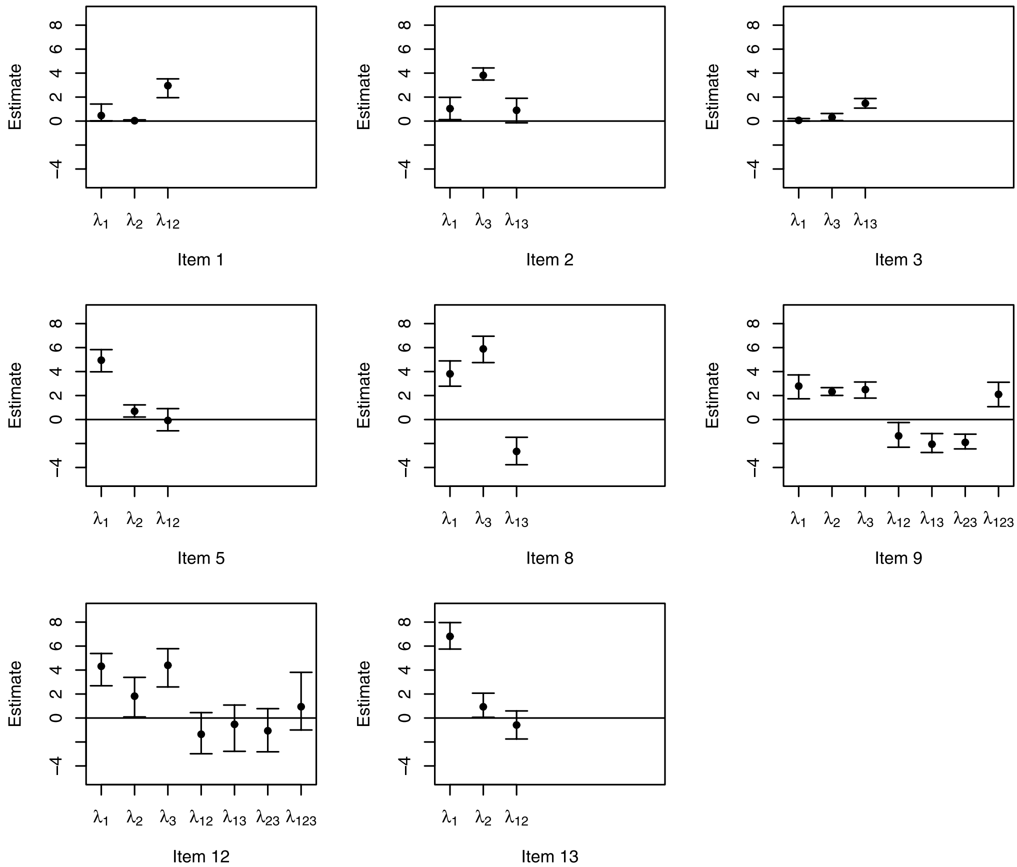

FIGURE 2.
Parameter estimates for the mixed fraction subtraction items requiring more than one attribute.

TABLE 2.
AIC and BIC to compare a full and reduced model.

Model	# of parameters	AIC	BIC
Full	48	22,541	22,813
Reduced	35	22,546	22,745

examinees have mastered the skill of determining a common denominator, they could set $2 = \frac{6}{3}$ as a way to obtain the correct answer.

In summary, the current model, which is the estimated LCDM with a total 48 item parameters, appears as though it may be simplified. Specifically, based on the estimated model parameters, it would appear that Items 1 and 3 may be consistent with a DINA, Item 8 may be consistent with a DINO model and Items 2, 5, and 12 may be consistent with a compensatory RUM. Should these constraints be placed on the LCDM, a total 35 parameters (a reduction of 13 item parameters) would be estimated. Thus, the new constrained model is estimated. In this particular case, the AIC and BIC are used to compare the two models and aid in model selection. Table 2 contains a summary of the number of estimated item parameters, AIC, and BIC for the full model (the first row) and for the reduced Model (the second row). As can be seen from the table, the AIC is very similar between the two different models, although the full model has a slightly smaller value. Thus, the AIC provides weak evidence that the full model should be used. However, the BIC, which has a stronger penalty for additional parameters, is smaller for the reduced model suggesting that the reduced model should be used. Therefore, the AIC and BIC seem to suggest that the proposed reduction of the model is feasible. This example illustrates how the LCDM can be used as a method to identify reasonable models for each of the items.

In addition to model selection, this example can also be used to show how the LCDM can provide some evidence of a misspecified Q-matrix or a misspecified theory of cognition for items that need additional attention. A good example of such a case is Item 12. Notice that for Item 12 ($7 - 1\frac{4}{3}$) it was assumed that all three attributes are required, however, all weights associated with attribute two are essentially zero. Upon further review of this item, it appears that examinees may never use the "separating" attribute to define $1\frac{4}{3} = 2\frac{1}{3}$, but instead simply define $7 = 5\frac{6}{3}$ prior to completing this problem. In using this strategy, they would only need to know how to borrow and find a common denominator. In addition, items such as Item 9 (suggesting a DINO) may be difficult to interpret, which can be used to direct a researchers attention to those problems where cognition may not be well understood.

Therefore, using this small example, it can be seen that the LCDM can be a useful tool to provide possible item specific models which can give some insight as to the cognitive process that is being used (assuming that reasonable attributes have originally been identified). We should note that in application, further modification of the Q-matrix and attributes should be completed. However, the subset of items for mixed number subtraction does provide clear examples for which the LCDM can provide value over cognitive diagnosis models previously defined in the literature.

7. Discussion

In this paper, we have provided a discussion about the LCDM. In doing so, we have explicitly defined the relationship between a number of models that have currently been introduced allowing for natural comparisons between these models to be made. In addition, the LCDM provides a model that can describe the functional relationship between attribute mastery and the item response probability without having to specify a restrictive type of model such as conjunctive versus disjunctive. Therefore, the LCDM has the added benefit of estimating attribute mastery, while providing empirical information regarding a reasonable "model."

Also, because of its flexibility, the LCDM can be used to describe a family of models, and thus gives a clearer explanation of conjunctive models, compensatory models, and disjunctive models. In addition, the presented analysis provided a preliminary example of the value of the LCDM as opposed to using one of the specific models previously presented in the literature. In this particular case, it could be seen that the DINA model (which was originally fit) does not seem appropriate for all items. In addition, the LCDM provided some insight as to what model could be more reasonable for some items while allowing for limited inspection as to Q-matrix misspecification. However, the example also shows some of the limitations. As a result of the small number of items and the total number of parameters for each item, it can be seen that some of the estimates have large standard errors. By adding appropriate constraints, as were suggested by the initial parameters' estimates, estimation precision possibly could be improved.

Future research will focus on the estimation limitations of the LCDM. Specifically, as the number of attributes included in the Q-matrix increases and as its complexity increases, the number of parameters estimated by this model will also increase. In these cases, expectations of its performance in estimated attribute mastery and item parameters must be explored. In addition, model comparisons using common indices such as the AIC, BIC, and measures of absolute fit such as posterior predictive checks must continue to be explored, which could result to clear guidelines for model identification. Finally, possible expansions of this model such as the addition of a continuous ability measure to imply an incomplete Q-matrix (much like what is used in the full RUM) will be explored.

Acknowledgements

The work completed in this paper is partially funded by NSF Grant #SES-0648861 and by NSF MMS Grant #SES-0750859 for Drs. Henson and Templin, respectively, who contributed equally to this paper. In addition, we would like to thank Jeffrey Douglas, Brian Junker, Roger Millsap, and three anonymous reviewers for their comments and suggestions.

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

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

DiBello, L., Stout, W., & Roussos, L. (2007). Cognitive diagnosis Part I. In C.R. Rao & S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics*. Amsterdam: Elsevier.

Embretson, S. (1985). Studying intelligence with test theory models. In *Current topics in human intelligence* (Vol. 1, pp. 98–140).

Fischer, G., & Forman, A. (1982). Some Applications of Logistic Latent Trait Models with Linear Constraints on the Parameters. *Applied Psychological Measurement*, *6*(4), 397–416.

Fu, J. (2005). *A polytomous extension of the fusion model and its Bayesian parameter estimation*. Unpublished doctoral dissertation.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (1995). *Bayesian data analysis*. London: Chapman & Hall.

Gelman, A., Meng, X., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *6*, 733–807.

Haberman, S.J. (1974). Loglinear models for frequency tables derived by indirect observation: Maximum likelihood equations. *Annals of Statistics*, *2*, 911–924.

Haberman, S.J. (1979). *Qualitative data analysis*: Vol. *2*. *New developments*. New York: Academic 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.

Hagenaars, J. (1993). *Loglinear models with latent variables*. Thousand Oaks: Sage.

Hartz, S. (2002) *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality*. Unpublished doctoral dissertation.

Jang, E. (2005) *A validity narrative: Effects of reading skills diagnosis on teaching and learning in the context of NG TOEFL*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Janssen, R., & DeBoeck, P. (1997). Psychometric modeling of componentially designed synonym tasks. *Applied Psychological Measurement*, *21*, 37–50.

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.

Lynch, S., & Western, B. (2004). Bayesian posterior predictive checks for complex models. *Sociological Methods and Research*, *32*, 302–335.

Macready, G.B., & Dayton, C.M. (1977). The use of probabilistic models in the assessment of mastery. *Journal of Education Statistics*, *33*, 379–416.

Muthén, L.K., & Muthén, B.O. (1998–2006). *Mplus user's guide* (4th ed.). Los Angeles: Muthén & Muthén.

Rindskopf, D. (1983). Ageneral framework for using latent class analysis to test hierarchical and non hierarchical learning models. *Psychometrika*, *48*, 85–97.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Tatsuoka, K.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*. Hillsdale: Erlbaum.

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). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, *11*(3), 287–305.

Templin, J., Henson, R., & Douglas, J. (2006). *General theory and estimation of cognitive diagnosis models: Using Mplus to derive model estimates*. 2007 National Council on Measurement in Education training session, Chicago, Illinois.

Templin, J., Henson, R., & Douglas, J. (2008). *General theory and estimation of cognitive diagnosis models as constrained latent class models*. Manuscript under review.

Templin, J., Henson, R., Templin, S., & Roussos, L. (in press). Robustness of hierarchical modeling of attribute correlation in cognitive diagnosis models. *Applied Psychological Measurement*.

Vermunt, J.K., & Magidson, J. (2005). *Technical guide for latent GOLD choice 4.0: Basic and advanced*. Belmont: Statistical Innovations.

von Davier, M. (2005). *A general diagnostic model applied to language testing data* (ETS Research Report RR-05-16).

Manuscript Received: 22 APR 2007
Final Version Received: 9 SEP 2008
Published Online Date: 6 NOV 2008