Listening. Learning. Leading.®

Research Report

ETS RR–14-40

The Log-Linear Cognitive Diagnostic Model (LCDM) as a Special Case of the General Diagnostic Model (GDM)

Matthias von Davier

December 2014

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

ETS Research Report Series ISSN 2330-8516

RESEARCH REPORT

The Log-Linear Cognitive Diagnostic Model (LCDM) as a Special Case of the General Diagnostic Model (GDM)

Matthias von Davier

Educational Testing Service, Princeton, NJ

Diagnostic models combine multiple binary latent variables in an attempt to produce a latent structure that provides more information about test takers' performance than do unidimensional latent variable models. Recent developments in diagnostic modeling emphasize the possibility that multiple skills may interact in a conjunctive way within the item function, while individual skills still may retain separable additive effects. This extension of either the conjunctive deterministic-input-noisy-and (DINA) model to the generalized version (G-DINA) or the compensatory/additive general diagnostic model (GDM) to the log-linear cognitive diagnostic model (LCDM) is aimed at integrating models with conjunctive skills and those that assume compensatory functioning of multiple skill variables. More recently, a result was proven mathematically that the fully conjunctive DINA model, which combines all required skills in a single binary function, may be recast as a compensatory special case of the GDM. This can be accomplished in more than one form such that the resulting transformed skill-space definitions and design (Q) matrices are different from each other but mathematically equivalent to the DINA model, producing identical model-based response probabilities. In this report, I extend this equivalency result to the LCDM and show that a mathematically equivalent, constrained GDM can be defined that yields identical parameter estimates based on a transformed set of compensatory skills.

Keywords Diagnostic classification models; general diagnostic model; log-linear cognitive diagnosis model; model equivalency; parameter redundancy; identifiability

doi:10.1002/ets2.12043

Diagnostic models combine multiple binary latent variables in an attempt to produce a latent structure that provides more information about test takers' performance than do unidimensional latent variable models. Recent developments in diagnostic modeling emphasized the possibility that multiple skills may interact in a conjunctive way within the item function, while individual skills still may retain separable additive effects. This extension of either the conjunctive deterministic-input-noisy-and (DINA; Junker & Sijtsma, 2001; Macready & Dayton, 1977) model to the generalized version (G-DINA; de la Torre, 2011) or the compensatory/additive general diagnostic model (GDM; von Davier, 2005, 2008; von Davier & Yamamoto, 2004) to the log-linear cognitive diagnostic model (LCDM; Henson, Templin, & Willse, 2009) is aimed at integrating models with conjunctive skills and those that assume compensatory functioning of multiple skill variables.

More recently, a result was proven mathematically that the conjunctive DINA model, which combines all required skills into a single binary function, may be recast as a compensatory special case of the GDM (von Davier, 2011, 2014). This can be accomplished in more than one form such that the resulting transformed skill-space definitions and design (Q) matrices are different from each other but mathematically equivalent to the DINA model, producing identical model-based response probabilities. DeCarlo (2011) showed in a related result that the DINA model can be transformed in a way such that it can be estimated with general-purpose latent class software such as LEM (Vermunt, 1997). These results emphasize the need to reduce these models to their mathematical definitions and to try to understand them on this formal basis rather than based on verbal descriptions of how compensatory or conjunctive skills may or may not better represent the underlying cognitive functions required to solve a task. Moreover, if multiple mathematically equivalent, but structurally different, sets of skill definitions that produce the same conditional response probabilities with different underlying assumptions exist, conjectures about whether skills function in a compensatory or conjunctive manner become questionable at best.

This report extends the equivalency result that shows the DINA model to be a special case of the GDM (von Davier, 2011, 2014) to the LCDM and shows that a mathematically equivalent, constrained GDM that yields identical model-based

Corresponding author: M. von Davier, E-mail: mvondavier@ets.org

response probabilities, as well as equivalent parameter estimates based on an alternative compensatory latent skill structure, can be defined.

The remainder of the report is structured as follows. The next section introduces the GDM and shows why the LCDM may be conceptualized (e.g., Rupp, Templin, & Henson, 2010) as a generalization of the GDM. It is followed by a section that recasts the LCDM as a special case of a constrained GDM with a modified skill space and a constrained skill distribution, extending the results presented in von Davier (2011, 2014). An empirical example is then presented. The report closes with a discussion of the main result.

The General Structure, and Examples, of Diagnostic Models

Diagnostic models are latent structure models (von Davier, 2005, 2008, 2009), that is, models that relate observed random variables to unobserved random variables that are considered probabilistic causes (Suppes & Zanotti, 1981) of the observed variables. More specifically, assume that there are $i = 1, \ldots, I$ multinomial random variables X_i with values in $\Omega_i = \{0, 1, \ldots, m_i\}$ observed on $n = 1, \ldots, N$ respondents. That is, let $x_{ni} \in \{0, 1, \ldots, m_i\}$ denote the response of test taker n on variable (item or task) i. Then, denote

$$\boldsymbol{x_n} = \left(x_{n1}, \ldots, x_{nI}\right) \in \otimes_{i=1}^{I} \Omega_i = \Omega_O, \tag{1}$$

the vector of observed responses for respondent n.

In addition to these observed variables, diagnostic models assume $d = 1, \ldots, D$ unobserved variables A_d, with values in $\Omega_d^A = \left\{l_{1d}, \ldots, l_{s_d d}\right\}$, called attributes or skills in this context. That is, let

$$\boldsymbol{a} = \left(a_1, \ldots, a_D\right) \in \prod_{d=1}^{D} \Omega_d^A = \Omega_L, \tag{2}$$

where $\boldsymbol{a}_n$ denotes the attribute vector of respondent n. There are $\prod_{d=1}^{D} s_d$ distinct elements in Ω_L. Assume that there is a discrete probability measure π with

$$\pi\left(a_1, \ldots, a_D\right) \geq 0 \text{ and } \sum_{\Omega_L} \pi\left(a_1, \ldots, a_D\right) = 1. \tag{3}$$

In many cases, only binary skills are assumed and binary observed response variables are analyzed. I therefore assume that all skills/attributes are binary, that is, $\Omega_d^A = \{0, 1\}$ for all $d = 1, \ldots, D$. Note, however, that the GDM is also well defined for ordinal and interval-level skill variables as well as for polytomous response variables (von Davier, 2008).

The joint space of observed and unobserved variables is given by $\Omega = \Omega_O \times \Omega_L$, and the complete data for each respondent can be written as $(\boldsymbol{x_n}, \boldsymbol{a_n})$.

Like all latent variable models, diagnostic models can be viewed as incomplete data problems as only the $\boldsymbol{x_n}$ part is observed and a model is needed to make inferences about the unobserved $\boldsymbol{a}_n$.

An important point to make is that the latent variable space $\boldsymbol{a}_n \in \otimes_{d=1}^{D} \Omega_d^A = \Omega_L$ is nothing but an assumption or a hypothesis, based on either a substantive theory or (unfortunately, and not infrequently) some ad hoc choice of the analyst. It is merely one of many potential conceptualizations of additional (but unobservable) variables that are utilized to “explain” the associations seen in the observed variables.

More specifically, most (if not all) latent variable models assume that the associations among the $X_1, \ldots, X_I$ vanish once we condition on the $A_1, \ldots, A_D$. More formally, it is assumed that the conditional probability of the observed variables given latent variables exists and that

$$P\left(x_1, \ldots, x_I | a_1, \ldots, a_D\right) = \prod_{i=1}^{I} P_i\left(X = x_i | a_1, \ldots, a_D\right). \tag{4}$$

The assumption in Equation 4 is commonly referred to as local (stochastic) independence. What remains to be defined is the functional form of the probability functions in the product on the right-hand side of Equation 4. More specifically, an item response function

$$P_i\left(X = x_i | a_1, \ldots, a_D\right) = g\left(x_i, \lambda_i, \boldsymbol{q_i}, a_1, \ldots, a_D\right), \tag{5}$$

has to be specified to allow inferences about $a_1, \ldots, a_D$ given the observed variables $x_1, \ldots, x_I$. In Equation 5, the $\boldsymbol{q_i}$ denotes a D-dimensional binary vector that determines which of the $a_1, \ldots, a_D$ are utilized in the item function, and the λ_i is a U_i-dimensional vector of item parameters.

A very general approach was suggested by von Davier and Yamamoto (2004) and later implemented in the GDM (von Davier, 2005, 2008). More specifically, the item function in Equation 5 is written such that linear functions of skill components and their combinations can be used to predict the log-odds of response probabilities.

That is, for binary or ordinal response variables $x = 1, \ldots, m_i$, let

$$g\left(x, \lambda_i, \boldsymbol{q_i}, \boldsymbol{a}\right) = \frac{\exp\left(f_x\left(\lambda_i, \boldsymbol{h}\left(\boldsymbol{q_i}, \boldsymbol{a}\right)\right)\right)}{1 + \sum_{y=1}^{m_i} \exp\left(f_y\left(\lambda_i, \boldsymbol{h}\left(\boldsymbol{q_i}, \boldsymbol{a}\right)\right)\right)}, \tag{6}$$

(see, e.g., von Davier, 2005, 2008; von Davier & Yamamoto, 2004), and for $x \in \{0, 1\}$ binary only, a very similar setup has been suggested in Henson et al. (2009) and Rupp et al. (2010). Obviously, Equation 6 can be transformed into

$$\ln\left[\frac{g\left(x, \lambda_i, \boldsymbol{q_i}, \boldsymbol{a}\right)}{g\left(0, \lambda_i, \boldsymbol{q_i}, \boldsymbol{a}\right)}\right] = f_x\left(\lambda_i, \boldsymbol{h}\left(\boldsymbol{q_i}, \boldsymbol{a}\right)\right), \tag{7}$$

a customary approach frequently used in categorical data analysis (Agresti, 1990; Haberman, 1978). The advantage of Equation 7 over Equation 5 is that the left-hand side of Equation 7 is not a bounded quantity, so the image of $f_x(.)$ can cover the real numbers. In contrast, Equation 5 parameterizes a probability such that $0 < g(x, \lambda_i, \boldsymbol{q_i}, \boldsymbol{a}) < 1$ needs to be ensured.

The General Diagnostic Model (GDM)

The GDM, as given in the developments presented by von Davier and Yamamoto (2004), defines the right-hand side of Equation 7 as

$$f_x\left(\lambda_i, \boldsymbol{h}\left(\boldsymbol{q_i}, \boldsymbol{a}\right)\right) = \lambda_{ix}^T \boldsymbol{h}\left(\boldsymbol{q_i}, \boldsymbol{a}\right), \tag{8}$$

with $\boldsymbol{h}(\boldsymbol{q_i}, \boldsymbol{a}) = (1, q_{i1}a_1, \ldots, q_{iD}a_D)$ and $\lambda_{ix} = (\lambda_{i0x}, \lambda_{i1x}, \ldots, \lambda_{iDx})$. Often, we will assume that $\lambda_{idx} = x\lambda_{id}$ for all $d = 1, \ldots, D$. That is, we obtain the customary additive form of the GDM for dichotomous and polytomous item responses:

$$f_x\left(\lambda_i, \boldsymbol{h}\left(\boldsymbol{q_i}, \boldsymbol{a}\right)\right) = \lambda_{i0x} + \sum_{d=1}^{D} x\lambda_{id} q_{id} a_d. \tag{9}$$

The GDM was subsequently extended to mixture distribution GDM (von Davier, 2007), which allows different skill distributions in different latent subpopulations, as well as a multigroup GDM with constraints across groups (e.g., Xu & von Davier, 2008) and multilevel diagnostic models (von Davier, 2007, 2010).

The Log-Linear Cognitive Diagnostic Model (LCDM)

The GDM served as the basis for the derivation of the LCDM (Rupp et al., 2010). More specifically, Henson et al. (2009) suggested extending the function $\boldsymbol{h}(\boldsymbol{q_i}, \boldsymbol{a})$ used in Equations 6–9, which connects the skill pattern $\boldsymbol{a} = (a_1, \ldots, a_D)$ and the item-skill requirements $\boldsymbol{q_i} = (q_{i1}, \ldots, q_{iD})$ to account for potential skill interactions. More specifically, the LCDM, as Henson et al. suggest, can be understood as a simple extension of the binary special case of the GDM, as given in Equation 9. That is, the LCDM assumes

$$f\left(\lambda_i^*, \boldsymbol{h}^*\left(\boldsymbol{q_i}, \boldsymbol{a}\right)\right) = \lambda_{i0} + \sum_{d=1}^{D} \lambda_{id} q_{id} a_d + \sum_{d<e} \lambda_i^{d,e} q_{id} q_{ie} a_d a_e + \ldots \tag{10}$$

for binary skills $a_d \in \{0, 1\}$ and binary item response data. The dots also indicate that higher order terms $\lambda_i^{d,e,f} q_{id} q_{ie} q_{if} a_d a_e a_f$ of three or more skills also may be included in the item function. The parameters $\lambda_i^{d,e}$, $\lambda_i^{d,e,f}$, … associated with these interactions quantify the *conjunctive* effects of having two, three, or more of the required skills that are not already explained by the main effects λ_{id} of required ($q_{id} = 1$) skills.

The Deterministic-Input-Noisy-And (DINA) Model

The DINA model (Haertel, 1989; Junker & Sijtsma, 2001; Macready & Dayton, 1977) produces only two levels of probabilities for the correct response per item. These two probabilities are the guessing probability, g_i, and the complement of the slipping probability, $1 - s_i$, no matter how many skills are involved. For example, an item that requires four different skills has, potentially, 16 different probabilities of a correct response, one for each of the 16 skill combinations that are possible in a four-skill item, while the DINA constrains these 16 levels to only two different values.

The DINA model can be written as a set of two nested equations: a first expression

$$h\left(\boldsymbol{q}_i, \boldsymbol{a}\right) = \prod_{d=1}^{D} a_d^{q_{id}} \in \{0, 1\} = \xi_{\boldsymbol{a}}^{\boldsymbol{q}_i}, \tag{11}$$

that maps the skill pattern and the Q-matrix vector onto a single binary variable, and a second equation that combines the expression in Equation 11 with the slipping and guessing parameters into a simple item function

$$P_i\left(X = 1|\boldsymbol{a}\right) = \left(1 - s_i\right)^{\xi_{\boldsymbol{a}}^{\boldsymbol{q}_i}} g_i^{\left(1-\xi_{\boldsymbol{a}}^{\boldsymbol{q}_i}\right)}, \tag{12}$$

and, consequently,

$$P_i\left(X = 0|\boldsymbol{a}\right) = s_i^{\xi_{\boldsymbol{a}}^{\boldsymbol{q}_i}} \left(1 - g_i\right)^{\left(1-\xi_{\boldsymbol{a}}^{\boldsymbol{q}_i}\right)}. \tag{13}$$

The DINA model is referred to as a conjunctive model. Any combination of required skills is projected on a 0/1 variable; only if all required skills are present do we have $\xi_{\boldsymbol{a}}^{\boldsymbol{q}_i} = 1$. The DINA model as given in Equation 12 may be transformed into the logistic form as given in Equation 6. That is, one may write

$$\ln\left[\frac{P_i\left(X = 1|\boldsymbol{a}\right)}{P_i\left(X = 0|\boldsymbol{a}\right)}\right] = \lambda_{i0} + \xi_{\boldsymbol{a}}^{\boldsymbol{q}_i}\lambda_{i1}, \tag{14}$$

with

$$\ln\left[\frac{1 - s_i}{s_i}\right] - \ln\left[\frac{g_i}{1 - g_i}\right] = \lambda_{i1} \text{ and } \ln\left[\frac{g_i}{1 - g_i}\right] = \lambda_{i0}. \tag{15}$$

DeCarlo (2011) provided a reparameterization of the DINA model along the same lines. This transformed version of the DINA can be estimated with LEM (Vermunt, 1997).

While the conjunctive nature of the DINA model appears to be one of the defining characteristics of this model, the model representation is not unique. von Davier (2011, 2014) showed that the DINA model can be recast as a compensatory, nonconjunctive model by transforming the skill space and deriving a mapping of the original skills and the Q-matrix onto the transformed space. Two variants of this equivalency were presented, and the equivalency was proven mathematically. That is, the DINA can be transformed in two different ways, and each variant can be written as a special case of the GDM. Estimation of the DINA equivalent model using the GDM produces parameter estimates that are identical up to differences in convergence criteria, and the skill-space restrictions can be relaxed such that the DINA-equivalent GDM can be tested against less constrained GDMs (von Davier, 2014). I do not include details of the approach here because a variant of this approach is described in detail and used in the next section to prove that the LCDM is a special case of the GDM.

The Main Result: The Log-Linear Cognitive Diagnostic Model (LCDM) Is a Special Case of the General Diagnostic Model (GDM)

With the notations introduced earlier, let $\boldsymbol{a}^*$ denote a t-dimensional skill vector that we refer to as the transformed skill space. Similarly, let $\boldsymbol{q}_i^*$ denote the t-dimensional transformed Q-matrix entry for item i. We define the entries of the transformed $\boldsymbol{q}_i^*$ for a given source $\boldsymbol{q}_i$ and the distributional constraints required for $\pi(\boldsymbol{a}^*)$ in the following. More specifically, the extended Q-matrix $\boldsymbol{Q}^*$ is defined as follows:

Let

$$k_{Dv} = \binom{D}{v} = \frac{D!}{v!\,(D - v)!} \tag{16}$$

denote the number of transformed skill columns required for each interaction of order $v-1$. For main effects, these can be considered as interactions of 0th order, that is, $v=1$, and for first-order interactions, we have $v=2$, and so on. The index v shows how many skills are involved, and $k_{Dv} = \binom{D}{v}$ is the potential number of these types of interactions. For example, consider a Q-matrix with four skills: There are four main effects, one for each skill, corresponding to the λ_{id} parameters. There are $\binom{4}{2} = 6$ interactions involving two skills (for the $\lambda_i^{d,e}$ parameters) and $\binom{4}{3} = 4$ interactions involving three skills (for the $\lambda_i^{d,e,f}$ parameters), and one, $\binom{4}{4} = 1$, that involves all four skills, $\lambda_i^{1,2,3,4}$. Implicitly, there is also $\binom{4}{0} = 1$ zero-skill "interaction" that corresponds to the baseline item parameter λ_{i0} present in all items. Note that $1+4+6+4+1=16=2^4$ in this example. More generally, we find for the number of transformed skills that

$$D^* = \sum_{v=1}^{D} k_{Dv} = 2^D - 1 \text{ since} \sum_{v=0}^{D} \binom{D}{v} = 2^D. \tag{17}$$

The equality on the right is a widely known result that finds applications, among other things, in the binomial distribution. For the purpose of mapping the main effects and first-order (two-skill) interactions, define

$$q^*_{id} = q_{id} \text{ for } 0 < d < D = k_{D1}, \tag{18}$$

and

$$q^*_{is} = q_{id}\, q_{ie} \text{ with } s = k_{D1} + \text{rank}\left((d,e) : d < e \in \{1, \ldots, D\}\right), \tag{19}$$

where $\text{rank}((d,e) : d < e \in \{1, \ldots, D\})$ is a well-defined function that rank orders tuples (d,e) with $d<e$ from the set $\{1, \ldots, D\}$. Similarly, for three-skill (second-order) interactions, let

$$q^*_{it} = q_{id}\, q_{ie} q_{if}$$

$$\text{with}$$

$$t = k_{D1} + k_{D2} + \text{rank}\left(\left(d,e,f\right) : d < e < f \in \{1, \ldots, D\}\right), \tag{20}$$

and so on, for Q-matrix entries $\boldsymbol{q}_i$ and assumed four or more skill interactions. This assignment is unique and well defined and produces Q-matrix entries $\boldsymbol{q}_i^*$ based on a transformed set of $D^\star = 2^D - 1$ skills. Obviously, this produces a Q-matrix that includes all potential skill interactions and represents a GDM that is equivalent to what is called the "saturated" item model in LCDM and G-DINA. Obviously, this matrix can be reduced by removing interactions that are not required as well as all-zero columns that may occur if the Q-matrix for a given test does not contain all possible skill combinations.

A general form for all interactions, that is, all transformations together, can be written as a function from $\boldsymbol{q}_i \in \{0,1\}^D$ to $\boldsymbol{q}_i^* \in \{0,1\}^{D^*}$, such that

$$\boldsymbol{q}_i^* = G\left(\boldsymbol{q}_i\right), \text{ with } q^*_{it} = \prod_{\left(j_1 < \ldots < j_g\right) \in \{1, \ldots, D\}^g} q_{ij_r},$$

$$\text{where}$$

$$t = \sum_{v=1}^{g} \binom{D}{v} + \text{rank}\left(\left(j_1 < \ldots < j_g\right) \in \{1, \ldots, D\}^g\right)$$

$$\text{for all } g = 1, \ldots, D. \tag{21}$$

This function $G(\boldsymbol{q}_i)$ transforms all Q-matrix entries of a log-linear cognitive diagnosis model (LCDM) source matrix into entries based on an extended skill space for the LCDM-equivalent version written as a constrained GDM.

The proof showing that this produces an LCDM-equivalent compensatory skill space that is used in the GDM becomes evident when inspecting Equation 21. For each of the potential skill interactions in the source LCDM, Equation 21 defines the required skill entries in the transformed skill space. The transformation in Equation 21, together with the general definition in Equations 8 and 9, defines an item model that is equivalent to the LCDM definition given in Equation 10 but that operates on a transformed additive/compensatory skill space.

The same function can be used to define the constrained skill space for this LCDM-equivalent GDM. Because Equation 21 defines function mapping binary vectors from $\{0,1\}^D$ to $\{0,1\}^{D^*}$, we can utilize $G(.)$ for skill patterns as well as Q-matrix vectors (as both are elements of $\{0,1\}^D$). Define

$$\mathrm{CS}_D = \left\{ \boldsymbol{a}^* = G(\boldsymbol{a}) : \boldsymbol{a} \in \{0,1\}^D \right\} \subseteq \{0,1\}^{D^*}, \tag{22}$$

as the set of transformed skill vectors $\boldsymbol{a}^*$ in $\{0,1\}^{D^*}$ that correspond to a skill vector $\boldsymbol{a}$ in $\{0,1\}^D$. Now, we define a measurable function on the constrained skill space CS_D, which is a subset of $\{0,1\}^{D^*}$, rather than on the whole set. That is, we use constrained skill distributions over $\{0,1\}^{D^*}$ that have positive mass on CS_D only.

More formally, let

$$\pi\left(\boldsymbol{a}^*\right) > 0 \text{ iff } \boldsymbol{a}^* \in \mathrm{CS}_D$$

with

$$\sum_{\boldsymbol{a}^* \in \mathrm{CS}_D} \pi\left(\boldsymbol{a}^*\right) = 1 \text{ and } \pi\left(\boldsymbol{a}^*\right) = 0 \text{ iff } \boldsymbol{a}^* \notin \mathrm{CS}_D. \tag{23}$$

The condition expressed in Equation 23 ensures that the resulting constrained GDM defined on the transformed skill space is equivalent to the source LCDM. However, if it turns out that the LCDM does not fit the data, the constraints defined in Equation 23 can be relaxed and a GDM that is more general than the LCDM on the transformed skill space can be utilized.

An Application

One curious observation related to research on diagnostic models is that many articles tend to use the same data sets. One is the widely used fraction subtraction data set, originally described in the seminal work of Kikumi Tatsuoka (1983); the other is the Examination for the Certificate of Proficiency in English (ECPE) data set, which is more recent but has been used in various articles.

The frequent reuse of only a handful of data sets in the field of diagnostic modeling is in part due to the availability of a Q-matrix associated with the data. However, this practice might be counterproductive to the field because practitioners may come to one of the following conclusions:

1. There may be a few data sets for which any diagnostic model produces skill estimates that can be made sense of, or can be considered preferable over model estimates using customary unidimensional latent traits or other types of simpler score variables.
2. Researchers developing new diagnostic models are not that interested in applying these models to current problems in educational and psychological measurement. Instead, they are more driven by "scoring" the next publication, whether or not the model developed in the process is useful for a range of real data applications.

To avoid these types of potential concerns regarding limited utility or practical applicability, we agree it would be preferable to present work on diagnostic models that is driven by answering concrete research questions using data collected in current assessment programs. But we feel it is appropriate in this instance to reuse the ECPE data because they will allow us to maintain comparability with published results, keeping two goals in mind: (a) to show that the LCDM is, and can be estimated as, a constrained GDM, and hence no special purpose implementation or treatment of the LCDM is required; and (b) to demonstrate that even for this widely used data set, the tools available through the GDM framework, as described in von Davier (2005, 2008, 2011, 2014), can be used to gain a deeper understanding of the appropriateness of diagnostic models in general, and the LCDM in particular, for this and other data sets. The tools available in the GDM software *mdltm* (von Davier, 2005) allow evaluation of the fit of the LCDM compared to a range of other models and a check on whether this model is identified for the ECPE data.

The ECPE data are available in the R package CDM (Robitzsch, Kiefer, George, & Ünlü, 2011–2014) and online (see Templin, 2012), which provides materials on a workshop conducted in 2012. Analyses of this data set with the LCDM were described in Templin and Bradshaw (2014) and Templin and Hoffman (2013). Chiu (2008) also used the ECPE data but

Table 1 Transformed Q-Matrix for the Examination for the Certificate of Proficiency in English (ECPE) Data

	Skills						
Items	1	2	3	4	5	6	~~7~~
10, 13, 14, 25, 27	1	0	0	0	0	0	0
2, 8, 23, 24	0	1	0	0	0	0	0
4, 5, 6, 9, 15, 18, 19, 22, 26, 28	0	0	1	0	0	0	0
1	1	1	0	1	0	0	0
3, 7, 11, 12, 16, 20, 21	1	0	1	0	1	0	0
17	0	1	1	0	0	1	0
~~NA~~	~~1~~	~~1~~	~~1~~	~~1~~	~~1~~	~~1~~	~~1~~

studied different modeling approaches, and Robitzsch et al. (2011 – 2014) presented exemplary analyses with the ECPE data for Robitzsch's R package, CDM. The ECPE file contains responses from 2,922 test takers to 28 items, each of which is associated to one or two out of a total of three skills.

Comparing a Log-Linear Cognitive Diagnostic Model (LCDM)–Equivalent General Diagnostic Model (GDM) With Published LCDM Results

The LCDM-equivalent GDM was used to estimate the parameters of the LCDM for the ECPE data using the transformed Q-matrix given in Table 1. All analyses were carried out with the software *mdltm* (von Davier, 2005) and compared against results published in Templin and Bradshaw (2014); Templin and Hoffman (2013), and Robitzsch et al. (2011 – 2014). The software *mdltm* is available on request based on a license agreement that allows noncommercial use for research and teaching.

None of the items in the ECPE Q-matrix required all three skills, so the transformed Q-matrix requires only $2^3 - 2 = 6$ columns. Also note that two of the required-skill combinations appear on single items only in the Q-matrix used with LCDM for the ECPE data.

The transformed skills space is given by $\{G(\boldsymbol{a}) : \boldsymbol{a} \in \{0, 1\}^3\}$ according to Equation 22 and is reduced by eliminating the last column. Note that this does not change the distribution, because the original skill space still contains (1,1,1), which is mapped onto either (1,1,1,1,1,1,1) or the reduced (1,1,1,1,1,1) with six skills, respectively. The last column represents the interaction of three skills and hence is obsolete (and its entries crossed out) because the original Q-matrix contains at most two nonzero skill entries. The constrained skill distribution is given, as defined in Equation 23, by $\pi(\boldsymbol{a}^*) > 0$ for all $\boldsymbol{a}^* \in \mathrm{CS}_3$ and $\pi(\boldsymbol{a}^*) = 0$ for $\boldsymbol{a}^* \notin \mathrm{CS}_3$.

When estimating the equivalent model, the parameters of the LCDM-equivalent GDM were initially not range constrained, whereas in a second run, range constraints required for identification that are equivalent to those found in Templin and Bradshaw (2014); Templin and Hoffman (2013), and Robitzsch et al. (2011 – 2014) were imposed on three items.

The second set of analyses was needed because results from the unconstrained analysis suggest a problem with identifiability. Among other things, some very large standard errors (e.g., model-based *SE* 1.83, jackknife-based *SE* 5.93 for two parameters of item 12) were obtained, and the condition number was 3.64e-08, which is indicative of identification issues (Muthén & Muthén, 2012). In addition, the correlation of certain pairs of parameters was found to be −0.99, which is another indicator of lack of parameter identification.

Therefore we only report the results on constrained analyses here. We report parameter estimates for the LCDM-equivalent GDM and associated standard errors, the condition number, and parameter correlations and compare these parameter estimates with the values published by Templin and Hoffman (2013), which agree with those obtained from the CDM R-package (Robitzsch et al., 2011 – 2014) and the parameters available from output files on the workshop Web site for the data set (Templin, 2012).

Table 2 contains a comparison of the skill distribution published in the sources listed, whereas Table 3 presents an overview of the item location and slope (λ) parameters obtained from estimating the constrained equivalent GDM. Both tables include the standard errors. Estimates were obtained using the jackknife procedure available in *mdltm* (von Davier, 2005).

Table 2 Model Parameters Obtained From the Constrained Analysis of the Examination for the Certificate of Proficiency in English (ECPE) Data With the Log-Linear Cognitive Diagnostic Model (LCDM) – Equivalent General Diagnostic Model (GDM)

Item	s1	s1	s2	s2	s3	s3	s4	s4	s5	s5	s6	s6	Diff	Diff
1*	***0.00***	*	0.60	[0.16]	–	–	1.22	[0.24]	–	–	–	–	0.83	[0.07]
2	–	–	1.24	[0.13]	–	–	–	–	–	–	–	–	1.03	[0.08]
3	0.75	[0.24]	–	–	0.34	[0.10]	–	–	0.53	[0.26]	–	–	−0.34	[0.06]
4	–	–	–	–	1.69	[0.10]	–	–	–	–	–	–	−0.13	[0.08]
5	–	–	–	–	2.01	[0.14]	–	–	–	–	–	–	1.08	[0.08]
6	–	–	–	–	1.69	[0.13]	–	–	–	–	–	–	0.86	[0.08]
7*	2.85	[0.19]	–	–	0.95	[0.12]	–	–	***−0.95***	*	–	–	−0.10	[0.07]
8	–	–	1.92	[0.25]	–	–	–	–	–	–	–	–	1.48	[0.09]
9	–	–	–	–	1.19	[0.09]	–	–	–	–	–	–	0.11	[0.06]
10	2.05	[0.14]	–	–	–	–	–	–	–	–	–	–	0.05	[0.06]
11	0.82	[0.29]	–	–	0.96	[0.12]	–	–	0.77	[0.30]	–	–	−0.03	[0.07]
12*	***0.00***	*	–	–	1.29	[0.11]	–	–	1.51	[0.12]	–	–	−1.76	[0.09]
13	1.63	[0.15]	–	–	–	–	–	–	–	–	–	–	0.66	[0.06]
14	1.36	[0.12]	–	–	–	–	–	–	–	–	–	–	0.17	[0.05]
15	–	–	–	–	2.11	[0.14]	–	–	–	–	–	–	0.99	[0.09]
16	2.32	[**0.90**]	–	–	0.89	[0.11]	–	–	−0.84	[**0.91**]	–	–	−0.10	[0.07]
17	–	–	0.77	[0.37]	0.59	[0.21]	–	–	–	–	0.06	[0.39]	1.35	[0.05]
18	–	–	–	–	1.38	[0.11]	–	–	–	–	–	–	0.92	[0.08]
19	–	–	–	–	1.84	[0.10]	–	–	–	–	–	–	−0.19	[0.08]
20	0.250	[0.33]	–	–	0.90	[0.11]	–	–	1.40	[0.32]	–	–	−1.39	[0.08]
21	1.054	[0.29]	–	–	1.13	[0.12]	–	–	0.04	[0.30]	–	–	0.16	[0.07]
22	–	–	–	–	2.24	[0.10]	–	–	–	–	–	–	−0.87	[0.09]
23	–	–	2.07	[0.17]	–	–	–	–	–	–	–	–	0.66	[0.08]
24	–	–	1.52	[0.10]	–	–	–	–	–	–	–	–	−0.67	[0.08]
25	1.136	[0.11]	–	–	–	–	–	–	–	–	–	–	0.09	[0.05]
26			–	–	1.11	[0.09]	–	–	–	–	–	–	0.16	[0.06]
27	1.713	[0.10]	–	–	–	–	–	–	–	–	–	–	−0.88	[0.06]
28	–	–	–	–	1.74	[0.11]	–	–	–	–	–	–	0.56	[0.08]

Note. Items with constrained parameters are marked with an asterisk; fixed parameters are given in bold + italics font. Standard errors are given in parentheses, large values in bold font. The root mean squared deviation of parameter estimates compared to parameters published elsewhere is 0.004, indicating close agreement.

Table 3 Skill Distributions (With Standard Errors) Obtained for the Log-Linear Cognitive Diagnostic Model (LCDM) – Equivalent General Diagnostic Model (GDM) Compared to Publicly Available Estimates

Skill pattern	LCDM (T&H)	LCDM (GDM)	Δ	*SE*
(0,0,0)	0.3007	0.3006	0.0002	0.0149
(1,0,0)	0.1290	0.1288	0.0002	0.0173
(0,1,0)	0.0119	0.0119	0.0000	0.0101
(1,1,0)	0.1751	0.1750	0.0001	0.0174
(0,0,1)	0.0087	0.0090	−0.0002	0.0046
(1,0,1)	0.0182	0.0182	0.0000	0.0078
(0,1,1)	0.0108	0.0110	−0.0002	0.0053
(1,1,1)	0.3462	0.3457	0.0004	0.0157

Note. T&H = Templin & Hoffman.

The root mean squared deviation between estimates published by Templin and Hoffman (2013) and the LCDM estimated using an LCDM-equivalent GDM is

$$\sqrt{\frac{1}{\#\mathrm{par}}\sum_{i,d^*}\left(\lambda_{id^*\mathrm{pub}} - \lambda_{id^*\mathrm{LCDM(GDM)}}\right)^2} = 0.004,$$

indicating close agreement between the two sets of estimates. The sum of squared differences runs over all estimated (not constrained or by definition zero) parameters, and #par denotes the number of (previously published and equivalent model-based) estimated parameters.

Note, however, that the maximum standard error is about 0.9 for two parameters associated with Item 16, so there may still be issues with weak identifiability of the LCDM.

Evidently, the LCDM-equivalent GDM produces parameter estimates that are practically identical to the published values. This is somewhat surprising because the LCDM seems weakly identified at best, as the condition number for the constrained LCDM is 1.74e-06, and the (in absolute terms) highest correlation between two parameter estimates is −0.977.

Note that the weak identification is not a consequence of using an equivalent model specification (because the models are mathematically equivalent but use transformed parameterizations). The weak identification is rather a consequence of the assumptions made in the LCDM relative to the empirical evidence present in the ECPE data. It appears that a model with multiple skills and interactions between skills is not supported by the data. von Davier and Haberman (2014) pointed out that the skill distribution presented for the ECPE data set provides support for a single ordered latent trait or a continuous latent trait being sufficient.

Examining Table 3 shows that only four of the eight skill vectors carry substantive mass. The other four estimates are numerically rather small, suggesting that four nonzero probabilities associated with the skill patterns (0,0,0), (1,0,0), (1,1,0), and (1,1,1) would suffice to fit the data. This implies a strict order of skills and suggests that a model with a limited number of ordered latent classes may be sufficient to fit the data. von Davier and Haberman (2014) pointed this out and conjectured that the skill distribution presented for the ECPE data set provides support for a single underlying skill, not multiple skills.

In the next section, we delve deeper into the issue of appropriateness of the LCDM for the ECPE data by comparing the model data fit of the LCDM with more parsimonious models, including item response theory (IRT) and located latent class models.

Comparison of the Log-Linear Cognitive Diagnostic Model (LCDM) Model Data Fit to Other Models

This section provides two additional lines of analysis. First, the model constraints required to make a weakly identifiable model out of the unidentified unconstrained LCDM are further investigated. Second, the constrained-LCDM equivalent model is compared to customary IRT models, to a three-skill GDM, and to two- and three-skill exploratory item factor analysis models on the basis of the GDM.

The constraints found in the Mplus-based estimation of the LCDM provided by Templin and Bradshaw (2014) as well as Templin and Hoffman (2013) and in the R-package CDM are as follows:

1. Ensure that all skill main effects are nonnegative: $\lambda_{id} > 0$.
2. Ensure that the negative skill interaction is not larger than the main effects of any of the involved skills: $-\lambda_i^{d,e} < \lambda_{id}$ and $-\lambda_i^{d,e} < \lambda_{ie}$.

These were also implemented (on the corresponding compensatory skill variables) during estimation by means of the LCDM-equivalent (constrained) D^* -skill GDM to ensure comparability with published results and to improve identifiability. However, though these constraints were effective in reducing the threat of lack of identifiability, the choice of these constraints is somewhat arbitrary, as Table 4 shows.

There may be a different rationale that assumes skill main effects can be negative (unless combined with the other required skill) or that skill interactions lead to a probability of success that is lower than in the presence of one or the other, but not both skills, for example, if skills represent two alternative but not compatible strategies to solve a task. That means that different rationales for different skill constraints may be found, so it is prudent to compare the effects of different constraints on the results. In Table 4, eight alternative constraints imposed on the same parameters as marked constrained in Table 2 are compared. The comparison involved the effect on the estimated skill distribution as well as the achieved log-likelihood under the different constraints.

As can be seen when comparing the log-likelihood of the models associated with different constraints to improve identifiability, none of these have profound effects on the fit of the model. If anything, the constraint that assigns −1.0 to all three parameters that were treated to achieve weak identifiability results in the largest log-likelihood. However, the difference from the log-likelihood based on published results is marginal. The constraint variants with more negative skill main effects have a small advantage over the constraint that assumes vanishing main effects, while the effect on skill distributions is likewise marginal. This result gives credence to any (or none) of these constraints, as they can hardly be distinguished empirically, and each of the constraints implies somewhat different interpretations of the resulting LCDM

Table 4 A Comparison of the Effects of Eight Different Constraints (a to h), in Addition to the One Previously Published, Imposed on Items 1, 7, and 12 Used to Achieve (Weak) Identifiability of the Log-Linear Cognitive Diagnostic Model (LCDM) for the Examination for the Certification of Proficiency in English (ECPE) Data

	Different model constraints to achieve (weak) identifiability								
Effect	pub	a	b	c	d	e	f	g	h
I01 S1	0.000	−0.500	−0.200	−0.700	−0.100	−0.400	**−1.000**	−0.100	−0.800
I12 S1	0.000	−0.500	−0.200	−0.100	−0.700	−0.400	**−1.000**	−0.100	−0.800
I07 S5	−0.952	−0.952	−0.952	−0.952	−0.520	−1.320	**−1.000**	−1.000	0.000
	Effects on fit and skill distribution								
Distr.	pub	a	b	c	d	e	f	g	h
π (000)	0.301	0.302	0.301	0.302	0.301	0.302	**0.302**	0.300	0.303
π (100)	0.129	0.127	0.128	0.126	0.128	0.127	**0.125**	0.129	0.125
π (010)	0.012	0.011	0.012	0.012	0.011	0.012	**0.012**	0.012	0.010
π (110)	0.175	0.177	0.176	0.178	0.175	0.177	**0.180**	0.175	0.179
π (001)	0.009	0.005	0.007	0.005	0.007	0.006	**0.002**	0.010	0.003
π (101)	0.018	0.018	0.018	0.017	0.019	0.018	**0.016**	0.018	0.017
π (011)	0.011	0.013	0.012	0.013	0.013	0.012	**0.015**	0.011	0.016
π (111)	0.346	0.347	0.346	0.347	0.346	0.346	**0.349**	0.346	0.348
LogLik	−42739.7	−42739.2	−42739.4	−42739.4	−42739.4	−42739.2	**−42739.0**	−42739.9	−42739.4

Note. The constraints used in variant *f* (boldfaced column) achieve the best fit.

Table 5 Comparison of Model–Data Fit Among Various Different Models

Model	#par	Likelihood	AIC	BIC
2PL	56	−42546.84	85205.68	**85540.56**
LLCA	65	−42516.52	85163.04	85551.74
GDM	72	−42744.76	85633.52	86064.08
GDM3	91	−42503.34	85188.68	85732.86
IFA 2	86	−42520.47	85212.95	85727.23
IFA 3	112	**−42432.34**	**85088.68**	85758.45
LCDM	81	−42738.74	85639.47	86123.85
LCDM2F	78	−42739.72	85635.43	86101.87
LCDM_FIX	81	−42739.71	85641.43	86125.81
?*JT_LCDM*?	81	−42487.40	85136.80	85621.18

Note. AIC = Akaike information criterion; BIC = Bayesian information criterion. Models compared: 2PL IRT, located latent class (LLCA) with 12 ordered levels, general diagnostic model (GDM) with binary (0,1) latent variables, GDM3 with trichotomous (−1, 0, +1) latent variables, item factor analysis with 2 or 3 binary latent variables, and constrained (LCDM2F) as well as unconstrained log-linear cognitive diagnostic model (LCDM), and an LCDM_FIX with parameters fixed to published estimates. Numbers in boldface indicate the minimum in each column.

estimates. Therefore, a decision about which set of constraints is most defensible or sensible remains mainly a matter of interpretation, not a question that can be answered by quantitative methods. The differences between the skill distribution and goodness of fit results are small, whereas the differences between estimates of parameters fixed at certain values are substantial but almost inconsequential owing to the lack of identifiability of the underlying unconstrained LCDM.

Moving to the comparison of the LCDM with customary IRT models and diagnostic models that require or imply no skill interactions, Table 5 presents the log-likelihood and the associated information criteria—the Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwarz, 1978).

The fit indices reported in Table 5 show that neither the unconstrained (but not identified) nor the weakly identified constrained LCDM are among the preferred models. The three-skill GDM achieves lower, and hence more desirable, fit indices than the constrained and unconstrained LCDM variants. The equivalent LCDM-estimated parameters, as well as the LCDM with parameters fixed to previous values, achieve practically the same values, which coincide with what Templin and Bradshaw (2014)[1] report as well as with the LCDM (Mplus-based) output on the LCDM workshop Web site (Templin, 2012) and with the values calculated with the CDM R-package by Robitzsch et al. (2011–2014).

The relatively best fitting model, when looking at the AIC, is the exploratory item factor analysis model with three latent variables (IFA3), whereas the preferred model when looking at the BIC is the unidimensional two-parameter-logistic IRT model. The BIC was found to be superior compared to the AIC in terms of identification of the correct model (Bozdogan, 1987; Li, Cohen, Kim, & Cho, 2009 ; Nylund, Asparouhov, & Muthén, 2007; Schwarz, 1978).

Conclusion

This report shows an example of how a general class of models can be used to facilitate understanding that a model including complex skill interactions may not be uniquely defined. This means that alternative parameterizations that are mathematically equivalent but based on or representing a different set of assumptions may exist. More specifically, we showed that any model the LCDM can represent, including the DINA and the saturated LCDM, can be recast as a GDM with a transformed set of skills that do not involve skill interactions. This LCDM-equivalent GDM can be used to provide parameter estimates that agree with other approaches to estimate the LCDM. In addition, however, the software implementation of the GDM allows an assessment of model identification of the LCDM, and the LCDM can be tested against more parsimonious models such as customary IRT models or located latent class models, which provide, for the case studied here, better model–data fit than the LCDM.

This report also shows that a constrained version of the LCDM that appears to provide reasonable results can, on closer inspection, be shown to achieve identification only after making some additional, largely arbitrary adjustments in terms of parameter constraints (see Table 4).

The need for replication and model comparisons cannot be overstated. New models are proposed on a frequent basis, but these new developments have to be studied rigorously. This includes (a) close inspection regarding potential overlap or even equivalency to other models as well as (b) an empirical comparison to customary models that involve less complex structures. Such comparisons should ideally be carried out on real data, because simulated data can only give us answers that we already provided in the conditions selected and the models used to generate the data.

Real data, however, come in two variants: (a) data that have been used previously and were selected mainly because the data appear to fit the model and (b) data that are of genuine interest and are collected to answer a research question, not merely to serve only the purpose of being the data that happened to be there for an apparently new model. In this report, data of Type a were used to deliver comparisons of Types a and b, as described. It appears necessary to use existing data to deliver the comparisons of Types a and b that were either missing or misleading owing to incomplete or incongruent[1] reporting of model–fit indices in previous publications.

The model comparisons and the main equivalency result presented in this report extend previous findings showing that the assumptions made in cognitive diagnosis do not lead to uniquely defined Q-matrices (Maris & Bechger, 2009). Moreover, the result presented here extends the equivalency result derived for the DINA (von Davier, 2011, 2014). This new result shows that the assumptions underlying the LCDM do not lead to a unique specification of a model, because multiple skill-space definitions and different conjunctive and nonconjunctive modeling approaches can be devised that produce the same model-based probabilities of observed responses.

For the practice of using these models, these results indicate that interpretations made based on model choices can be contrasted by rather different alternative conceptualizations and specifications of model structures that do not assume conjunctive skills. If there are multiple latent structures that lead to the same model-based response probabilities, interpretations that use one specific variant as evidence of certain response processes or cognitive patterns do not appear to be on solid ground.

Note

1 However, Templin and Hoffman (2013) report a more *optimistic* AIC comparing the LCDM to the 2PL. That value appears to indicate a lower AIC for the LCDM than the 2PL. The AIC cannot be replicated, and it differs from the AIC reported by Templin and Bradshaw (2014) and the analyses provided on the 2012 workshop site, and from the Robitzsch R-package, which all coincide with the values reported for the LCDM-equivalent GDM in this report. Also, when recreating the BIC based on incongruent AIC in the 2013 publication, the 2PL would still be favored over the LCDM.

References

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

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. doi: 10.1109/TAC.1974.1100705

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.

Chiu, C. Y. (2008). *Cluster analysis for cognitive diagnosis*: *Theory and applications*. Unpublished doctoral dissertation, Rutgers, New Brunswick, NJ.

DeCarlo, L. T. (2011). On the analysis of fraction subtraction data: The DINA model, classification, latent class sizes, and the Q-matrix. *Applied Psychological Measurement*, *35*, 8–26.

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, *76*, 179–199.

Haberman, S. J. (1978). *Analysis of qualitative data* (Vol. 1). New York, NY: Academic Press.

Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, *26*, 301–323.

Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log linear models with latent variables. *Psychometrika*, *74*, 191–210.

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.

Li, F., Cohen, A. S., Kim, S.-H., & Cho, S.-J. (2009). Model selection methods for mixture dichotomous IRT models. *Applied Psychological Measurement*, *33*, 353–373.

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, G., & Bechger, T. (2009). Equivalent diagnostic classification models. *Measurement: Interdisciplinary Research and Perspective*, *7*(1), 41–46.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modelling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 535–569.

Robitzsch, A., Kiefer, T., George, A. C., & Ünlü, A. (2011–2014). *R-package documentation for CDM*: *Cognitive diagnosis modeling*. Retrieved from http://cran.r-project.org/web/packages/CDM/index.html

Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York, NY: Guilford.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464. doi: 10.1214/aos/1176344136.MR468014

Suppes, P., & Zanotti, M. (1981). When are probabilistic explanations possible? *Synthese*, *48*, 191–199.

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

Templin, J. (2012). *NCME workshop website—with Mplus based LCDM syntax and ECPE data*. Retrieved from http://psych.unl.edu/jtemplin/teaching/dcm/dcm12ncme/

Templin, J., & Bradshaw, L. (2014). Hierarchical diagnostic classification models: A family of models for estimating and testing attribute hierarchies. *Psychometrika*, *79*(2), 317–339.

Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational Measurement: Issues and Practice*, *32*, 37–50.

Vermunt, J. K. (1997). *LEM 1.0: A general program for the analysis of categorical data*. Tilburg, the Netherlands: Tilburg University.

von Davier, M. (2005). *A general diagnostic model applied to language testing data* (Research Report No. RR-05-16). Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1348/000711007x193957

von Davier, M. (2007). *Hierarchical general diagnostic models* (Research Report No. RR-07-19). Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1002/j.2333-8504.2007.tb02061.x

von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, *61*, 287–307.

von Davier, M. (2009). Some notes on the reinvention of latent structure models as diagnostic classification models. *Measurement: Interdisciplinary Research and Perspectives*, *7*(1), 67–74.

von Davier, M. (2010). Hierarchical mixtures of diagnostic models. *Psychological Test and Assessment Modeling*, *52*(1), 8–28, Retrieved from http://www.psychologie-aktuell.com/fileadmin/download/ptam/1–2010/02_vonDavier.pdf

von Davier, M. (2011). *Equivalency of the DINA model and a constrained general diagnostic model* (Research Report No. RR-11-37). Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1002/j.2333-8504.2011.tb02273.x von Davier, M. (2014). The DINA model as a constrained general diagnostic model—Two variants of a model equivalency. *British Journal of Mathematical and Statistical Psychology*, *67*(1), 49–71. doi: 10.1111/bmsp.12003

von Davier, M., & Haberman, S. (2014). Diagnostic classification models morphing into unidimensional "diagnostic" classification models—A commentary. *Psychometrika*, *79*, 340–346. doi: 10.1007/s11336-013-9363-z

von Davier, M., & Yamamoto, K. (2004). *A class of models for cognitive diagnosis*. Paper presented at the 4th Spearman conference "Diagnostics for Education: Theory, Measurement, Applications," Philadelphia, PA.

Xu, X., & von Davier, M. (2008). *Comparing multiple-group multinomial loglinear models for multidimensional skill distributions in the general diagnostic model* (Research Report No. RR-08-35). Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1002/j.2333-8504.2008.tb02121.x

Suggested citation:

von Davier, M. (2014). *The log-linear cognitive diagnostic model as a special case of the general diagnostic model* (ETS Research Report No. RR-14-40). Princeton, NJ: Educational Testing Service. doi:10.1002/ets2.12043

Action Editor: James Carlson

Reviewers: Xueli Xu and Shelby Haberman