Journal of Classification 33:171-184 (2016)
DOI: 10.1007/s00357-016-9202-x

A Proof of the Duality of the DINA Model and the DINO Model

Hans-Friedrich Köhn

University of Illinois at Urbana-Champaign, U.S.A.

Chia-Yi Chiu

Rutgers, the State University of New Jersey, U.S.A.

Abstract: The Deterministic Input Noisy Output "AND" gate (DINA) model and the Deterministic Input Noisy Output "OR" gate (DINO) model are two popular cognitive diagnosis models (CDMs) for educational assessment. They represent different views on how the mastery of cognitive skills and the probability of a correct item response are related. Recently, however, Liu, Xu, and Ying demonstrated that the DINO model and the DINA model share a "dual" relation. This means that one model can be expressed in terms of the other, and which of the two models is fitted to a given data set is essentially irrelevant because the results are identical. In this article, a proof of the duality of the DINA model and the DINO model is presented that is tailored to the form and parameterization of general CDMs that have become the new theoretical standard in cognitively diagnostic modeling.

Keywords: Cognitive diagnosis; DINA model; DINO model; General cognitive diagnosis models.

1. Introduction

The Deterministic Input Noisy Output "AND" gate (DINA) model (Junker and Sijtsma 2001; Macready and Dayton 1977) and the Deterministic Input Noisy Output "OR" gate (DINO) model (Templin and Henson

---

Corresponding Author's Address: Hans-Friedrich Köhn, Department of Psychology, University of Illinois, Urbana-Champaign, email: hkoehn@illinois.edu.

Published online: 21 July 2016

2006) are two popular cognitive diagnosis models (CDMs). CDMs for educational assessment (DiBello, Roussos, and Stout 2007; Haberman and von Davier 2007; Leighton and Gierl 2007; Rupp, Templin, and Henson 2010) decompose an examinee's ability in a domain into binary cognitive skills called *attributes*, each of which an examinee may or may not have mastered. Distinct profiles of attributes define different proficiency classes. From the observed item scores, maximum likelihood estimates of the model parameters are obtained that are then used to assign examinees to the different proficiency classes. Software for fitting the DINA model and the DINO model using marginal maximum likelihood estimation via the Expectation Maximization (EM) algorithm (MMLE-EM) is available through the package `CDM` implemented in `R` (Robitzsch, Kiefer, George, and Uenlue 2016).

The DINA model and the DINO model represent different views on how the mastery of attributes and the probability of a correct item response are related. The DINA model is a conjunctive model, meaning that only mastery of all attributes required for an item maximizes the probability of a correct response. In contrast, the DINO model is a disjunctive model, which means that mastery of a subset of the required attributes is a sufficient condition for maximizing the probability of a correct response (for a detailed discussion of these concepts, consult Henson, Templin, and Willse 2009).

Recently, however, Liu, Xu, and Ying (2011) demonstrated that the DINO model and the DINA model share a "dual" relation: One model can be expressed in terms of the other, and which of the two models is fitted to a given data set is essentially irrelevant because, after appropriate transformations, the item parameter estimates are identical (as is shown in detail below) and thus, the estimates of examinees' proficiency class memberships are identical too. This also means that the two models must share the same theoretical properties—what applies to one model automatically holds for the other model. Hence, one proof fits both models, and one set of simulations suffices to cover both models.

General CDMs have become the new theoretical standard in cognitively diagnostic modeling (de la Torre 2011; Henson, Templin, and Willse 2009; Rupp, Templin, and Henson 2010; von Davier 2005, 2008, 2014). In this article, a proof of the duality of the DINA model and the DINO model is presented that is tailored to the form and parameterization of general CDMs. The presentation is preceded by a brief review of some key technical concepts concerning CDMs. As an example of how the duality of the DINA model and the DINO model allows to condense separate proofs for the two models into a single proof, a compact proof of the condition of completeness of the Q-matrix is presented that covers both models. The Discussion summarizes the practical and theoretical implications of the DINA-DINO duality.

2. Review: Key Technical Concepts

2.1 Cognitive Diagnosis Models

Models for cognitive diagnosis are constrained latent class models. Let Y_{ij} denote the observed response of examinee i, $i = 1, \ldots, N$, to binary item j, $j = 1, \ldots, J$. Consider N examinees who belong to M distinct latent proficiency classes. The general latent class model defines the conditional probability of examinee i in proficiency class $\mathcal{C}_m$, $m = 1, \ldots, M$, answering correctly item j by the item response function (IRF), $P(Y_{ij} = 1|i \in \mathcal{C}_m) = \pi_{jm}$, where π_{jm} is constant for item j across all members i in proficiency class $\mathcal{C}_m$. The proficiency-class membership of the examinees is estimated from the observed item responses, Y_{ij}, using either MMLE-EM or Markov chain Monte Carlo (MCMC) techniques. The observed item responses are assumed independent conditional on proficiency-class membership (i.e, local independence). No further restrictions are imposed on the relation between the latent variable—proficiency-class membership—and the observed item response.

In contrast, CDMs constrain the relation between the latent variable and the observed item response such that the membership in a certain proficiency class is associated with the mastery of particular cognitive attributes that in turn determines the item response probabilities. Suppose that K latent binary attributes constitute a given ability domain; there are then 2^K distinct attribute profiles composed of these K attributes representing $M = 2^K$ distinct classes of proficiency. (Note that an attribute profile for a proficiency class can consist of all zeroes, because it is possible for an examinee not to have mastered any attributes at all.) Let the K-dimensional vector, $\boldsymbol{\alpha}_m = (\alpha_{m1}, \ldots, \alpha_{mK})^T$, represent the binary attribute profile of proficiency class $\mathcal{C}_m$, where the k^{th} entry indicates whether the respective attribute has been mastered. The attribute profile of examinee $i \in \mathcal{C}_m$, $\boldsymbol{\alpha}_{i \in \mathcal{C}_m}$, is usually written as $\boldsymbol{\alpha}_i = (\alpha_{i1}, \ldots, \alpha_{iK})^T$. (Throughout the text, the superscript T denotes the transpose of vectors or matrices; the "prime notation" is reserved for distinguishing between vectors or their scalar entries. For brevity, the examinee index, i, is omitted if the context permits; for example, $\boldsymbol{\alpha}_i$ is simply written as $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^T$.)

Consider a test of J items for assessing ability in the domain. Each individual item j is associated with a K-dimensional binary vector, $\mathbf{q}_j$, called item-attribute profile, where $q_{jk} = 1$ if a correct answer requires mastery of the k^{th} attribute, and 0 otherwise. Note that item-attribute profiles consisting entirely of zeroes are inadmissible, because they correspond to items that require no skills at all. Hence, given K attributes, there are at most $2^K - 1$ distinct item-attribute profiles. The J item-attribute profiles of a test constitute its Q-matrix, $\mathbf{Q} = \{q_{jk}\}_{(J \times K)}$, (Tatsuoka 1985), that summarizes the constraints specifying the associations between items and attributes.

2.2 General Cognitive Diagnosis Models

CDMs differ in the way in which mastery and nonmastery of the attributes are believed to affect an examinee's performance on a test item (e.g., compensatory models versus non-compensatory models; conjunctive models versus disjunctive models; for a detailed discussion, see Henson, Templin, and Willse 2009). General CDMs provide a theoretical framework for expressing the functional relation between attribute mastery and the probability of a correct item response in a unified mathematical form and parameterization that are applicable to "recognizable" CDMs (de la Torre 2011, p. 181), as discussed previously in the literature, and CDMs "that have not yet been defined" (Henson, Templin, and Willse 2009, p. 199), thereby establishing a general standard for model comparison and evaluation.

The General Diagnostic Model (GDM; von Davier, 2005, 2008) is the archetypal general CDM. Von Davier defined $h_j = h(\mathbf{q}_j, \boldsymbol{\alpha}_i)$ as a general function of the attribute profile of item j and the attribute profile of examinee i to allow for maximal flexibility in modeling examinees' responses to item j. Presumably, the most popular form of von Davier's GDM uses the logistic function as the link with h_j resulting in the IRF

$$\begin{aligned} P(Y_{ij} = 1 \mid \boldsymbol{\alpha}_i) &= \frac{e^{\beta_{j0} + \boldsymbol{\beta}_j' h(\boldsymbol{q}_j, \boldsymbol{\alpha}_i)}}{1 + e^{\beta_{j0} + \boldsymbol{\beta}_j' h(\boldsymbol{q}_j, \boldsymbol{\alpha}_i)}} \\ &= \frac{e^{\beta_{j0} + \sum_{k=1}^{K} \beta_{jk} q_{jk} \alpha_{ik}}}{1 + e^{\beta_{j0} + \sum_{k=1}^{K} \beta_{jk} q_{jk} \alpha_{ik}}}. \end{aligned}$$

q_{jk} indicates whether mastery of attribute α_{ik} is required for item j. (See Equations 1 and 2; von Davier, 2005.) Henson, Templin, and Willse (2009) specified ν_j as the linear combination of the K attribute main effects, α_k, and all their two-way, three-way, ..., K-way interactions

$$\begin{aligned} \nu_j = h(\mathbf{q}_j, \boldsymbol{\alpha}_i) = \beta_{j0} + \sum_{k=1}^{K} \beta_{jk} q_{jk} \alpha_{ik} + \sum_{k'=k+1}^{K} \sum_{k=1}^{K-1} \beta_{j(kk')} q_{jk} q_{jk'} \alpha_{ik} \alpha_{ik'} \\ + \cdots + \beta_{j12\ldots K} \prod_{k=1}^{K} q_{jk} \alpha_{ik}. \end{aligned}$$

Based on this form of ν_j, Henson, Templin, and Willse (2009) defined the IRF of a general CDM termed the (saturated) Loglinear Cognitive Diagnosis Model (LCDM) as

$$P(Y_{ij}=1\mid\boldsymbol{\alpha}_i)$$
$$=\frac{e^{\beta_{j0}+\sum_{k=1}^{K}\beta_{jk}q_{jk}\alpha_k+\sum_{k'=k+1}^{K}\sum_{k=1}^{K-1}\beta_{jkk'}q_{jk}q_{jk'}\alpha_k\alpha_{k'}+\cdots+\beta_{j12\ldots K}\prod_{k=1}^{K}q_{jk}\alpha_k}}{1+e^{\beta_{j0}+\sum_{k=1}^{K}\beta_{jk}q_{jk}\alpha_k+\sum_{k'=k+1}^{K}\sum_{k=1}^{K-1}\beta_{jkk'}q_{jk}q_{jk'}\alpha_k\alpha_{k'}+\cdots+\beta_{j12\ldots K}\prod_{k=1}^{K}q_{jk}\alpha_k}}.$$

(see Equation 11 in Henson, Templin, and Willse 2009). By imposing appropriate constraints on the β-coefficients in ν_j, the IRFs of specific CDMs can be expressed as (constrained) submodels of the (saturated) LCDM—among them also the DINA model and the DINO model. Recently, von Davier (2011, 2014) demonstrated that by appropriately transforming the Q-matrix the DINA model can be reparameterized as a submodel of the GDM.

2.3 The Deterministic Input Noisy Output "AND" Gate Model

The traditional form of the IRF of the DINA model is $P(Y_j = 1 \mid \boldsymbol{\alpha}) = (1-s_j)^{\eta_j} g_j^{(1-\eta_j)}$ subject to $0 < g_j < 1 - s_j < 1$ for all items j (Junker and Sijtsma 2001; Macready and Dayton 1977). The conjunction parameter η_j is defined as $\eta_j = \prod_{k=1}^{K}\alpha_k^{q_{jk}}$; η_j indicates whether examinee i has mastered all the attributes needed to answer item j correctly. The item-related parameters $s_j = P(Y_j = 0|\eta_j = 1)$ and $g_j = P(Y_j = 1|\eta_j = 0)$ formalize the probabilities of *slipping* (failing to answer item j correctly despite having the skills required to do so) and *guessing*, (answering item j correctly despite lacking the skills required to do so), respectively.

Without loss of generality, assume that $K^* \leq K$ of the K attributes are required for item j; in addition, assume that their indices are ordered such that the K^* attributes occupy the first K^* positions, $1, 2, \ldots, K^*$, of $\mathbf{q}_j$. Because the DINA model is a conjunctive CDM, its conversion to the logit form of the LCDM requires that all attribute main effects and interactions be restricted to zero except for the highest order interaction involving all K^* attributes. Hence, the IRF of the DINA model in terms of the LCDM is

$$\begin{aligned} P(Y_j=1\mid\boldsymbol{\alpha}) &= \frac{e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}} \\ \text{subject to:} \quad & \beta_{j12\ldots K^*} > 0. \end{aligned} \tag{1}$$

(Note that if the K^* attributes occupy the first K^* positions, $1, 2, \ldots, K^*$, of $\mathbf{q}_j$, then $q_{jk} = 1$ is always true and can be dropped from the IRF.) The constraint is implied by the traditional parameterization of the DINA model. From $P(Y_j = 1|\boldsymbol{\alpha}) = g_j$ if $\eta_j = 0$ follows $e^{\beta_{j0}}/\big(1+e^{\beta_{j0}}\big) = g_j$; and $P(Y_j = 1|\boldsymbol{\alpha}) = 1 - s_j$ if $\eta_j = 1$ suggests $e^{\beta_{j0}+\beta_{j12\ldots K^*}}/\big(1+e^{\beta_{j0}+\beta_{j12\ldots K^*}}\big) = 1 - s_j$. The constraint then follows from the restriction $0 < g_j < 1 - s_j < 1$.

2.4 The Deterministic Input Noisy Output "OR" Gate Model

The DINO model (Templin and Henson 2006) is a disjunctive CDM (i.e., mastery of a subset of the required attributes is a sufficient condition for maximizing the probability of a correct item response). Define the disjunction parameter $\omega_j = 1 - \prod_{k=1}^{K}(1-\alpha_k)^{q_{jk}}$ that indicates whether at least one of the attributes associated with item j has been mastered. (Like η_j in the DINA model, ω_j represents the ideal item response when neither slipping nor guessing occurs.) The traditionally parameterized IRF of the DINO model is $P(Y_{ij}=1 \mid \boldsymbol{\alpha}) = (1-s_j)^{\omega_j} g_j^{(1-\omega_j)}$.

Without loss of generality, assume again that $K^* \leq K$ of the K attributes are required for item j; in addition, assume that their indices are ordered such that the K^* attributes occupy the first K^* positions, $1, 2, \ldots, K^*$, of $\mathbf{q}_j$. The condition that mastery of just one attribute of those required for item j already maximizes the probability of a correct response translates into the following constraint to be imposed on the coefficients

$$\begin{aligned}\beta_{j1} &= \beta_{j2} = \ldots = \beta_{jK^*} = (-1)\beta_{j12} = (-1)\beta_{j13}\\ &= \ldots = (-1)\beta_{j(K^*-1)K^*} = \ldots = (-1)^{K^*+1}\beta_{j12\cdots K^*}.\end{aligned}$$

The IRF of the DINO model in terms of the LCDM is then

$$\begin{aligned}&P(Y_j = 1 \mid \boldsymbol{\alpha})\\ &= \frac{e^{\beta_{j0}+\beta_{j1}\sum_{k=1}^{K^*}\alpha_k+(-1)\beta_{j1}\sum_{k'=k+1}^{K^*}\sum_{k=1}^{K^*-1}\alpha_k\alpha_{k'}+\ldots+(-1)^{K^*+1}\beta_{j1}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{\beta_{j0}+\beta_{j1}\sum_{k=1}^{K^*}\alpha_k+(-1)\beta_{j1}\sum_{k'=k+1}^{K^*}\sum_{k=1}^{K^*-1}\alpha_k\alpha_{k'}+\ldots+(-1)^{K^*+1}\beta_{j1}\prod_{k=1}^{K^*}\alpha_k}}\\ &= \frac{e^{\beta_{j0}+\beta_{j1}\left(\sum_{k=1}^{K^*}\alpha_k+(-1)\sum_{k'=k+1}^{K^*}\sum_{k=1}^{K^*-1}\alpha_k\alpha_{k'}+\ldots+(-1)^{K^*+1}\prod_{k=1}^{K^*}\alpha_k\right)}}{1+e^{\beta_{j0}+\beta_{j1}\left(\sum_{k=1}^{K^*}\alpha_k+(-1)\sum_{k'=k+1}^{K^*}\sum_{k=1}^{K^*-1}\alpha_k\alpha_{k'}+\ldots+(-1)^{K^*+1}\prod_{k=1}^{K^*}\alpha_k\right)}}.\end{aligned}$$

In assuming that $1 \leq k \leq K^*$ attributes are mastered, the expression in parentheses

$$\sum_{k=1}^{K^*}\alpha_k + (-1)\sum_{k'=k+1}^{K^*}\sum_{k=1}^{K^*-1}\alpha_k\alpha_{k'} + \ldots + (-1)^{K^*+1}\prod_{k=1}^{K^*}\alpha_k$$

on the right-hand side in the previous equation can be written as

$$\binom{k}{1} + (-1)\binom{k}{2} + \ldots + (-1)^{k+1}\binom{k}{k}.$$

Because

$$\begin{aligned}(-1+1)^k = 0 &= 1 + (-1)\binom{k}{1} + \binom{k}{2} + \ldots + (-1)^k\binom{k}{k}\\ \Rightarrow\ 1 &= \binom{k}{1} + (-1)\binom{k}{2} + \ldots + (-1)^{k+1}\binom{k}{k}.\end{aligned}$$

Hence,

$$P(Y_j = 1 \mid \boldsymbol{\alpha}) \quad = \quad \begin{cases} \frac{e^{\beta_{j0}}}{1+e^{\beta_{j0}}} & \text{if } \boldsymbol{\alpha} = \mathbf{0} \\ \frac{e^{\beta_{j0}+\beta_{j1}}}{1+e^{\beta_{j0}+\beta_{j1}}} & \text{otherwise.} \end{cases}$$

Combining these two conditions into the expression

$$1 - \prod_{k=1}^{K^*}(1-\alpha_k) \quad = \quad \begin{cases} 0 \text{ if } \boldsymbol{\alpha} = \mathbf{0} \\ 1 \text{ otherwise} \end{cases}$$

leads to the compact form of the IRF of the DINO model in terms of the LCDM

$$\begin{aligned} P(Y_j = 1 \mid \boldsymbol{\alpha}) \quad &= \quad \frac{e^{\beta^\circ_{j0}+\beta^\circ_{j1}\left(1-\prod_{k=1}^{K^*}(1-\alpha_k)\right)}}{1 + e^{\beta^\circ_{j0}+\beta^\circ_{j1}\left(1-\prod_{k=1}^{K^*}(1-\alpha_k)\right)}} \\ \text{subject to:} \quad & \quad \beta^\circ_{j1} > 0. \end{aligned} \tag{2}$$

(Henceforth the notation β°_j is used to distinguish the DINO model from the DINA model.) Constraints 1 and 2 are implied by the "traditional" parameterization of the DINO model. From $P(Y_j = 1 \mid \boldsymbol{\alpha}) = g_j$ if $\omega_j = 0$ follows $g_j = e^{\beta^\circ_{j0}}/\left(1 + e^{\beta^\circ_{j0}}\right)$; and from $P(Y_j = 1 \mid \boldsymbol{\alpha}) = 1 - s_j$ if $\omega_j = 1$, follows $1 - s_j = e^{\beta^\circ_{j0}+\beta^\circ_{j1}}/\left(1 + e^{\beta^\circ_{j0}+\beta^\circ_{j1}}\right)$. Because the slipping and guessing parameters are restricted to $0 < g_j < 1 - s_j < 1$, the constraint follows directly.

3. Proof of the Duality of the DINA Model and the DINO Model Based on the Loglinear Cognitive Diagnosis Model

In this section, the proof of the duality of the DINA model and the DINO model is developed. Specifically, it is shown that the two models are technically identical under certain transformations of (a) the examinees' attribute profiles, (b) their observed item scores, and (c) the model parameters. (As an aside, note that the characterization of the special relation of the DINA model and the DINO model as "dual" deviates from the well-defined meaning of this term in operations research; for details, consult Papadimitriou and Steiglitz 1998.)

3.1 Transformation of Examinees' Attribute Profiles

Consider the attribute profile $\boldsymbol{\alpha}^* = \mathbf{1} - \boldsymbol{\alpha}$. Then, the conditional expectation of the DINO model, given $\boldsymbol{\alpha}^*$, is

$$
\begin{aligned}
E_{\text{DINO}}(Y_j \mid \boldsymbol{\alpha}^*) &= \frac{e^{\beta^\circ_{j0}+\beta^\circ_{j1}\left(1-\prod_{k=1}^{K^*}(1-\alpha_k^*)\right)}}{1+e^{\beta^\circ_{j0}+\beta^\circ_{j1}\left(1-\prod_{k=1}^{K^*}(1-\alpha_k^*)\right)}} \\
&= \frac{e^{\beta^\circ_{j0}+\beta^\circ_{j1}\left(1-\prod_{k=1}^{K^*}\alpha_k\right)}}{1+e^{\beta^\circ_{j0}+\beta^\circ_{j1}\left(1-\prod_{k=1}^{K^*}\alpha_k\right)}}.
\end{aligned}
$$

3.2 Transformation of Examinees' Observed Item Response Scores

Let $Y_j^* = 1 - Y_j$; then, the conditional expectation of Y_j^*, given $\boldsymbol{\alpha}^*$, for the DINO model is

$$
\begin{aligned}
E_{\text{DINO}}(Y_j^* \mid \boldsymbol{\alpha}^*) &= 1 - E_{\text{DINO}}(Y_j \mid \boldsymbol{\alpha}^*) \\
&= 1 - \frac{e^{\beta^\circ_{j0}+\beta^\circ_{j1}(1-\prod_{k=1}^{K^*}\alpha_k)}}{1+e^{\beta^\circ_{j0}+\beta^\circ_{j1}(1-\prod_{k=1}^{K^*}\alpha_k)}} \\
&= \frac{1}{1+e^{\beta^\circ_{j0}+\beta^\circ_{j1}(1-\prod_{k=1}^{K^*}\alpha_k)}} \\
&= \frac{e^{-\beta^\circ_{j0}-\beta^\circ_{j1}(1-\prod_{k=1}^{K^*}\alpha_k)}}{1+e^{-\beta^\circ_{j0}-\beta^\circ_{j1}(1-\prod_{k=1}^{K^*}\alpha_k)}} \\
&= \frac{e^{(-\beta^\circ_{j0}-\beta^\circ_{j1})+\beta^\circ_{j1}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{(-\beta^\circ_{j0}-\beta^\circ_{j1})+\beta^\circ_{j1}\prod_{k=1}^{K^*}\alpha_k}}. \qquad (3)
\end{aligned}
$$

3.3 Transformation of the Model Parameters

Compare Equation 3 with the conditional expectation of Y_j, given $\boldsymbol{\alpha}$, for the DINA model

$$
E_{\text{DINA}}(Y_j \mid \boldsymbol{\alpha}) = \frac{e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}}.
$$

Setting $-\beta^\circ_{j0} - \beta^\circ_{j1} = \beta_{j0}$ and $\beta^\circ_{j1} = \beta_{j12\ldots K^*}$ results in

$$
\begin{aligned}
E_{\text{DINO}}(Y_j^* \mid \boldsymbol{\alpha}^*) &= \frac{e^{(-\beta^\circ_{j0}-\beta^\circ_{j1})+\beta^\circ_{j1}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{(-\beta^\circ_{j0}-\beta^\circ_{j1})+\beta^\circ_{j1}\prod_{k=1}^{K^*}\alpha_k}} = \frac{e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}} \\
&= E_{\text{DINA}}(Y_j \mid \boldsymbol{\alpha}), \qquad (4)
\end{aligned}
$$

which completes the proof.

3.4 Results and Conclusions Based on Equation 4

First, the observed item responses, Y_j^*, can be fitted by the DINO model via software for the DINA model using the transformed item responses, $Y_j = 1 - Y_j^*$, as input. Second, the estimates of examinees' attribute profiles in terms of the DINO model, $\widehat{\boldsymbol{\alpha}}^*$, are then computed from the DINA attribute profile estimates, $\widehat{\boldsymbol{\alpha}}$, through the transformation $\widehat{\boldsymbol{\alpha}}^* = \mathbf{1} - \widehat{\boldsymbol{\alpha}}$. Third, the item parameter estimates of the DINO model can be derived from those of the DINA model as $-\widehat{\beta^\circ}_{j0} - \widehat{\beta^\circ}_{j1} = \hat{\beta}_{j0}$ and $\widehat{\beta^\circ}_{j1} = \hat{\beta}_{j12\ldots K^*}$. Obviously, these transformations "work both ways": Y_j^* can be fitted by the DINA model via software for the DINO model using the transformed item responses, $Y_j = 1 - Y_j^*$, as input, and so on.

From a practical point of view, these results appear as a mere curiosity—who would get the idea to fit data with the DINO model through software for the DINA model if routines for fitting the DINO model directly are available? However, these results are relevant because they imply that the DINA model and the DINO model must share the same theoretical properties: What applies to one model automatically holds for the other model. Hence, one proof fits both models, As an illustration, the next section demonstrates how the DINA-DINO-duality allows to condense the separate proofs of the condition of completeness of the Q-matrix into a single proof covering both models.

4. A Proof of Q-Matrix Completeness for the DINA and the DINO Model Using Their Duality

Recall that the $J \times K$ Q-matrix, $\mathbf{Q} = \{q_{jk}\}$, collects the constraints specifying the associations between all J test items and K attributes, where $q_{jk} = 1$ if a correct answer to the j^{th} item requires mastery of the k^{th} attribute, and 0 otherwise. The Q-matrix is an integral component of all CDMs.

A Q-matrix is said to be complete if it allows identification of all possible proficiency classes among examinees. For the DINA model, Chiu, Douglas, and Li (2009) proved that the Q-matrix is complete if and only if it contains all K single-attribute items. (A single-attribute item has the unit vector $\mathbf{e}_k$ as its $\mathbf{q}$-vector, with the k^{th} element equal to 1 and all other elements equal to 0.) (Chiu & Köhn, 2015a, proved that this condition of Q-completeness also applies to the DINO model.) The original proof by Chiu, Douglas, and Li (2009) was tailored to the DINA model and used the traditional parameterization of this model in terms of η, s_j and g_j. This parameterization is not suitable for general CDMs; hence, Chiu and Köhn (2015a) proposed a general definition of Q-completeness in terms of the expected item responses: $\boldsymbol{S}(\boldsymbol{\alpha}) = \boldsymbol{S}(\boldsymbol{\alpha}') \Rightarrow \boldsymbol{\alpha} = \boldsymbol{\alpha}'$, where $\boldsymbol{S}(\boldsymbol{\alpha}) = E(\boldsymbol{Y} \mid \boldsymbol{\alpha})$ denotes the expectation of an examinee's item-score profile, $\boldsymbol{Y} = (Y_1, \ldots, Y_j, \ldots, Y_J)^T$, given attribute profile $\boldsymbol{\alpha}$. (For the DINA model, the j^{th} entry of $\boldsymbol{S}(\boldsymbol{\alpha})$ is defined as $S_j(\boldsymbol{\alpha}) = E(Y_j \mid \boldsymbol{\alpha}) = (1 - s_j)^{\eta_j} g_j^{(1-\eta_j)}$.)

As a small-scale example, consider the Q-matrix $\mathbf{Q}^*$, with $K = 3$ attributes and $J = 3$ items

$$\mathbf{Q}^* = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

that is not complete, as the computation of the expected item-response profiles, $\boldsymbol{S}(\boldsymbol{\alpha}) = \big(S_1(\boldsymbol{\alpha}), S_2(\boldsymbol{\alpha}), S_3(\boldsymbol{\alpha})\big)^T$, for the DINA model demonstrates

$$S_j(\boldsymbol{\alpha}) = \frac{e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}}{1+e^{\beta_{j0}+\beta_{j12\ldots K^*}\prod_{k=1}^{K^*}\alpha_k}}.$$

Note that due to space restrictions the subsequent table only reports the coefficients that are retained in the expression of $S_j(\boldsymbol{\alpha})$. Observe that in case of the single-attribute item 1, the coefficient of the highest-order interaction term, $\beta_{j12\ldots K^*}$, reduces to the coefficient of the corresponding main effect, β_{12}.

$\boldsymbol{\alpha}$	**Q-Vectors** $\mathbf{q}_1^* = (010)$	$\mathbf{q}_2^* = (101)$	$\mathbf{q}_3^* = (111)$
	$S_1(\boldsymbol{\alpha})$	$S_2(\boldsymbol{\alpha})$	$S_3(\boldsymbol{\alpha})$
(000)	β_{10}	β_{20}	β_{30}
(100)	β_{10}	β_{20}	β_{30}
(010)	$\beta_{10} + \beta_{12}$	β_{20}	β_{30}
(001)	β_{10}	β_{20}	β_{30}
(110)	β_{10}	β_{20}	β_{30}
(101)	β_{10}	$\beta_{20} + \beta_{2(13)}$	β_{30}
(011)	β_{10}	β_{20}	β_{30}
(111)	$\beta_{10} + \beta_{12}$	$\beta_{20} + \beta_{2(13)}$	$\beta_{30} + \beta_{3(123)}$

Clearly, $\mathbf{Q}^*$ is not complete because, for example, $\boldsymbol{\alpha}_1 = (000)^T \neq \boldsymbol{\alpha}_2 = (100)^T \Rightarrow \boldsymbol{S}(\boldsymbol{\alpha}_1) = \boldsymbol{S}(\boldsymbol{\alpha}_2) = (\beta_{10}, \beta_{20}, \beta_{30})$. (Using the DINO model would change the specific mathematical expressions of the $\boldsymbol{S}(\boldsymbol{\alpha})$, but not the general result that the proficiency classes are not well-separated.) An incomplete Q-matrix prevents identification of all possible attribute-profiles among examinees. Thus, completeness of the Q-matrix is a general requirement for any CDM. But Q-completeness can only be determined in connection with a specific CDM supposed to underly the data—a Q-matrix can be shown to be complete for one CDM, but incomplete for another. Hence, the condition under which a Q-matrix is complete must be proven for each CDM individually (for the DINA model, this was proven in Chiu, Douglas, and Li 2009). The duality of the DINA model and the DINO model, however, allows to combine the two proofs of the condition of Q-completeness, as is demonstrated next.

Define $\mathbf{e}_k$ as the K-dimensional unit vector, with the k^{th} entry equal to 1 and all remaining entries equal to 0.

Proposition: *For the DINO model and the DINA model, a $J \times K$ matrix $\mathbf{Q}$ is complete if and only if it contains the K vectors, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_K$, of all single-attribute items among its J rows.*

Proof:
($\Rightarrow$) Consider the attribute profiles $\boldsymbol{\alpha} = (0, 0, \ldots, 0)^T$ and $\boldsymbol{\alpha}' = \mathbf{e}_k$. Thus,

$$\begin{aligned} S_{j(\text{DINA})}(\boldsymbol{\alpha}) &= \frac{e^{\beta_{j0}}}{1 + e^{\beta_{j0}}} \\ S_{j(\text{DINA})}(\boldsymbol{\alpha}') &= \begin{cases} \frac{e^{\beta_{j0}+\beta_{jk}}}{1+e^{\beta_{j0}+\beta_{jk}}} & \text{if } \mathbf{q}_j = \mathbf{e}_k \\ \frac{e^{\beta_{j0}}}{1+e^{\beta_{j0}}} & \text{otherwise.} \end{cases} \end{aligned}$$

Because $\beta_{j0} + \beta_{jk} > \beta_{j0}$, $\frac{e^{\beta_{j0}+\beta_{jk}}}{1+e^{\beta_{j0}+\beta_{jk}}} > \frac{e^{\beta_{j0}}}{1+e^{\beta_{j0}}}$. Assume that $\mathbf{e}_k$ is missing from $\mathbf{Q}$; thus, $S_{j(\text{DINA})}(\boldsymbol{\alpha}) = S_{j(\text{DINA})}(\boldsymbol{\alpha}') = \frac{e^{\beta_{j0}}}{1+e^{\beta_{j0}}}$ for all j, which implies that $\boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}) = \boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}')$. Hence, $\mathbf{Q}$ is not complete.

Define $\boldsymbol{\alpha}^* = \mathbf{1} - \boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'^* = \mathbf{1} - \boldsymbol{\alpha}'$; because $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$, $\boldsymbol{\alpha}^* \neq \boldsymbol{\alpha}'^*$. From the duality of the DINA model and the DINO model then follows:

$$\begin{aligned} \boldsymbol{S}_{\text{DINO}}(\boldsymbol{\alpha}^*) &= E_{\text{DINO}}(\boldsymbol{Y}^* \mid \boldsymbol{\alpha}^*) = E_{\text{DINA}}(\boldsymbol{Y} \mid \boldsymbol{\alpha}) = E_{\text{DINA}}(\boldsymbol{Y} \mid \boldsymbol{\alpha}') \\ &= E_{\text{DINO}}(\boldsymbol{Y}^* \mid \boldsymbol{\alpha}'^*) = \boldsymbol{S}_{\text{DINO}}(\boldsymbol{\alpha}'^*), \end{aligned}$$

(recall that $\boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}) = E_{\text{DINA}}(\boldsymbol{Y} \mid \boldsymbol{\alpha})$). Therefore, $\boldsymbol{Q}$ is not complete.

($\Leftarrow$) Assume that K of the J rows of $\mathbf{Q}$ consist of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_K$. Reorder the rows of $\mathbf{Q}$ by moving these K rows to the first K row positions. Thus,

$$S_{k(\text{DINA})}(\boldsymbol{\alpha}) = \begin{cases} \frac{e^{\beta_{k0}+\beta_{kk}}}{1+e^{\beta_{k0}+\beta_{kk}}} & \text{if } \alpha_k = 1 \\ \frac{e^{\beta_{k0}}}{1+e^{\beta_{k0}}} & \text{if } \alpha_k = 0. \end{cases} \tag{5}$$

Consider $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$ such that $\alpha_k = 1$ and $\alpha'_k = 0$. Then, according to Equation 5, $S_{k(\text{DINA})}(\boldsymbol{\alpha}) > S_{k(\text{DINA})}(\boldsymbol{\alpha}')$ because $\beta_{k0} + \beta_{kk} > \beta_{k0}$, which implies $\boldsymbol{S}_{1:K(\text{DINA})}(\boldsymbol{\alpha}) \neq \boldsymbol{S}_{1:K(\text{DINA})}(\boldsymbol{\alpha}')$, where $\boldsymbol{S}_{1:K(\text{DINA})}(\boldsymbol{\alpha})$ denotes the first K entries in $\boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha})$. Hence, $\boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}) \neq \boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}')$, regardless of whether $\boldsymbol{S}_{(K+1):J(\text{DINA})}(\boldsymbol{\alpha})$ is identical to $\boldsymbol{S}_{(K+1):J(\text{DINA})}(\boldsymbol{\alpha}')$. Therefore, $\mathbf{Q}$ is complete.

Recall that $\boldsymbol{\alpha}^* \neq \boldsymbol{\alpha}'^*$ because $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$. From the duality of the DINA model and the DINO model follows then $\boldsymbol{S}_{\text{DINO}}(\boldsymbol{\alpha}^*) \neq \boldsymbol{S}_{\text{DINO}}(\boldsymbol{\alpha}'^*)$ because $\boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}) \neq \boldsymbol{S}_{\text{DINA}}(\boldsymbol{\alpha}')$. Therefore, $\boldsymbol{S}_{\text{DINO}}(\boldsymbol{\alpha}^*) \neq \boldsymbol{S}_{\text{DINO}}(\boldsymbol{\alpha}'^*)$ if $\boldsymbol{\alpha}^* \neq \boldsymbol{\alpha}'^*$. So, $\boldsymbol{Q}$ must be complete.
■

In summary, the DINA model and the DINO model share the theoretical property that completeness of the Q-matrix is guaranteed if and only if it includes all K single-attribute items.

5. Discussion

The DINA model and the DINO model have always been perceived as two conceptually different models of the relation between attribute mastery and the probability of a correct item response. Liu, Xu, and Ying (2011), however, showed that one model can be expressed in terms of the other. They called this the duality of the DINA model and the DINO model. In this article, the duality of the two models was proven using their parameterization as (constrained) LCDMs: Under certain (linear) transformations, the DINA model and the DINO model are technically identical. As an immediate practical consequence of the DINA-DINO duality, both models can be fitted by the same software. Of course, this appears as a mere curiosity: No one would consider fitting the DINO model by using DINA software with the transformed data, while software for fitting the DINO model is readily available. The importance of the DINA-DINO duality is rather theoretical: If the two models are shown to be identical under certain (linear) transformations, then they must share the same theoretical properties. A methodological development or discovery concerning one model automatically also applies to the other model. Hence, instead of individual proofs for each model, a single proof covers both models. As an example, a compact proof of the condition of Q-completeness for the DINA model and the DINO model was presented that used their duality.

Finally, the DINA-DINO duality also suggests that instead of separate numerical studies, one set of simulations suffices for both models. As a concluding example, consider a study conducted by de la Torre and Lee (2013) comparing the power of the Wald test for several recognizable CDMs including the DINA model and the DINO model. De la Torre and Lee (2013) ran separate simulations for the two models. The results for the power of the Wald test for the DINA model are reported in Table 4 (p. 366, De la Torre and Lee 2013) and for the DINO model in Table 7 (p. 369, De la Torre and Lee 2013). A comparison of the power values shows that they are almost identical for the two models (the few slight deviations are due to sampling variability).

As a concluding remark, it should be noted that the theoretical interest in equivalencies among CDMs is relatively recent. Maris and Bechger (2009) define that two CDMs are formally equivalent if and only if for both models distinct parameterizations can be identified that generate identical item response probabilities. For example, the Reduced Reparameterized Unified Model (Reduced RUM; Hartz 2002; Hartz & Roussos 2008) is equivalent to the Generalized Noisy Input Deterministic Output "And" gate (G-NIDA) model (Maris 1999). In using a log-link function, de la Torre (2011) derived a reparameterization of the Reduced RUM that he called the Additive Cognitive Diagnosis Model (ACDM). Von Davier (2011, 2014) showed that by transforming the Q-matrix the DINA model can be reparameterized as a submodel of the GDM. Maris and Bechger (2009) address the potential damage that can arise from undetected equivalencies among seemingly different CDMs because they may lead to different diagnoses, and subsequently, different treatments. On the other hand, model equivalencies—if they are known—can be exploited for the development of new algorithms for fitting CDMs. Examples are DeCarlo's (2011) Reparameterized DINA model (RDINA) and the reparameterization of the Reduced RUM based on the log-link function that Chiu and Köhn (2015b) used to derive an algorithm for fitting the Reduced RUM with an unlimited number of attributes, which, until then, was technically not possible.

References

CHIU, C.-Y., and KÖHN, H.-F. (2015a), "Consistency of Cluster Analysis for Cognitive Diagnosis: The DINO Model and the DINA Model Revisited," *Applied Psychological Measurement, 39*, 465–479.

CHIU, C.-Y., and KÖHN, H.-F. (2015b), "The Reduced RUM as a Logit Model: Parameterization and Constraints," *Psychometrika*, in press.

CHIU, C.-Y., DOUGLAS, J.A., and LI, X. (2009), "Cluster Analysis for Cognitive Diagnosis: Theory and Applications," *Psychometrika, 74*, 633–665.

DE LA TORRE, J. (2011), "The Generalized DINA Model Framework," *Psychometrika, 76*, 179–199.

DE LA TORRE, J., and LEE, Y.-S., (2013), "Evaluating the Wald Test for Item-Level Comparison of Saturated and Reduced Models in Cognitive Diagnosis," *Journal of Educational Measurement, 50*, 355–373.

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.

DIBELLO, L.V., ROUSSOS, L.A., and STOUT, W. (2007), "Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models," in *Handbook of Statistics: Vol. 26. Psychometrics*, eds. C.R. Rao and S. Sinharay, Amsterdam: North-Holland, pp. 979–1030.

HABERMAN, S.J., and VON DAVIER, M. (2007), "Some Notes on Models for Cognitively Based Skills Diagnosis," in *Handbook of Statistics: Vol. 26. Psychometrics*, eds. C.R. Rao and S. Sinharay, Amsterdam: North-Holland, pp. 1031–1038.

HARTZ, S.M. (2002), "A Bayesian Framework for the Unified Model for Assessing Cognitive Abilities: Blending Theory with Practicality," Doctoral dissertation, available from ProQuest Dissertations and Theses database (UMI No. 3044108).

HARTZ, S.M., and ROUSSOS, L.A. (2008), "The Fusion Model for Skill Diagnosis: Blending Theory with Practicality," Research Report No. RR-08-71, Princeton, NJ: Educational Testing Service.

HENSON, R.A., TEMPLIN, J.L., and WILLSE, J.T. (2009), "Defining a Family of Cognitive Diagnosis Models Using Log-Linear Models with Latent Variables," *Psychometrika, 74*, 191–210.

JUNKER, B.W., and SIJTSMA, K. (2001), "Cognitive Assessment Models with Few Assumptions, and Connections with Nonparametric Item Response Theory," *Applied Psychological Measurement, 25*, 258–272.

LEIGHTON, J., and GIERL, M. (2007), *Cognitive Diagnostic Assessment for Education: Theory and Applications*, Cambridge, UK: Cambridge University Press.

LIU, Y., XU, G., and YING, Z. (2011), "Learning Item-Attribute Relationship in Q-Matrix Based Diagnostic Classification Models," Report: arXiv:1106.0721, New York City, NY: Columbia University, Department of Statistics. Retrieved from the website http://arxiv.org/pdf/1106.0721v1.pdf.

MACREADY, G.B., and DAYTON, C.M. (1977), "The Use of Probabilistic Models in the Assessment of Mastery," *Journal of Educational Statistics, 2*, 99–120.

MARIS, E. (1999), "Estimating Multiple Classification Latent Class Models," *Psychometrika, 64*, 187–212.

MARIS, E., and BECHGER, T. (2008), "Equivalent Diagnostic Classification Models," *Measurement, 7*, 41–46.

PAPADIMITRIOU, C.H., and STEIGLITZ, K. (1998), *Combinatorial Optimization: Algorithms and Complexity* (2nd ed.), Mineola, NY: Dover.

ROBITZSCH, A., KIEFER, T., GEORGE, A.C., and UENLUE, A. (2016), *CDM: Cognitive Diagnosis Modeling. R Package Version 3.1-14*, Retrieved from the Comprehensive `R` Archive Network [CRAN] website http://CRAN.R-project.org/package=CDM.

RUPP, A.A., TEMPLIN, J.L., and HENSON, R.A. (2010), *Diagnostic Measurement. Theory, Methods, and Applications*, New York: Guilford.

TATSUOKA, K.K. (1983), "Rule-Space: An Approach for Dealing with Misconceptions Based on Item-Response Theory," *Journal of Educational Measurement, 20*, 345–354.

TATSUOKA, K.K. (1985), "A Probabilistic Model for Diagnosing Misconception by the Pattern Classification Approach," *Journal of Educational Statistics, 10*, 55–73.

TEMPLIN, J.L., and HENSON, R.A. (2006), "Measurement of Psychological Disorders Using Cognitive Diagnosis Models," *Psychological Methods, 11*, 287–305.

VON DAVIER, M. (2005), "A General Diagnostic Model Applied to Language Testing Data," Research Report No. RR-05-16, Princeton, NJ: Educational Testing Service.

VON DAVIER, M. (2008), "A General Diagnostic Model Applied to Language Testing Data," *British Journal of Mathematical and Statistical Psychology, 61*, 287–301.

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.

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*, 49–71.