*Center for Advanced Studies in Measurement and Assessment*

*CASMA Research Report*

*Number 38*

# Hierarchical Cognitive Diagnostic Analysis: Simulation Study

*Yu-Lan Su, Won-Chan Lee, & Kyong Mi Choi*†

Dec 2013

†Yu-Lan Su is program development associate at ACT (email: yulan.su@act.org). Won-Chan Lee is Associate Professor and Co-director, Center for Advanced Studies in Measurement and Assessment (CASMA), 210 Lindquist Center, College of Education, University of Iowa, Iowa City, IA 52242 (email: won-chan-lee@uiowa.edu). Kyong Mi Choi is Assistant Professor in Mathematics Education, Department of Teaching and Learning, N291 Lindquist Center, University of Iowa (email: kyongmi-choi@-uiowa.edu).

Center for Advanced Studies in
Measurement and Assessment (CASMA)
College of Education
University of Iowa
Iowa City, IA 52242
Tel: 319-335-5439
Web: www.education.uiowa.edu/casma

# Contents

# List of Tables

# List of Figures

# Abstract

This study conducts simulation analyses using two modified cognitive diagnostic models: the deterministic, inputs, noisy, and gate (DINA) model with hierarchical configuration and the deterministic, inputs, noisy, or gate (DINO) model with hierarchical configuration. The simulation study evaluates the effectiveness of incorporating hierarchical structures of the cognitive skills in the model estimation process under various conditions (e.g., various numbers of attributes, test lengths, sample sizes, and hierarchical structures). The simulation study attempts to address the model fits, items fit, and accuracy of item parameter recovery when the skills are in a specified hierarchy and varying estimation models are applied. The simulation analysis examines and compares the impacts of the misspecification of a skill hierarchy on various estimation models under their varying assumptions of dependent or independent attributes. The results of the study demonstrate the proposed approaches with comparing the results of conventional DINA and DINO models to their hierarchical counterparts.

# 1 Introduction

Cognitive Diagnostic Models (CDMs) are psychometric models developed to identify examinees' ability to master fine-grained skills based on a pre-specified matrix. Several CDMs have been applied to parameterize the latent attribute space to model the relationships among attributes and help improve the efficiency in estimating parameters.

In some school subjects, skills are ordered hierarchically. For example, number and operation, algebra, and geometry are not completely independent knowledge segments. The current deterministic, inputs, noisy, and gate (DINA; Haertel, 1989, 1990; Junker & Sijtsma, 2001) model and the deterministic, inputs, noisy, or gate (DINO; Templin & Henson, 2006) model assume independent skills and up to $2^K$ of attribute profiles ($K=$ numbers of attributes), without considering situations where skills are hierarchically related in a certain structure. If the skills are hierarchically related and the conventional models are applied, the parameter estimates are biased and less accurate. There is a need for a model whose specifications, relationships among attributes, possible attribute profiles, and Q-matrix are consistent with the theoretical background and test blueprint.

This study uniquely incorporates hierarchically structured skills into the estimation process of the conventional DINA/DINO models. Additionally, the proposed approaches also reduce the number of possible reasonable latent classes in order to promote computing efficiency. The simulation analysis aims to investigate the inaccuracy of parameter estimates due to misspecification of the relationships among attributes and possible attribute profiles. The simulation study examines model fit and item parameter recovery when the data simulation models are different from the estimation models. Specifically, both the conventional DINA and DINO models and new hierarchical models are applied when skills are independent or dependent in a specified hierarchical structure.

# 2 Methodology

## 2.1 Modified Methods

The detailed discussion about the conventional and hierarchical DINA and DINO models can be found in Su, Choi, Lee, Choi, and McAninch (2013). In addition to the conventional models, Su et al. (2013) also used two modified approaches: The DINA model with hierarchical configurations (DINA-H) and the DINO model with hierarchical configurations (DINO-H). These models involve the hierarchical structures of the cognitive skills in the estimation process. The pre-specified possible attribute profiles under a certain skill hierarchy are adapted in the DINA-H and DINO-H models, whereas the number of possible attribute profiles ($L$) is equal to $2^K$ in the conventional models.

Four common types of cognitive attribute hierarchies are linear, convergent, divergent, and unstructured (Gierl, Leighton, & Hunka, 2007; Leighton, Gierl,

& Hunka, 2004). The skill hierarchies for the condition of six attributes are as discussed in Gierl et al. (2007) and Leighton et al. (2004). Figures 1 to 4 show the attribute hierarchies for the condition of eight attributes. The condition of eight attributes includes two more attributes at the end of the hierarchy than the condition of six attributes. Like the condition of six attributes, the linear attribute hierarchy of eight attributes requires all attributes to be ordered sequentially. In a linear hierarchy, for an examinee to have mastered attribute 8, he or she must have also mastered attributes 1 through 7. The convergent attribute hierarchy specifies a situation in which a single attribute could be the prerequisite to multiple different attributes and situations in which a single attribute could require mastering one or more of the multiple preceding attributes. The relationships among attributes 2 to 5 are the same as specified in the condition of six attributes. In the condition of eight attributes, attribute 5 is the prerequisite to attributes 6 and 7, and mastering either attribute 6 or 7 could lead to mastering attribute 8. The divergent attribute hierarchy refers to different distinct paths originating from the same single attribute. The relationships among attributes 1 and 4 to 6 are the same as specified in the condition of six attributes. In the condition of eight attributes, attributes 7 and 8 appear parallel at the end of attribute 3. That means when an examinee has mastered attributes 7 or 8, he or she has also mastered attributes 1 to 3. The unstructured attribute hierarchy describes cases when a single attribute could be prerequisite to multiple different attributes which have no direct relationships to each other. Similar to the condition of six attributes, attribute 1 is the common prerequisite to attributes 2 to 8 in the condition of eight attributes. However, a certain type of a hierarchy model could have various types of structures, except for the liner hierarchy model. The specified conditions in this study are ones among many possible structures under a hierarchy. This is especially so for the convergent and the divergent hierarchies.

Based on the attribute hierarchies, the number of attribute profiles can be found for each hierarchical model. The attribute profiles for the condition of six attributes are listed in Tables 1 to 4. Tables 5 to 8 list the attribute profiles of each attribute hierarchy for the condition of eight attributes. Table 9 presents the total number of attribute profiles of each attribute hierarchy model for the condition of both six and eight attributes. The number of possible attribute profiles is different for various attribute hierarchies. The more independent the attributes, the larger the number of possible attribute profiles. The higher the dependency among the attributes, the fewer the number of possible attribute profiles. Since the convergent and the divergent hierarchies could have varying structures, the numbers of possible attribute profiles are different for various structures.

## 2.2 Simulation Design

This section includes two parts. The first part describes each factor manipulated in the simulation study, and the second part describes the steps carried out in the simulation process.

### 2.2.1 Simulation Factors

The simulation design considered five factors: the number of attributes, test length, data generating model, sample size, and estimation model. Table 10 displays various levels of these simulation factors.

The first simulation factor was the number of attributes. There were six or eight attributes assessed in each test, which were within the usual range of those found with current applications of CDMs (Rupp & Templin, 2008a). A review of the literature on multiple classification models showed that most application examples used about four to eight attributes (Hartz, 2002; Maris, 1999; Rupp & Templin, 2008b). Six and eight attributes were chosen in the current study because the intention of the study was to maintain consistency with the previous studies using six attributes (Gierl et al., 2007; Leighton et al., 2004; Rupp, Templin, & Henson, 2010), and also to understand the effect of reducing the sample size requirements for tests measuring more attributes.

The second simulation factor was test length (i.e., the number of items). There were 12 or 30 items in each test. These numbers were chosen because the usual range of items measured in the current applications of CDMs is two to four items for every single attribute (i.e., all columns of the Q-matrix sum up to 2 or 4) (Rupp & Templin, 2008a).

The third simulation factor was the identity of the data-generating model. To demonstrate the proposed model and to effectively answer the research questions, the study chose two skill hierarchies to apply in the simulation analysis, which were the linear hierarchy and the unstructured hierarchy. The linear hierarchy is the most constrained hierarchy with the least number of possible attribute profiles. The unstructured hierarchy is the least restricted hierarchy with about half of the number of possible attribute profiles of the conventional DINA/DINO model. Six different models were considered: the DINA model, the DINA model with linear hierarchy (DINA-$\mathrm{H_L}$), the DINA model with unstructured hierarchy (DINA-$\mathrm{H_U}$), the DINO model, the DINO model with linear hierarchy (DINO-$\mathrm{H_L}$), and the DINO model with unstructured hierarchy (DINO-$\mathrm{H_U}$). The baseline DINA and DINO models represented the conditions in which no constraint was imposed on the relationships among attributes.

The fourth simulation factor was sample size. Three levels were used: 300, 1,000, and 3,000 examinees. Simulation studies use a wide range of sample sizes ranging from 500 to 10,000 respondents (e.g., Rupp & Templin, 2008a). The smallest sample size in the current study was chosen to examine the performance of the proposed DINA-H and DINO-H models under small sample-size conditions.

The fifth factor was the identity of the estimation model. Six different models were considered: the DINA, DINA-$\mathrm{H_L}$, DINA-$\mathrm{H_U}$, DINO, DINO-$\mathrm{H_L}$, and DINO-$\mathrm{H_U}$ models. While estimating with each skill hierarchical model, the specified possible attribute profiles of the model was applied to be the examinee's binary skills vector $\alpha_{\mathbf{i}} = \{\alpha_{ik}\}$(where $i$ refers to examinee $i$) the EM algorithem, and to replace the amount of $2^K$ attribute profiles in the conventional DINA and DINO models. In addition, the study only considered the cross comparisons of calibrating data generated from the DINA, DINA-$H_L$, and DINA-$H_U$ models by using DINA-based models, and calibrating data generated from the DINO, DINO-$H_L$, and DINO-$H_U$ models by using the DINO-based models. Cross comparisons between the DINA and DINO models were not considered.

### 2.2.2 Simulation Procedure

The simulation steps were listed below.

1. The study first simulated four item-by-attribute Q-matrices (i.e., 12 items measuring 6 attributes, 12 items measuring 8 attributes, 30 items measuring 6 attributes, and 30 items measuring 8 attributes).

To be closer to practical testing, the distribution of the percentage of items measuring varying numbers of attributes from TIMSS data was used as the guideline in simulating the Q-matrices. The current study adapted the distribution based on the Q-matrix in Park, Lee, and Choi (2010) and Choi (2011). Their Q-matrix was developed by using the attributes form the National Council of Teachers of Mathematics (NCTM) *Principles and Standards for School Mathematics* (NCTM, 2000) for coding, and from the consensus of three mathematics educators. In Park et al. (2010)'s and Choi (2011)'s studies, 39.4% of the coded attributes measured number and operations; 28.2% of the coded attributes measured algebra; 16.9% of the coded attributes measured geometry; 5.6% of the coded attributes measured measurement; 9.9% of the coded attributes measured data analysis and probability. Table 11 shows the Q-matrix from Park et al. (2010) and Choi (2011) for the eighth grade TIMSS 2007 mathematics test. In their Q-matrix, three (10%) of the items measured one attribute, 13 (45%) of the items measured two attributes, ten (35%) of the items measured three attributes, and three (10%) of the items measured four attributes. This Q-matrix consists of 29 items measuring 12 attributes which were coded 71 times in total. Hence, in the current study the distribution of item percentages was set to 30% of items measuring one attribute, 60% of items measuring two attributes, and 10% of items measuring three attributes for the condition of six attributes in the Q-matrix. The distribution of item percentages was set to 10% of items measuring one attribute, 45% of items measuring two attributes, 35% of items measuring three attributes, and 10% of items measuring four attributes for the condition of eight attributes in the Q-matrix. Each row of the Q-matrix sums up to at least 1, which means that each item assesses at least one attribute.

2. The true item parameters of guessing and slip for each corresponding Q-matrix were generated from a random uniform distribution with a lower bound of 0.05 and an upper bound of 0.4 using equations $s_j = P(X_{ij} = 0 \,|\, \eta_{ij} = 1)$ and $g_j = P(X_{ij} = 1 \,|\, \omega_{ij} = 0)$. The true item parameters were generated differently for each unique Q-matrix, but were set the same for various models, different sample sizes, and across replications.

3. Based on the generated Q-matrix and the true guessing and slip parameters, and the possible $\alpha$-matrix for each hierarchy, the $\eta$-matrix was computed for the $N$ examinees (i.e., 300, 1000, or 3000) for the DINA(-H) model (see

Equation 2), and the $\omega$-matrix was computed using Equation 4 for the DINO(-H) model. Once the $\eta$-matrix and $\omega$-matrix were obtained, the probability to answer each item correctly $P(\alpha_i)$ was computed for each examinee. Given the true probabilities and using the cut point of $\alpha = 0.5$ to determine mastery, examinee item responses were simulated randomly from the binomial distribution for each examinee under different simulation conditions. Data were simulated differently for each replication under each condition; however, their corresponding true item parameters were set the same for the 50 replications under each condition to make comparisons easier. Below is the equation of the item response function for the DINA(-H) model:

$$p_j(\alpha_i) = P(X_{ij} = 1 \,|\, \alpha_i) = g_j^{1-\eta_{ij}}(1 - s_j)^{\eta_{ij}}, \tag{1}$$

where $\eta$ refers to a binary indicator showing whether the examinee attribute profile pattern $i$ has mastered all of the required skills for item $j$. The formula for $\eta_{ij}$ is defined as:

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

where $\alpha_{ik}$ refers to the binary mastery status of the $k^{th}$ skill of the $i^{th}$ skill pattern; and, $q_{jk}$ is the Q-matrix entries specifying whether the $j^{th}$ item requires the $k^{th}$ skill. Below is the item response function in the DINO(-H) model:

$$p_j(\omega_{ij}) = P(X_{ij} = 1 \,|\, \omega_{ij}) = g_j^{1-\omega_{ij}}(1 - s_j)^{\omega_{ij}}, \tag{3}$$

where

$$\omega_{ij} = 1 - \prod_{k=1}^{K} (1 - \alpha_{ik})^{q_{jk}}. \tag{4}$$

4. While simulating data for the DINA and DINO models, the number of possible attribute profiles ($L$) was set to $2^K$, and $\alpha$ was a matrix consisting of all the possible combinations of 0s and 1s for all attributes. While simulating data for the DINA-H and DINO-H models, the number of possible attribute profiles ($L$) was set to the pre-specified reasonable examinee attribute profiles from various cognitive skill hierarchies, and $\alpha$ was a matrix consisting of all the pre-specified examinee attribute profiles.

In sum, the 2 (the numbers of attributes: 6 and 8) x 2 (the numbers of items: 12 and 30) x 3 (data generating models: conventional, H-linear, H-unstructured) x 3 (sample sizes: 300, 1000, 3000) x 3 (estimation models: conventional, H-linear, H-unstructured) x 2 (CDMs: DINA and DINO) design resulted in a total of 216 conditions. Each condition was replicated 50 times. Tables 12 and 13 list all the conditions for the DINA/DINA-H and DINO/DINO-H models, respectively. The DINA/DINO-$\mathrm{H_L}$ models refer to the DINA/DINO models with the linear hierarchy, and the DINA/DINO-$\mathrm{H_U}$ models refer to the DINA/DINO models with the unstructured hierarchy.

These simulation studies examined the question of which simulated data best fit the estimating models, based on different skill hierarchies under different conditions. Using simulated data, model parameter recovery was evaluated. Special attention was paid to the robustness of the models, and whether they were able to fit the data that were consistent or inconsistent with the assumptions of the relationship among attributes held by the estimating models.

## 2.3 Analysis

### 2.3.1 Parameter Estimation Process

To estimate examinee attribute profiles and item parameters, the procedure outlined by de la Torre (2009) was followed. The EM algorithm was implemented by first setting initial parameter values, then estimating the expected value of the unknown variables, and giving the current parameter estimates. The final step was to re-estimate the posterior distribution, by maximizing the likelihood of the data by computing the marginal maximum likelihood estimation given the expected estimates of the unknown variables. The estimation steps were repeated until convergence was achieved. Because the estimation process was initialized with a flat prior distribution, the prior was updated in each iteration using the expected proportions of examinees in each latent class (Huebner & Wang, 2011). This method is referred to as the empirical Bayesian method in Carlin and Louis (1996). The estimation began with all the slip and guessing parameters set to 0.2 (Huebner, 2009; Huebner & Wang, 2011). It was shown in a simulation study that the results from the EM estimation were not sensitive to the initial values of the parameters as long as the true guess and slip parameters reasonably fall between 0 and 0.5, and it was demonstrated that the differences in results at varying levels of initial values was negligible (Huebner, 2009). In the DINA and DINA-H models, the $\eta$-matrix is used to indicate whether examinee $i$ has mastered all of the required skills for item $j$, whereas in the DINO and DINO-H models the $\omega$-matrix is used. In the DINA and DINO models, the number of possible attribute profiles $L$ equals $2^K$; whereas $L$ equals the maximum number of possible attribute profiles specified for each unique DINA-H and DINO-H models. In the DINA and DINO models, the initial possible attribute profiles ($\alpha$) contain all $2^K$ possible combinations of 0s and 1s, whereas $\alpha$ contains the possible attribute profiles specified for each unique DINA-H and DINO-H models. The major steps of EM computation described in de la Torre (2009) were outlined as follows, using the notation for the DINA model as an example. The first step in computing the posterior matrix is to calculate the matrix of marginalized maximum likelihood estimation likelihood. For the observed data $X$ and the attribute profiles $\alpha$:

$$Likelihood(X) = \prod_{i=1}^{I} Likelihood(X_i) = \prod_{i=1}^{I} \sum_{l=1}^{L} Likelihood(X_i \mid \alpha_{\mathbf{l}}) p(\alpha_{\mathbf{l}}), \quad (5)$$

where $Likelihood(X)$ is the marginalized likelihood of the response vector of examinee $i$, and $p(\alpha_{\mathbf{l}})$ is the prior probability of the attribute profile vector $\alpha_{\mathbf{l}}$.

The next step toward computing the posterior matrix 1s to multiply the columns of the likelihood matrix by the prior probability for the corresponding skill pattern with a flat (non-informative) prior distribution, meaning that each skill pattern have a probability of $1/L$, where $L$ equals $2^K$ in the conventional DINA and DINO models or $L$ equals the number of all possible attribute profiles specified for each DINA-H or DINO-H model.

Parameter estimation based on the marginalized likelihood (i.e., the marginal maximum likelihood estimation) was implemented using the EM algorithm. To obtain the maximum likelihood estimate, the following is maximized:

$$ln(X) = log \prod_{i=1}^{I} Likelihood(X_i) = \sum_{i=1}^{I} logLikelihood(X_i). \tag{6}$$

The expected number of examinees with attribute profile $\alpha_{\mathbf{l}}$ is computed from

$$I_l = \sum_{i=1}^{I} p(\alpha_{\mathbf{l}} \mid X_i), \tag{7}$$

where $p(\alpha_{\mathbf{l}} \mid X_i)$ is the posterior probability that examinee $i$ has the attribute profile $\alpha_{\mathbf{l}}$ where $l$ ranging form $i$ to $L$ (the number of possible attribute profiles). Moreover, the expected number of examinees with attribute profile $\alpha_{\mathbf{l}}$ answering item $j$ correctly is defined as:

$$R_{jl} = \sum_{i=1}^{I} p(\alpha_{\mathbf{l}} \mid X_i) X_{ij}. \tag{8}$$

Finally, item parameters were estimated when $\eta = 0$, using

$$\hat{g_j} = \frac{R_{jl}^{(0)}}{I_{jl}^{(o)}}, \tag{9}$$

and

$$\hat{s_j} = \frac{[I_{jl}^{(1)} - R_{jl}^{(1)}]}{I_{jl}^{(1)}}, \tag{10}$$

where $I_{jl}^{(0)}$ is the expected number of examinees lacking at least one of the required attributes for item $j$, and $R_{jl}^{(0)}$is the expected number of examinees among $I_{jl}^{(0)}$ correctly answering item $j$. $I_{jl}^{(1)}$ and $R_{jl}^{(1)}$ have the same interpretation except that they pertain to the examinees with all the required attributes for item $j$. $I_{jl}^{(0)} + I_{jl}^{(1)}$ is equal to $I_l$ for all $j$.

As mentioned earlier in this section, the algorithm started with initial values for $g$ and $s$ both equal to 0.2. Next, $I_{jl}^{(0)}$, $R_{jl}^{(0)}$, $I_{jl}^{(1)}$ and $R_{jl}^{(0)}$ were computed based on the current values of $g$ and $s$. Then, the values of $g$ and $s$ were found and updated. The steps were repeated until convergence was achieved. The criterion for convergence was set to be smaller than 0.001. The number of iteration cycles was smaller than 100 across all the conditions of the simulation.

### 2.3.2 Evaluation Indices

**Fit Indices**. Model fit evaluations were to address the question of whether the data simulated based on different hierarchical models under varying conditions fit the estimating models. Again, special attention was paid to the robustness of the model to fit the simulated data that were not consistent with the estimating model. The mean of AIC and BIC for each condition are reported in the result section in which the terms MAIC and MBIC are used. AIC is defined as:

$$AIC = -2ln(Likelihood) + 2p, \tag{11}$$

where ln(Likelihood) is the log-likelihood of the data under the model and $p$ is the number of parameters in the model. For the conventional DINA and DINO models, $p = 2J + 2^K - 1$. For the DINA-H and DINO-H models, $p = 2J + L - 1$ where $L$ is equal to the maximum number of possible attribute profiles specified for each unique model. For a given dataset, the larger the log-likelihood, the better the model fit; the smaller the AIC value, the better the model fit (Xu & von Davier, 2008). BIC is defined as:

$$BIC = -2ln(Likelihood) + p\,ln(N), \tag{12}$$

where $N$ is the sample size. Again, the smaller the BIC value, the better the model fit. The AIC and BIC for each condition are reported in the results section.

**Summary Statistics**. Three bias indices were used to evaluate model parameter recovery: the average squared bias (ASB), the average variance (AVAR), and the average mean squared error (AMSE) of item parameter estimates ($s$ and $g$) for each condition. Under each simulation condition, the bias for each item was defined as the difference between the average item parameter estimates over replications and their corresponding true generating values. The following formulas use the guessing parameter as an example.

$$Bias_j = \bar{g}_j - g_j^* = (\frac{1}{R}\sum_{r=1}^{R} g_{jr}) - g_j^*, \tag{13}$$

where $\bar{g}_j$ is the average $g$ parameter estimate for item $j$ over replications; $g_j^*$ is the true guessing parameter for item $j$; $g_{jr}$ is the guessing parameter estimate for item $j$ for replication $r$; $R$ refers to the number of replications under each condition. The ASB for each condition is defined as the average squared bias:

$$ASB(g) = \frac{1}{J}\sum_{j=1}^{J} Bias_j^2 = \frac{1}{J}\sum_{j=1}^{J}(\bar{g}_j - g_j^*)^2. \tag{14}$$

Furthermore, AVAR is defined as the average variance of an item parameter across replications for each condition, which is given by:

$$AVAR(g) = \frac{1}{J}\sum_{j=1}^{J}\frac{(g_{jr} - \bar{g}_j)^2}{R}. \tag{15}$$

Finally, since the mean squared error is equal to the squared bias plus variance, the AMSE is regarded as a combination of information from variance and bias, and is defined as:

$$AMSE(g) = \frac{1}{J}\sum_{j=1}^{J}\sum_{r=1}^{R}\frac{(g_{jr} - g_j^*)^2}{R}. \tag{16}$$

# 3 Results

This section describes the detailed results of the fit indices and summary statistics for varying conditions under the DINA(O) and DINA(O)-H models from the simulation study. The conditions in the tables refer to the data generating models and the estimation models, respectively. For example, the condition of HL_NA refers to the condition of using the DINA-H-linear data generating model and the conventional DINA estimation model. The same logic of notation naming is used throughout the whole paper.

## 3.1 DINA and DINA-H

This part shows the results of the model fit and summary statistics from the DINA and DINA-H models in the simulation analysis. It includes the main effects of model consistency, numbers of attribute, test length, and sample size; the interaction effects of test length by attribute, sample size by attribute, and sample size by test length; and the three-way interaction effect of sample size by test length by attribute.

### 3.1.1 Main Effect of Model Consistency

The following paragraphs present the results of the fit indices and summary statistics for the main effect of model consistency testing varying estimation models, in terms of their being consistent or inconsistent with the specifications on the skill hierarchies, under the DINA and DINA-H models from the simulation study. The values used for comparisons were computed by averaging over other conditions of sample size, test length, and number of attribute.

**Fit Indices**. For the model fit indices, MAIC and MBIC, all the conditions that used the estimation models consistent with their data generating models show better model fit results (i.e., smaller MAICs and smaller MBICs) than the conditions that used the estimation models inconsistent with their data generating models, as shown in Table 14. The results confirm the assumption that a better model fit can be obtained from the calibration results when the relationships among attributes specified in the data generating model are consistent with those in the estimation model. The results also confirm that when skills are ordered hierarchically, the proposed DINA-H models perform better.

**Summary Statistics**. The results of the summary statistics for the main effect of model consistency show that using the consistent estimation model with the data generating model obtained better item parameter recovery (i.e., the smaller ASB and AMSE of guessing and slip parameter estimates) when compared to the results of using the inconsistent model, as shown in Table 15. Similar to the results from the evaluation of the fit indices, the results of the summary statistics confirm the assumption that better item parameter recovery can be obtained from the calibration results when the relationships among attributes specified in the data generating model are consistent with those in the estimation model, and also confirm that the proposed DINA-H models should be used to calibrate items when cognitive skills are ordered hierarchically.

### 3.1.2 Main Effect of Number of Attribute

Since the main effect of model consistency was confirmed and supported, the following main effects of other study variables and their interaction effects were examined by comparing only the results of using the consistent models in generating and estimating data. Note also that magnitudes of the model fit indices (i.e., MAIC and MBIC) depend heavily on the numbers of items, attributes, and examinees, and thus the comparison of the fit indices across various conditions does not provide any meaningful information. The MAIC and MBIC values are not compared for the subsequent main and interaction effects. This section presents the results of the summary statistics for the main effect of number of attribute under the DINA and DINA-H models from the simulation study. When interpating the values, focus on the boldfaced mean values.

The results of the summary statistics for the main effect of number of attribute show that the conditions of eight attributes obtain better item parameter recovery with smaller ASB, AVAR, and AMSE when compared to the results of six attributes (see Table 16). The only exception is that AVAR of slip parameter estimates is smaller for the condition of eight attributes. For both conditions of six and eight attributes, the guessing parameter estimates show smaller ASB, but larger AMSE than the slip parameter estimates. Increasing the number of attribute from six to eight improves the recovery of both parameter estimates, with the amount of ASB, AVAR, and AMSE decreasing. This is especially so when attributes are more related to each other.

### 3.1.3 Main Effect of Test Length

The results of the summary statistics for the main effect of test length show that the condition of 12 items in a test has smaller ASB and AMSE for both the guessing and slip parameter estimates when compared to the results of 30 items, as shown in Table 17. In general, the results of the summary statistics confirm that tests that consisted of 12 items obtained better item parameter recovery than those that used 30 items although there is larger variance when using 12 items in a test. When reducing the number of items from 30 to 12 items, more reduction in estimation error (i.e., AMSE and ASB) is found for the slip parameter estimates than for the guessing parameter estimates. For both the conditions of 12 and 30 items, guessing parameter estimates are more accurate than slip parameter estimates.

### 3.1.4 Main Effect of Sample Size

The results of the summary statistics for the main effect of sample size consistently show that using N=3000 has the least average ASB, AVAR, and AMSE for both the slip and guessing parameter estimates, compared to the results of using N=300 or 1000 (see Table 18). It means that using the larger sample size would result in better item parameter recovery. The results of the summary statistics support the assumption that better item parameter recovery can be obtained from the calibration results when large sample sizes are applied. The magnitude of AVAR consistently decreases when the sample size increases for both the guessing and slip parameter estimates. For all conditions of sample size, the DINA-$H_L$ condition shows better item parameter recovery results of smaller AMSE than the DINA and DINA-$H_U$ models. This may be due to the fact that the linear model has the highest level of dependency of attributes, comparing to the other models.

### 3.1.5 Interaction Effect of Test Length by Number of Attribute

The results of the summary statistics for the interaction effect of different test lengths by varying numbers of attributes show that the condition of 30 items with eight attributes performed the best (see Table 19) although the results of the main effects show that the condition of 12 items and the condition of eight attributes tend to perform better than other conditions (see Tables 16 and 17). Using 30 items measuring eight attributes show the smallest mean ASB, AVAR, and AMSE for both guessing and slip parameter estimates. As shown in Figure 5, the interaction effect between test length and number of attribute can be observed under the DINA(-H) model.

Comparing to the NA_NA and HU_HU conditions, the condition of HL_HL obtains the best item parameter recovery results, except for the larger AMSE and ASB under the condition of 30 items with six attributes for guessing. Both AMSE and ASB increase when the number of item increases under the condition of six attributes, whereas the error decreases when the number of item increases under the condition of eight attributes for both the guessing and slip parameter estimates. Hence, the interaction effect is found for 30 items with eight attributes performing better. All summary statistics results show that the guessing parameter recovery is better than the slip parameter recovery.

### 3.1.6 Interaction Effect of Sample Size by Number of Attribute

Table 20 and Figure 6 present the ASB, AVAR, and AMSE of guessing and slip parameter estimates for the interaction effect of N by K for the DINA and DINA-H models. The results of summary statistics for the guessing and slip parameter estimates are different. The smallest mean AMSE was shown in the condition of N=3000 and K=8 for the guessing parameter estimates. The smallest mean ASB was shown in the condition of N=1000 and K=8 for the guessing parameter estimates although there are just slightly different from the conditions of N=3000 and K=8 and N=3000 and K=6. For the slip parameter estimates, the best item parameter recovery was observed in the condition of N=3000 and K=6. For the main effect, the condition of eight attributes and the condition of N=3000 perform better. In terms of interaction effect, the condition of N=3000 and K=8 provides the best results for the guessing parameter estimates, and the condition of N=3000 and K=6 provides the best results for the slip parameter estimates.

In terms of AMSE for the guessing parameter estimates, the condition of N=3000 and K=8 performs the best, next the condition of N=3000 and K=6 followed by the condition of N=1000 and K=8. For the slip parameter estimates, the top three conditions are the condition of N=3000 and K=6, the condition of N=3000 and K=8, and the condition of N=1000 and K=6. The item parameter recovery results are better for the guessing parameters than for the slip parameter estimates.

Comparing the conditions of different estimation models, for the guessing parameter estimates the DINA-$H_L$ model shows better item parameter recovery of the smaller AMSE than the DINA and DINA-$H_U$ models. When sample sizes are larger, the three models perform similarly for their guessing parameter estimates. For the slip parameter estimates, the DINA-$H_L$ model performs the best, and the conventional DINA model obtains the poorest item parameter recovery results. Consistently across all conditions, the guessing parameter results are better than the slip parameter results.

The sample size by number of attribute interaction effect is more obvious when the sample size is smaller (see Figure 6). It again reconfirms the sample size concern while conducting CDM calibrations.

#### 3.1.7 Interaction Effect of Sample Size by Test Length

Table 21 lists the ASB, AVAR, and AMSE of guessing and slip parameter estimates for the interaction effect of sample size by test length for the DINA and DINA-H models. For both parameter estimates, the condition of N=3000 and J=30 obtains the best item parameter recovery results and the next better results are shown under the condition of N=1000 and J=30 and the condition of N=3000 and J=12, while the condition of N=300 and J=30 obtains the worst item parameter recovery results. The sample size by test length interaction effect can also be observed in Figure 7, especially for the condition of smaller sample size. For the main effects, the condition of N=3000 and the condition of J=12 produce the best item parameter recovery results (see Tables 17 and 18).

The mean AMSE, ASB, and AVAR decrease when sample size increases for both guessing and slip parameter estimates (see Table 21). When the sample size is larger (i.e.,1000 or 3000), the conditions with more items produce more accurate item parameter recovery results for both guessing and slip parameter estimates (see Figure 7). However, when the sample size is smaller, the conditions with more items produce less accurate item parameter recovery results for both parameter estimates. The guessing parameters are recovered better than the slip parameters with smaller error across all conditions.

The DINA-$H_L$ model shows the best item parameter recovery results and the DINA-$H_U$ model shows better item parameter recovery than the conventional DINA model under all conditions for the slip parameter estimates although this is not necessary true fot the guessing parameter estimates (see Table 21).

#### 3.1.8 Interaction Effect of N by J by K

Table 22 lists the ASB, AVAR, and AMSE of guessing and slip parameter estimates for the N by J by K three-way interaction effect under the DINA and DINA-H models. These three-way interaction tables also show a comprehensive picture of which combinations of various conditions provide the best results. The values shown in the tables are averaging over the values with consistent data-generating and estimation models for each N by J by K condition.

For the main effects of both the guessing and slip parameter estimates, best results were obtained from the conditions of N=3000, J=12, and K=8 (see Tables 16 to 18). For the three-way interaction effect for the guessing parameter estimates, across all samples size conditions, the condition of N=3000, J=30 and K=8 obtained the best item parameter recovery with the smallest ASB, AVAR, and AMSE, followed by the condition of N=3000 by J=30 by K=6, and then the condition of N=1000 by J=30 by K=8 (see Table 22 ). For the slip parameter estimates, the condition of N=3000 by J=30 by K=6 obtained the best item parameter recovery with the smallest ASB, AVAR, and AMSE. The next better results are shown under the condition of N=3000 by J=30 by K=8 and then the condition of N=3000 by J=12 by K=6 (see Table 22). Generally speaking, the guessing parameter estimates show better item parameter recovery results than the slip parameter estimates, with a few exceptions of smaller ASB and AVAR appearing for the slip parameter estimates, but not the AMSE values. The N by J by K three-way interaction was shown when sample size is smaller.

### 3.2 DINO and DINO-H

#### 3.2.1 Main Effect of Model Consistency

This part discusses the results of the fit indices and summary statistics for the main effect of model consistency when using varying estimation models consistent or inconsistent with the specifications of the skill hierarchies under the DINO and DINO-H models from the simulation study.

**Fit Indices**. For the model fit indices MAIC and MBIC, all conditions, except one, that used the estimation models consistent with their data generating models show better model fit results than the conditions that used the estimation models inconsistent with their data generating models (see Table 23). The only exception is the smaller MBIC that occurs when using the DINO-$H_L$ model to estimate data generated via the DINO-$H_U$ model. Generally speaking, the results confirm the main effect of model consistency that better model fit can be obtained from the calibration results when the relationships among attributes specified in the data generating model are consistent with those in the estimation model. Similar to what was found under the DINA-H model, the results also confirm that when skills are ordered hierarchically, the proposed DINO-H model performs better and is preferred to be applied, in terms of the model fit.

Due to the different levels of dependency among the hierarchically ordered skills, using the conventional DINO model to estimate data generated via the DINO-$H_L$ model produces poorer model fit results than those estimated by the DINO-$H_U$ model. However, the results do not show any consistent pattern for the data generated via the conventional DINO or the DINO-$H_U$ models.

**Summary Statistics**. The results of the summary statistics for the main effect of model consistency show that using the consistent estimation model with the data generating model produces more accurate results of guessing and slip parameter estimates with the least AMSE and ASB, compared to the results of using the inconsistent model (see Table 24). The two exceptions are that the smaller AVAR values appear when using the DINO model to estimate the DINO-$H_L$ model data for the guessing parameter estimates and when using the DINO-$H_L$ model to estimate the DINO-$H_U$ model data for the slip parameter estimates. Similar to the results from the evaluation of the fit indices, when the relationships among attributes specified in the data generating model are consistent with those in the estimation model, better item parameter recovery results can be obtained. The results of the summary statistics under the DINO-H models are consistent to what were found under the DINA-H models. The proposed DINO-H models should be applied to calibrate items when cognitive skills are in a certain hierarchical structure.

### 3.2.2 Main Effect of Number of Attribute

As with the DINA(-H) model, the results of model fit are not comparable here because the values of MAIC and MBIC are more sensitive to the number of attribute, test length, and sample size. Hence, only the summary statistics are evaluated for the following main and interaction effects.

The results of the summary statistics for the main effect of varying numbers of attributes show that measuring six attributes in a Q-matrix generates better item parameter recovery when compared to the results of measuring eight attributes for the guessing parameter estimates, as shown in Table 25. For the slip parameter estimates under the DINO(-H) model, the condition of eight attributes performs better in terms of item parameter recovery than the condition of six attributes. The results of the summary statistics show that the main effect of number of attribute is different for the guessing and slip parameter estimates. The reduction of the estimation error in the guessing parameter estimates when decreasing the number of attribute from eight to six is much larger than the improvement in the slip parameter estimates when increasing the number of attribute from six to eight.

For both conditions of six and eight attributes, the condition of HL_HL shows the least amount of error for both guessing and slip parameter estimates.

The condition of NO_NO shows more accurate estimation results with smaller AMSE and AVAR than the condition of HU_HU for the guessing parameter estimates although the smaller ASB appears under the condition of HU_HU. In contrast, for the slip parameter estimates the condition of HU_HU produces better item parameter recovery results than the condition of NO_NO, with only one exception of smaller ASB appearing for the condition of NO_NO when K=6.

#### 3.2.3 Main Effect of Test Length

The results of the summary statistics for the main effect of varying test lengths show that the condition of 30 items consistently shows better item parameter recovery than the condition of 12 items (see Table 26). The condition of 30 items produces smaller AMSE, ASB, and AVAR than the condition of 12 items for both the guessing and slip parameter estimates. The improvement of estimation accuracy is better for the guessing parameter estimates than for the slip parameter estimates when increasing number of items from 12 to 30.

The slip parameter estimates show better item parameter recovery results than the guessing parameter estimates. For both parameter estimates under both the conditions of 12 and 30 items, the DINO-$H_L$ model consistently performs better than the DINO and DINO-$H_U$ models. The only exception is that the smaller ASB appears under the conventional DINO model rather than under the DINO-$H_L$ model for the slip parameter estimates when J=30. For the guessing parameter estimates, the NO_NO condition shows better item parameter recovery than the condition of HU_HU for both conditions of 12 and 30 items although the larger ASB is found under the NO_NO condition when J=12. For the slip parameter estimates, the HU_HU condition is better when J=12 and the NO_NO condition is better when J=30.

#### 3.2.4 Main Effect of Sample Size

All the results of the summary statistics for the main effect of varying sample sizes show that the condition of the largest sample size (i.e., N=3000) produced the smallest ASB, AVAR, and AMSE for both the guessing and slip parameter estimates, as shown in Table 27. The values of ASB, AVAR, and AMSE for both the guessing and slip parameter estimates decrease when sample size increases. The slip parameter estimates consistently perform better than the guessing parameter estimates for all three conditions of sample sizes. In terms of the mean values of AMSE, ASB, and AVAR, the reduction of the estimation error when increasing sample size from 300 to 1000 and from 1000 to 3000 for the guessing parameter estimates is much larger than the improvement for the slip parameter estimates. Across various sample sizes, the condition of DINO-$H_L$ model shows less estimation error than the DINO and DINO-$H_U$ models.

#### 3.2.5 Interaction Effect of Test Length by Number of Attribute

The results of the summary statistics for the interaction effect of test length with varying numbers of attributes show that using more items measuring fewer attributes obtained better item parameter recovery, compared to the results of using fewer items measuring more attributes. As shown in Table 28, the condition of 30 items with 6 attributes produces the smallest ASB, AVAR, and AMSE for both the guessing and slip parameter estimates. The next best condition is 12 items with six attributes for the guessing parameter estimates, whereas it is the condition of 30 items with eight attributes for the slip parameter estimates. The main effect results show that the conditions of 30 items and six attributes are preferred for estimating the guessing parameters, while the conditions of 30 items and eight attributes are preferred for estimating the slip parameters (see Table 25). The interaction effect can be observed in Figure 8 showing that the conditions of more items with fewer attributes produce better item parameter recovery than the conditions of fewer items with more attributes.

Generally speaking, the slip parameter estimates show better item parameter recovery results than the guessing parameter estimates, except for the smaller ASB for the guessing parameter estimates when J=12 and K=6. For the guessing parameter estimates, the condition of DINO-$\mathrm{H_L}$ model produced less estimation error than the conditions of DINO and DINO-$\mathrm{H_U}$ models across all combinations of conditions of different test lengths by varying numbers of attributes. For the slip parameter estimates, the condition of DINO-$\mathrm{H_L}$ model is better than the DINO and DINO-$\mathrm{H_U}$ models under the conditions of J=12 with K=6, J=12 with K=8, and J=30 with K=8. However, the DINO model shows better slip parameter recovery results under the condition of J=30 with K=6.

### 3.2.6 Interaction Effect of Sample Size by Number of Attribute

The ASB, AVAR, and AMSE of guessing and slip parameter estimates for the interaction effect of N by K for the DINO and DINO-H models are presented in Table 29, respectively. The results of summary statistics for the guessing parameter estimates show that the condition of N=3000 with K=6 obtained the best item parameter recovery results with the least mean AMSE and AVAR, although the condition of N=1000 with K=6 produced the least mean ASB. The difference of mean ASB between the conditions of N=1000 with K=6 and N=3000 with K=6 is relatively small (i.e., 0.00002). The results of summary statistics for the slip parameter estimates show that the condition of N=3000 with K=8 obtained the least mean AVAR and AMSE, and the condition of N=1000 with K=8 produced the least mean ASB. The difference of mean ASB between the conditions of N=1000 with K=8 and N=3000 with K=6 is 0.000014.

For the main effect, the condition of six attributes performs better than eight attributes in estimating the guessing parameters, but the condition of eight attributes provides better results for the slip parameter estimates (see Table 25), and the condition of N=3000 provides the best results for both the guessing and slip parameter estimates (see Table 27). Generally speaking, the interaction effect shows that the condition of larger sample sizes with fewer attributes in the Q-matrix show better item parameter recovery for the guessing parameter estimates, but the condition of larger sample size with more attributes is better for the slip parameter estimates, as shown in Figure 9.

The slip parameters are, in general, recovered better than the guessing parameters, except for the ASB values under the conditions of N=300 with K=6 and N=1000 with K=6. In terms of AMSE for the guessing parameter estimates, the DINO-$H_L$ model shows better estimation accuracy than the DINO and DINO-$H_U$ models. The only exception is that the smaller AMSE occurs when N=1000 with K=6. For the slip parameter estimates, the DINO-$H_L$ model consistently shows better estimation accuracy than the DINO and DINO-$H_U$ models across all different conditions of sample sizes with varying numbers of attributes.

#### 3.2.7 Interaction Effect of Sample Size by Test Length

Table 30 lists the ASB, AVAR, and AMSE of guessing and slip parameter estimates for the interaction effect of sample size by test length for the DINO and DINO-H models, respectively. For both guessing and slip parameter estimates, the conditions of N=3000 with J=30 produced the least mean ASB, AVAR, and AMSE across all conditions. In terms of AMSE, the next best conditions are N=3000 with J=12 and N=1000 with J=30 for both guessing and slip parameter estimates. It shows that having larger sample sizes with more items in a test can obtain more accurate item parameter recovery results. The role of sample size is more important than the role of number of items in reducing estimation error. The values of ASB, AVAR, and AMSE increase when sample size decreases. The values of ASB, AVAR, and AMSE increase when test length decreases, except for the condition of N=300 for the guessing parameter estimates. The trend of the main effects of sample size and test length is clearer than the interaction effect between the two variables (see Figure 10), as the main effects show that the condition of N=3000 and the condition of J=30 are preferred (see Tables 26 and 27).

The slip parameter estimates show better item parameter recovery results than the guessing parameter estimates across all conditions of different sample sizes by various test lengths. The DINO-$H_L$ model produces more accurate estimation results than using the DINO and DINO-$H_U$ models.

#### 3.2.8 Interaction Effect of N by J by K

Table 31 presents the results of summary statistics for the N by J by K three-way interaction effect for the guessing and slip parameter estimates under the DINO and DINO-H models, respectively. The values were averaging over the three values with consistent data-generating and estimation models for each N by J by K condition.

For the main effects under the DINO(-H) model, the conditions of N=3000, J=30, and K=6 show better item parameter recovery for the guessing parameter estimates, and the conditions of N=3000, J=30, and K=8 show better item parameter recovery for the slip parameter estimates (see Tables 25 to 27). For the three-way interaction effect for the guessing parameter estimates, the condition of N=3000 by J=30 by K=6 produced the best item parameter recovery results with the least mean AMSE, ASB, and AVAR. For the slip parameter estimates, the condition of N=3000 by J=30 by K=8 shows the best item parameter recovery results with the least mean AMSE and AVAR, while the condition of N=3000 by J=30 by K=6 has the least mean ASB. With smaller sample sizes, the values of summary statistics are more sensitive to the number of attribute and test length. When sample size is smaller, test length is shorter and the number of attribute is large, item parameter recovery is poorer with larger ASB and AMSE for the slip parameter estimates. In general, the slip parameter estimates show more accurate estimation results than the guessing parameter estimates, except for the conditions of N=1000 by J=12 by K=6 and N=300 by J=12 by K=6 in which the guessing parameter estimates performed better.

## 3.3 DINA(-H) vs. DINO(-H)

### 3.3.1 Main Effect of Model Consistency

This part contrasts the results of the fit indices and summary statistics between the DINA(-H) and DINO(-H) models from the simulation study for the main effect of using varying estimation models consistent or inconsistent with the specifications on the skill hierarchies.

**Fit Indices**. Generally speaking, for the main effect of model consistency, both the DINA(-H) and DINO(-H) models confirm that using the estimation models consistent with the data generating models show better model fit results than the conditions using the estimation models inconsistent with their data generating models, as shown in Tables 14 and 23.

Table 32 shows the differences of model fit between the DINA(-H) and DINO(-H) models. The differences found in MAIC and MBIC between the DINA(-H) and DINO(-H) models were computed by subtracting the DINA(-H) condition values from those of the DINO(-H). The negative values in the differences of MAIC and MBIC, thus, indicate that the DINO(-H) model performs better than the DINA(-H) model. All the differences of the model fit indices show that the DINO(-H) model has better model fit with smaller MAIC and MBIC than those in the DINA(-H) model. In terms of model fit, the DINO(-H) model performs better.

**Summary Statistics**. All the results of summary statistics for the DINA(-H) and DINO(-H) models' main effect of model consistency show that using the estimation model consistent with the data generating model obtains better item parameter recovery (see Tables 15 and 24).

The differences of ASB, AVAR, and AMSE for the guessing and slip parameters between the DINA(-H) and DINO(-H) models, as shown in Table 33, were also computed from using the values in the DINO(-H) conditions minus those in the DINA(-H) conditions. The negative values mean that the values in the DINO(-H) conditions are smaller than the values in the DINA(-H) conditions, indicating that the DINO(-H) model performs better than the DINA(-H) model. Comparing the summary statistics between the DINA(-H) model and the DINO(-H) model, the DINA(-H) model outperforms the DINO(-H) model with smaller AMSE under two-thirds of the conditions for the guessing parameter estimates. The DINO(-H) model performs better than the DINA(-H) model with smaller AMSE values under all conditions for the slip parameter estimates. This finding generally reconfirms that the DINO(-H) model obtains better item parameter recovery than the DINA(-H) model although not for tte AMSE for guessing.

### 3.3.2 Main Effect of Number of Attribute

The condition of eight attributes produces more accurate item parameter recovery results for both the guessing and slip parameter estimates under the DINA(-H) model and for the slip parameter estimates under the DINO(-H) model, as shown in Tables 16 and 25, while the condition of six attribute is deemed to be a better condition in terms of the guessing parameter estimation. Comparing the differences of AMSE between the DINA(-H) and DINO(-H) models, the DINO(-H) model displays better item parameter recovery than the DINA(-H) model in estimating the slip parameters for most of the condition of six and eight attributes (See Table 34). The DINA(-H) model outperforms the DINO(-H) model for the guessing parameter estimates when K=8.

### 3.3.3 Main Effect of Test Length

The results of the summary statistics for the main effect of test length show that using 12 items in a test obtains better item parameter recovery under the DINA(-H) model and using 30 items in a test obtains better item parameter recovery under the DINO(-H) model (see Tables 17 and 26).

Generally speaking, the DINO model outperforms the DINA model in estimating slip parameters for both conditions of J=12 and J=30, whereas the DINA model outperforms the DINO model in obtaining better guessing parameter recovery with smaller AMSE when J=12 and J=12 (see Table 35).

### 3.3.4 Main Effect of Sample Size

All the results of the summary statistics for the main effect of sample size under both the DINA(-H) and DINO(-H) models show that using the larger sample size (i.e., N=3000) would result in better item parameter recovery for both the guessing and slip parameter estimates (see Tables 18 and 27).

Some mixed findings appear in the differences of item parameter recovery between the DINA(-H) and DINO(-H) models. For the guessing parameter recovery, the DINA model performs better in most of the conditions, whereas the DINO model outperforms the DINA model in obtaining the better slip parameter recovery, regardless of the sample size (see Table 36).

## 3.4 Summary of the Results

The main effect of model consistency from the simulation analysis shows that better results can be obtained when the relationships among attributes specified in the data generating model are consistent with those in the estimation model. Hence, when skills are ordered hierarchically, the proposed DINA-H and DINO-H models should be considered, rather than the conventional DINA and DINO models.

For the main effect of the number of attribute, the condition of eight attributes recovers item parameter estimates more accurately under both the DINA(-H) and DINO(-H) models, except for the guessing parameter estimates under the DINO(-H) model. For the main effect of test length, using 12 items in a test obtains better results for the DINA model, and using 30 items is better for the DINO model. For the main effect of sample size, using the larger sample size results in better item parameter recovery for all models. Generally speaking, having a larger sample size with more items in a test can obtain more accurate item parameter recovery results. The role of sample size is more important than the role of the number of items in reducing estimation error. With smaller sample sizes, the values of summary statistics are more sensitive to the number of attribute and test length. Using more items measuring fewer attributes displayed more accurate item parameter recovery under the DINO(-H) models. The conditions of larger sample sizes with fewer attributes in the Q-matrix show more accurate item parameter recovery results for the slip parameter estimates under the DINA(-H) model and for both parameter estimates under the DINO(-H) model.

### 3.4.1 Limitations

Due to the scope of the study, only a few variables and a limited number of conditions for each simulation factor were considered in the simulation design. Future research could incorporate other factors that were not considered in the current study. The comparisons between the DINA/DINA-H and the DINO/DINO-H models are only based on numeric data. In practical situations, it is crucial to consider the rationality of different assumptions of these two models. Additionally, other confounding factors such as the dependency among skills or the response patterns, dimensionalities and item difficulties, are not completely excluded in the analysis.

# 4 Conclusions

The calibration results of simulation analysis using different estimation models consistent or inconsistent with the specifications of the skill hierarchies were compared to each other. The item parameters are poorly recovered when a specified skill hierarchy is inconsistent with an estimation model. This situation is especially worse with smaller sample sizes, and fewer items with more attributes in the Q-matrices. The misspecification of a skill hierarchy has a negative impact on all models across most of the conditions of number of attribute, test length, and sample size. The DINA-H and the DINO-H models provide stable item parameter estimates even with smaller sample sizes. The results support the assumption that the DINA-H and the DINO-H models, instead of the conventional DINA and DINO models, should be considered when skills are in a certain hierarchical structure, no matter how long the test is, how many attributes are measured, or how small the sample size is.

The simulation study confirms that model specification needs to be consistent with the assumptions and characteristics of skills in order to obtain better model fit, and item parameter recovery. The current study contributes to the examination of the performance of the proposed DINA-H and DINO-H models, and provides information about model fit and item parameter recovery under varying conditions of number of attribute, test length, and sample size. The results of the study demonstrate the feasibility of the proposed DINA-H and DINO-H models, facilitate the reduction of possible attribute profiles in analyzing a CDM, allow analysis of tests that assess more attributes in the future, and promote computational efficiency.

REFERENCES

Carlin, B., & Louis, A. (1996). *Bayes and Empirical Bayes Methods for Data Analysis.* New York: Chapman & Hall.

Choi, K. M. (2011). *What make changes in US 8th graders' mathematics literacy? Using TIMSS 2003 and 2007 via cognitive diagnostic modeling.* Unpublished manuscript. The University of Iowa, IA.

de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics, 34*, 115-130. doi: 10.3102/1076998607309474

Gierl, M. J., Leighton, J. P., & Hunka, S. (2007). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 242–274). Cambridge, UK: Cambridge University Press.

Haertel, E. H. (1989). Using restricted latent class models to map skill structure of achievement items. *Journal of Educational Measurement, 26,* 301–321. doi:10.1111/j.1745-3984.1989.tb00336.x

Haertel, E. H. (1990). Continuous and discrete latent structure models of item response data. *Psychometrika, 55,* 477–494. doi:10.1007/BF02294762

Hartz, S. M. (2002). *A Bayesian framework for the Unified Model for assessing cognitive abilities: Blending theory with practicality.* Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign, IL.

Huebner, A. (2009). *Implementation of the DINA Cognitive Diagnostic Model.* Unpublished manuscript. Iowa city, IA: ACT, Inc.

Huebner, A., & Wang, C. (2011). A Note on Comparing Examinee Classification Methods for Cognitive Diagnosis Models. *Educational and Psychological Measurement, 71*, 407-419. doi: 10.1177/0013164410388832

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. doi: 10.1177/01466210122032064

Leighton, J. P., Gierl, M. J., & Hunka, S. (2004). The attribute hierarchy model: An approach for integrating cognitive theory with assessment practice. *Journal of Educational Measurement*, *41*, 205–236. doi: 10.1111/j.1745-984.2004.tb01163.x

Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika, 64*, 178-212. doi: 10.1007/BF02294535

National Council of Teachers of Mathematics (NCTM; 2000). *Principles and standards for school mathematics.* Reston, VA: NCTM.

Park, Y. S., Lee, Y.-S., & Choi, K. M. (2010, April). *Cognitive diagnostic analysis of timss 2007 using the DINA model: A multilevel comparison of high, average, and low performance.* Paper presented at the annual meeting of the American Educational Research Association, Denver, CO.

Rupp, A. A., & Templin, J. L. (2008a). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Educational and Psychological Measurement, 68,* 78-96. doi: 10.1177/0013164407301545

Rupp, A. A., & Templin, J. L. (2008b). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement, 6*, 219-262. doi: 10.1080/15366360802490866

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

Su, Y.-L. (2013). *Cognitive diagnostic analysis using hierarchically structured skills.* Unpublished doctoral dissertation. The University of Iowa, IA.

Su, Y.-L., Choi, K. M., Lee, W., Choi, T., & McAninch, M. (2013). Hierarchical cognitive diagnostic analysis for TIMSS 2003 mathematics. *CASMA Research Report, 35.* Retrived from http://www.education.uiowa.edu/docs/default-source/casma—research/test-elementary.pdf?sfvrsn=2

Table 1. Linear Attribute Hierarchy Profiles of Six Attributes

| | Attribute 1 | Attribute 2 | Attribute 3 | Attribute 4 | Attribute 5 | Attribute 6 |
|---|---|---|---|---|---|---|
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 5 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 6 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 7 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 2. Convergent Attribute Hierarchy Profiles of Six Attributes

| | Attribute 1 | Attribute 2 | Attribute 3 | Attribute 4 | Attribute 5 | Attribute 6 |
|---|---|---|---|---|---|---|
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 5 | 1 | 1 | 0 | 1 | 0 | 0 |
| Profile 6 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 7 | 1 | 1 | 0 | 1 | 1 | 0 |
| Profile 8 | 1 | 1 | 1 | 0 | 1 | 0 |
| Profile 9 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 10 | 1 | 1 | 0 | 1 | 1 | 1 |
| Profile 11 | 1 | 1 | 1 | 0 | 1 | 1 |
| Profile 12 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 3. Divergent Attribute Hierarchy Profiles of Six Attributes

| | Attribute 1 | Attribute 2 | Attribute 3 | Attribute 4 | Attribute 5 | Attribute 6 |
|---|---|---|---|---|---|---|
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 5 | 1 | 0 | 0 | 1 | 0 | 0 |
| Profile 6 | 1 | 0 | 0 | 1 | 1 | 0 |
| Profile 7 | 1 | 0 | 0 | 1 | 0 | 1 |
| Profile 8 | 1 | 0 | 0 | 1 | 1 | 1 |
| Profile 9 | 1 | 1 | 0 | 1 | 0 | 0 |
| Profile 10 | 1 | 1 | 0 | 1 | 1 | 0 |
| Profile 11 | 1 | 1 | 0 | 1 | 0 | 1 |
| Profile 12 | 1 | 1 | 0 | 1 | 1 | 1 |
| Profile 13 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 14 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 15 | 1 | 1 | 1 | 1 | 0 | 1 |
| Profile 16 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 4. Unstructured Attribute Hierarchy Profiles of Six Attributes

| | Attribute 1 | Attribute 2 | Attribute 3 | Attribute 4 | Attribute 5 | Attribute 6 |
|---|---|---|---|---|---|---|
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 0 | 1 | 0 | 0 | 0 |
| Profile 5 | 1 | 0 | 0 | 1 | 0 | 0 |
| Profile 6 | 1 | 0 | 0 | 0 | 1 | 0 |
| Profile 7 | 1 | 0 | 0 | 0 | 0 | 1 |
| Profile 8 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 9 | 1 | 1 | 0 | 1 | 0 | 0 |
| Profile 10 | 1 | 1 | 0 | 0 | 1 | 0 |
| Profile 11 | 1 | 1 | 0 | 0 | 0 | 1 |
| Profile 12 | 1 | 0 | 1 | 1 | 0 | 0 |
| Profile 13 | 1 | 0 | 1 | 0 | 1 | 0 |
| Profile 14 | 1 | 0 | 1 | 0 | 0 | 1 |
| Profile 15 | 1 | 0 | 0 | 1 | 1 | 0 |
| Profile 16 | 1 | 0 | 0 | 1 | 0 | 1 |
| Profile 17 | 1 | 0 | 0 | 0 | 1 | 1 |
| Profile 18 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 19 | 1 | 1 | 1 | 0 | 1 | 0 |
| Profile 20 | 1 | 1 | 1 | 0 | 0 | 1 |
| Profile 21 | 1 | 1 | 0 | 1 | 1 | 0 |
| Profile 22 | 1 | 1 | 0 | 1 | 0 | 1 |
| Profile 23 | 1 | 1 | 0 | 0 | 1 | 1 |
| Profile 24 | 1 | 0 | 1 | 1 | 1 | 0 |
| Profile 25 | 1 | 0 | 1 | 0 | 1 | 1 |
| Profile 26 | 1 | 0 | 1 | 1 | 0 | 1 |
| Profile 27 | 1 | 0 | 0 | 1 | 1 | 1 |
| Profile 28 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 29 | 1 | 1 | 1 | 0 | 1 | 1 |
| Profile 30 | 1 | 1 | 0 | 1 | 1 | 1 |
| Profile 31 | 1 | 0 | 1 | 1 | 1 | 1 |
| Profile 32 | 1 | 1 | 1 | 1 | 0 | 1 |
| Profile 33 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 5. Linear Attribute Hierarchy Profiles of Eight Attributes

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 5 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 6 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 7 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 6. Convergent Attribute Hierarchy Profiles of Eight Attributes

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 5 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Profile 6 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| Profile 7 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Profile 8 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 9 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 10 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| Profile 11 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| Profile 12 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 13 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| Profile 14 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| Profile 15 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 |
| Profile 16 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| Profile 17 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 |
| Profile 18 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 19 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 |
| Profile 20 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| Profile 21 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| Profile 22 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| Profile 23 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| Profile 24 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| Profile 25 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| Profile 26 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| Profile 27 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 7. Divergent Attribute Hierarchy Profiles of Eight Attributes

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 5 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Profile 6 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| Profile 7 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| Profile 8 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| Profile 9 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Profile 10 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| Profile 11 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| Profile 12 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| Profile 13 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 14 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 15 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| Profile 16 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 17 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| Profile 18 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| Profile 19 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 |
| Profile 20 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| Profile 21 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 22 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| Profile 23 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| Profile 24 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 |
| Profile 25 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| Profile 26 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| Profile 27 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| Profile 28 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| Profile 29 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| Profile 30 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| Profile 31 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 8. Unstructured Attribute Hierarchy Profiles of Eight Attributes

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Profile 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 5 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Profile 6 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Profile 7 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Profile 8 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Profile 9 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Profile 10 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Profile 11 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Profile 12 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| Profile 13 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| Profile 14 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| Profile 15 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| Profile 16 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 17 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| Profile 18 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Profile 19 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| Profile 20 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Profile 21 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| Profile 22 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| Profile 23 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| Profile 24 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| Profile 25 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| Profile 26 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| Profile 27 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| Profile 28 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| Profile 29 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| Profile 30 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| Profile 31 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| Profile 32 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |

Table 8. (continued)

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 33 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| Profile 34 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| Profile 35 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| Profile 36 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| Profile 37 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| Profile 38 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| Profile 39 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| Profile 40 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| Profile 41 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| Profile 42 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| Profile 43 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| Profile 44 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Profile 45 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| Profile 46 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 47 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| Profile 48 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| Profile 49 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| Profile 50 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| Profile 51 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| Profile 52 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| Profile 53 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| Profile 54 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| Profile 55 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| Profile 56 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| Profile 57 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| Profile 58 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |
| Profile 59 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| Profile 60 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| Profile 61 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| Profile 62 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| Profile 63 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| Profile 64 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |

Table 8. (continued)

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 65 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| Profile 66 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| Profile 67 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| Profile 68 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| Profile 69 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| Profile 70 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| Profile 71 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| Profile 72 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| Profile 73 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| Profile 74 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| Profile 75 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| Profile 76 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| Profile 77 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| Profile 78 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| Profile 79 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| Profile 80 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| Profile 81 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| Profile 82 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 |
| Profile 83 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| Profile 84 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| Profile 85 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| Profile 86 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 87 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |
| Profile 88 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 |
| Profile 89 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| Profile 90 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| Profile 91 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| Profile 92 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| Profile 93 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| Profile 94 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| Profile 95 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| Profile 96 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |

Table 8. (continued)

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 97 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| Profile 98 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| Profile 99 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 |
| Profile 100 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| Profile 101 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| Profile 102 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 |
| Profile 103 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 |
| Profile 104 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| Profile 105 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| Profile 106 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| Profile 107 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 |
| Profile 108 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| Profile 109 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| Profile 110 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| Profile 111 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| Profile 112 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 |
| Profile 113 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| Profile 114 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| Profile 115 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| Profile 116 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 117 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| Profile 118 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
| Profile 119 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| Profile 120 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| Profile 121 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| Profile 122 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Profile 123 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| Profile 124 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| Profile 125 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |

Table 8. (continued)

| | Attribute | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Profile 126 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| Profile 127 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| Profile 128 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| Profile 129 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table 9. The Possible Number of Attribute Profiles for Each Hierarchy Model

| Attribute Hierarchy | Six Attributes | Eight Attributes |
|---|---|---|
| Linear | 7 | 9 |
| Convergent | 12 | 27 |
| Divergent | 16 | 31 |
| Unstructured | 33 | 129 |
| Baseline | $2^6 = 64$ | $2^8 = 256$ |

Table 10. The List of Simulation Factors

| | | | |
|---|---|---|---|
| Number of Attributes | 6 | 8 | |
| Test Length | 12 | 30 | |
| Data Generating Model | DINA | DINA-$\mathrm{H_L}$ | DINA-$\mathrm{H_U}$ |
| | DINO | DINO-$\mathrm{H_L}$ | DINO-$\mathrm{H_U}$ |
| Sample Size | 300 | 1000 | 3000 |
| Estimation Model | DINA | DINA-$\mathrm{H_L}$ | DINA-$\mathrm{H_U}$ |
| | DINO | DINO-$\mathrm{H_L}$ | DINO-$\mathrm{H_U}$ |

Table 11. Q-Matrix for the Eighth Grade TIMSS 2007 Mathematics Test

| Attribute Item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Total Number |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 3 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 4 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 6 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 7 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| 8 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 9 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 10 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 11 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 2 |
| 13 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 3 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 3 |
| 15 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 18 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 19 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 20 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 3 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 3 |
| 23 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 24 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 25 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 26 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 27 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 28 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 4 |
| 29 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| Sum | 9 | 4 | 15 | 9 | 8 | 3 | 5 | 2 | 5 | 1 | 3 | 7 | |

Note: The table was from Park et al. (2010) and Choi (2011).

Table 12. The List of Conditions for the DINA and DINA–H models

| Condition | Number of Attribute | Test Length | Generating Model | Sample Size | Estimation Model |
|---|---|---|---|---|---|
| K6_J12_NA_300_NA | | | | | DINA |
| K6_J12_NA_300_HL | | | | 300 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_NA_300_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_NA_1000_NA | | | | | DINA |
| K6_J12_NA_1000_HL | | | DINA | 1000 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_NA_1000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_NA_3000_NA | | | | | DINA |
| K6_J12_NA_3000_HL | | | | 3000 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_NA_3000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_HL_300_NA | | | | | DINA |
| K6_J12_HL_300_HL | | | | 300 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_HL_300_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_HL_1000_NA | | | | | DINA |
| K6_J12_HL_1000_HL | 6 | 12 | DINA-H$_{\mathrm{L}}$ | 1000 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_HL_1000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_HL_3000_NA | | | | | DINA |
| K6_J12_HL_3000_HL | | | | 3000 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_HL_3000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_HU_300_NA | | | | | DINA |
| K6_J12_HU_300_HL | | | | 300 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_HU_300_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_HU_1000_NA | | | | | DINA |
| K6_J12_HU_1000_HL | | | DINA-H$_{\mathrm{U}}$ | 1000 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_HU_1000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J12_HU_3000_NA | | | | | DINA |
| K6_J12_HU_3000_HL | | | | 3000 | DINA-H$_{\mathrm{L}}$ |
| K6_J12_HU_3000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J30_NA_300_NA | | | | | DINA |
| K6_J30_NA_300_HL | | | | 300 | DINA-H$_{\mathrm{L}}$ |
| K6_J30_NA_300_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J30_NA_1000_NA | | | | | DINA |
| K6_J30_NA_1000_HL | 6 | 30 | DINA | 1000 | DINA-H$_{\mathrm{L}}$ |
| K6_J30_NA_1000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J30_NA_3000_NA | | | | | DINA |
| K6_J30_NA_3000_HL | | | | 3000 | DINA-H$_{\mathrm{L}}$ |
| K6_J30_NA_3000_HU | | | | | DINA-H$_{\mathrm{U}}$ |
| K6_J30_HL_300_NA | | | | | DINA |
| K6_J30_HL_300_HL | | | DINA-H$_{\mathrm{L}}$ | 300 | DINA-H$_{\mathrm{L}}$ |
| K6_J30_HL_300_HU | | | | | DINA-H$_{\mathrm{U}}$ |

Table 12. (continued)

| Condition | Number of Attribute | Test Length | Generating Model | Sample Size | Estimation Model |
|---|---|---|---|---|---|
| K6_J30_HL_1000_NA | | | | | DINA |
| K6_J30_HL_1000_HL | | | DINA-$H_L$ | 1000 | DINA-$H_L$ |
| K6_J30_HL_1000_HU | | | | | DINA-$H_U$ |
| K6_J30_HL_3000_NA | | | | | DINA |
| K6_J30_HL_3000_HL | | | | 3000 | DINA-$H_L$ |
| K6_J30_HL_3000_HU | 6 | 30 | | | DINA-$H_U$ |
| K6_J30_HU_300_NA | | | | | DINA |
| K6_J30_HU_300_HL | | | | 300 | DINA-$H_L$ |
| K6_J30_HU_300_HU | | | | | DINA-$H_U$ |
| K6_J30_HU_1000_NA | | | | | DINA |
| K6_J30_HU_1000_HL | | | DINA-$H_U$ | 1000 | DINA-$H_L$ |
| K6_J30_HU_1000_HU | | | | | DINA-$H_U$ |
| K6_J30_HU_3000_NA | | | | | DINA |
| K6_J30_HU_3000_HL | | | | 3000 | DINA-$H_L$ |
| K6_J30_HU_3000_HU | | | | | DINA-$H_U$ |
| K8_J12_NA_300_NA | | | | | DINA |
| K8_J12_NA_300_HL | | | | 300 | DINA-$H_L$ |
| K8_J12_NA_300_HU | | | | | DINA-$H_U$ |
| K8_J12_NA_1000_NA | | | | | DINA |
| K8_J12_NA_1000_HL | | | DINA | 1000 | DINA-$H_L$ |
| K8_J12_NA_1000_HU | | | | | DINA-$H_U$ |
| K8_J12_NA_3000_NA | | | | | DINA |
| K8_J12_NA_3000_HL | | | | 3000 | DINA-$H_L$ |
| K8_J12_NA_3000_HU | | | | | DINA-$H_U$ |
| K8_J12_HL_300_NA | | | | | DINA |
| K8_J12_HL_300_HL | | | | 300 | DINA-$H_L$ |
| K8_J12_HL_300_HU | | | | | DINA-$H_U$ |
| K8_J12_HL_1000_NA | | | | | DINA |
| K8_J12_HL_1000_HL | 8 | 12 | DINA-$H_L$ | 1000 | DINA-$H_L$ |
| K8_J12_HL_1000_HU | | | | | DINA-$H_U$ |
| K8_J12_HL_3000_NA | | | | | DINA |
| K8_J12_HL_3000_HL | | | | 3000 | DINA-$H_L$ |
| K8_J12_HL_3000_HU | | | | | DINA-$H_U$ |
| K8_J12_HU_300_NA | | | | | DINA |
| K8_J12_HU_300_HL | | | | 300 | DINA-$H_L$ |
| K8_J12_HU_300_HU | | | | | DINA-$H_U$ |
| K8_J12_HU_1000_NA | | | | | DINA |
| K8_J12_HU_1000_HL | | | DINA-$H_U$ | 1000 | DINA-$H_L$ |
| K8_J12_HU_1000_HU | | | | | DINA-$H_U$ |

Table 12. (continued)

| Condition | Number of Attribute | Test Length | Generating Model | Sample Size | Estimation Model |
|---|---|---|---|---|---|
| K8_J12_HU_3000_NA | | | | | DINA |
| K8_J12_HU_3000_HL | 8 | 12 | DINA-$H_U$ | 3000 | DINA-$H_L$ |
| K8_J12_HU_3000_HU | | | | | DINA-$H_U$ |
| K8_J30_NA_300_NA | | | | | DINA |
| K8_J30_NA_300_HL | | | | 300 | DINA-$H_L$ |
| K8_J30_NA_300_HU | | | | | DINA-$H_U$ |
| K8_J30_NA_1000_NA | | | | | DINA |
| K8_J30_NA_1000_HL | | | DINA | 1000 | DINA-$H_L$ |
| K8_J30_NA_1000_HU | | | | | DINA-$H_U$ |
| K8_J30_NA_3000_NA | | | | | DINA |
| K8_J30_NA_3000_HL | | | | 3000 | DINA-$H_L$ |
| K8_J30_NA_3000_HU | | | | | DINA-$H_U$ |
| K8_J30_HL_300_NA | | | | | DINA |
| K8_J30_HL_300_HL | | | | 300 | DINA-$H_L$ |
| K8_J30_HL_300_HU | | | | | DINA-$H_U$ |
| K8_J30_HL_1000_NA | | | | | DINA |
| K8_J30_HL_1000_HL | 8 | 30 | DINA-$H_L$ | 1000 | DINA-$H_L$ |
| K8_J30_HL_1000_HU | | | | | DINA-$H_U$ |
| K8_J30_HL_3000_NA | | | | | DINA |
| K8_J30_HL_3000_HL | | | | 3000 | DINA-$H_L$ |
| K8_J30_HL_3000_HU | | | | | DINA-$H_U$ |
| K8_J30_HU_300_NA | | | | | DINA |
| K8_J30_HU_300_HL | | | | 300 | DINA-$H_L$ |
| K8_J30_HU_300_HU | | | | | DINA-$H_U$ |
| K8_J30_HU_1000_NA | | | | | DINA |
| K8_J30_HU_1000_HL | | | DINA-$H_U$ | 1000 | DINA-$H_L$ |
| K8_J30_HU_1000_HU | | | | | DINA-$H_U$ |
| K8_J30_HU_3000_NA | | | | | DINA |
| K8_J30_HU_3000_HL | | | | 3000 | DINA-$H_L$ |
| K8_J30_HU_3000_HU | | | | | DINA-$H_U$ |

Table 13. The List of Conditions for the DINO and DINO–H models

| Condition | Number of Attribute | Test Length | Generating Model | Sample Size | Estimation Model |
|---|---|---|---|---|---|
| K6_J12_NA_300_NA | | | | | DINO |
| K6_J12_NA_300_HL | | | | 300 | DINO-H$_{\text{L}}$ |
| K6_J12_NA_300_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_NA_1000_NA | | | | | DINO |
| K6_J12_NA_1000_HL | | | DINO | 1000 | DINO-H$_{\text{L}}$ |
| K6_J12_NA_1000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_NA_3000_NA | | | | | DINO |
| K6_J12_NA_3000_HL | | | | 3000 | DINO-H$_{\text{L}}$ |
| K6_J12_NA_3000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_HL_300_NA | | | | | DINO |
| K6_J12_HL_300_HL | | | | 300 | DINO-H$_{\text{L}}$ |
| K6_J12_HL_300_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_HL_1000_NA | | | | | DINO |
| K6_J12_HL_1000_HL | 6 | 12 | DINO-H$_{\text{L}}$ | 1000 | DINO-H$_{\text{L}}$ |
| K6_J12_HL_1000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_HL_3000_NA | | | | | DINO |
| K6_J12_HL_3000_HL | | | | 3000 | DINO-H$_{\text{L}}$ |
| K6_J12_HL_3000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_HU_300_NA | | | | | DINO |
| K6_J12_HU_300_HL | | | | 300 | DINO-H$_{\text{L}}$ |
| K6_J12_HU_300_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_HU_1000_NA | | | | | DINO |
| K6_J12_HU_1000_HL | | | DINO-H$_{\text{U}}$ | 1000 | DINO-H$_{\text{L}}$ |
| K6_J12_HU_1000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J12_HU_3000_NA | | | | | DINO |
| K6_J12_HU_3000_HL | | | | 3000 | DINO-H$_{\text{L}}$ |
| K6_J12_HU_3000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J30_NA_300_NA | | | | | DINO |
| K6_J30_NA_300_HL | | | | 300 | DINO-H$_{\text{L}}$ |
| K6_J30_NA_300_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J30_NA_1000_NA | | | | | DINO |
| K6_J30_NA_1000_HL | 6 | 30 | DINO | 1000 | DINO-H$_{\text{L}}$ |
| K6_J30_NA_1000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J30_NA_3000_NA | | | | | DINO |
| K6_J30_NA_3000_HL | | | | 3000 | DINO-H$_{\text{L}}$ |
| K6_J30_NA_3000_HU | | | | | DINO-H$_{\text{U}}$ |
| K6_J30_HL_300_NA | | | | | DINO |
| K6_J30_HL_300_HL | | | DINO-H$_{\text{L}}$ | 300 | DINO-H$_{\text{L}}$ |
| K6_J30_HL_300_HU | | | | | DINO-H$_{\text{U}}$ |

Table 13. (continued)

| Condition | Number of Attribute | Test Length | Generating Model | Sample Size | Estimation Model |
|---|---|---|---|---|---|
| K6_J30_HL_1000_NA | | | | | DINO |
| K6_J30_HL_1000_HL | | | DINO-$H_L$ | 1000 | DINO-$H_L$ |
| K6_J30_HL_1000_HU | | | | | DINO-$H_U$ |
| K6_J30_HL_3000_NA | | | | | DINO |
| K6_J30_HL_3000_HL | | | | 3000 | DINO-$H_L$ |
| K6_J30_HL_3000_HU | 6 | 30 | | | DINO-$H_U$ |
| K6_J30_HU_300_NA | | | | | DINO |
| K6_J30_HU_300_HL | | | | 300 | DINO-$H_L$ |
| K6_J30_HU_300_HU | | | | | DINO-$H_U$ |
| K6_J30_HU_1000_NA | | | | | DINO |
| K6_J30_HU_1000_HL | | | DINO-$H_U$ | 1000 | DINO-$H_L$ |
| K6_J30_HU_1000_HU | | | | | DINO-$H_U$ |
| K6_J30_HU_3000_NA | | | | | DINO |
| K6_J30_HU_3000_HL | | | | 3000 | DINO-$H_L$ |
| K6_J30_HU_3000_HU | | | | | DINO-$H_U$ |
| K8_J12_NA_300_NA | | | | | DINO |
| K8_J12_NA_300_HL | | | | 300 | DINO-$H_L$ |
| K8_J12_NA_300_HU | | | | | DINO-$H_U$ |
| K8_J12_NA_1000_NA | | | | | DINO |
| K8_J12_NA_1000_HL | | | DINO | 1000 | DINO-$H_L$ |
| K8_J12_NA_1000_HU | | | | | DINO-$H_U$ |
| K8_J12_NA_3000_NA | | | | | DINO |
| K8_J12_NA_3000_HL | | | | 3000 | DINO-$H_L$ |
| K8_J12_NA_3000_HU | | | | | DINO-$H_U$ |
| K8_J12_HL_300_NA | | | | | DINO |
| K8_J12_HL_300_HL | | | | 300 | DINO-$H_L$ |
| K8_J12_HL_300_HU | | | | | DINO-$H_U$ |
| K8_J12_HL_1000_NA | | | | | DINO |
| K8_J12_HL_1000_HL | 8 | 12 | DINO-$H_L$ | 1000 | DINO-$H_L$ |
| K8_J12_HL_1000_HU | | | | | DINO-$H_U$ |
| K8_J12_HL_3000_NA | | | | | DINO |
| K8_J12_HL_3000_HL | | | | 3000 | DINO-$H_L$ |
| K8_J12_HL_3000_HU | | | | | DINO-$H_U$ |
| K8_J12_HU_300_NA | | | | | DINO |
| K8_J12_HU_300_HL | | | | 300 | DINO-$H_L$ |
| K8_J12_HU_300_HU | | | | | DINO-$H_U$ |
| K8_J12_HU_1000_NA | | | | | DINO |
| K8_J12_HU_1000_HL | | | DINO-$H_U$ | 1000 | DINO-$H_L$ |
| K8_J12_HU_1000_HU | | | | | DINO-$H_U$ |

Table 13. (continued)

| Condition | Number of Attribute | Test Length | Generating Model | Sample Size | Estimation Model |
|---|---|---|---|---|---|
| K8_J12_HU_3000_NA | | | | | DINO |
| K8_J12_HU_3000_HL | 8 | 12 | DINO-$H_U$ | 3000 | DINO-$H_L$ |
| K8_J12_HU_3000_HU | | | | | DINO-$H_U$ |
| K8_J30_NA_300_NA | | | | | DINO |
| K8_J30_NA_300_HL | | | | 300 | DINO-$H_L$ |
| K8_J30_NA_300_HU | | | | | DINO-$H_U$ |
| K8_J30_NA_1000_NA | | | | | DINO |
| K8_J30_NA_1000_HL | | | DINO | 1000 | DINO-$H_L$ |
| K8_J30_NA_1000_HU | | | | | DINO-$H_U$ |
| K8_J30_NA_3000_NA | | | | | DINO |
| K8_J30_NA_3000_HL | | | | 3000 | DINO-$H_L$ |
| K8_J30_NA_3000_HU | | | | | DINO-$H_U$ |
| K8_J30_HL_300_NA | | | | | DINO |
| K8_J30_HL_300_HL | | | | 300 | DINO-$H_L$ |
| K8_J30_HL_300_HU | | | | | DINO-$H_U$ |
| K8_J30_HL_1000_NA | | | | | DINO |
| K8_J30_HL_1000_HL | 8 | 30 | DINO-$H_L$ | 1000 | DINO-$H_L$ |
| K8_J30_HL_1000_HU | | | | | DINO-$H_U$ |
| K8_J30_HL_3000_NA | | | | | DINO |
| K8_J30_HL_3000_HL | | | | 3000 | DINO-$H_L$ |
| K8_J30_HL_3000_HU | | | | | DINO-$H_U$ |
| K8_J30_HU_300_NA | | | | | DINO |
| K8_J30_HU_300_HL | | | | 300 | DINO-$H_L$ |
| K8_J30_HU_300_HU | | | | | DINO-$H_U$ |
| K8_J30_HU_1000_NA | | | | | DINO |
| K8_J30_HU_1000_HL | | | DINO-$H_U$ | 1000 | DINO-$H_L$ |
| K8_J30_HU_1000_HU | | | | | DINO-$H_U$ |
| K8_J30_HU_3000_NA | | | | | DINO |
| K8_J30_HU_3000_HL | | | | 3000 | DINO-$H_L$ |
| K8_J30_HU_3000_HU | | | | | DINO-$H_U$ |

Table 14. Results of Fit Indices of the Main Effect of Model Consistency for DINA(-H)

| Overall | MAIC | MBIC |
|---|---|---|
| HL_NA | 33578 | 34557 |
| **HL_HL** | **33281** | **33519** |
| HL_HU | 33423 | 34017 |
| HU_NA | 35772 | 36751 |
| HU_HL | 36685 | 36924 |
| **HU_HU** | **35619** | **36213** |
| **NA_NA** | **36161** | **37140** |
| NA_HL | 37013 | 37251 |
| NA_HU | 36562 | 37157 |

Table 15. Summary Statistics for the Main Effect of Model Consistency for DINA(-H)

| Overall | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|
| HL_NA | 0.00268 | **0.00061** | 0.00329 | 0.00289 | 0.00155 | 0.00443 |
| **HL_HL** | **0.00144** | 0.00063 | **0.00206** | **0.00186** | **0.00149** | **0.00334** |
| HL_HU | 0.00227 | 0.00079 | 0.00306 | 0.00294 | 0.00156 | 0.00450 |
| HU_NA | 0.00253 | **0.00172** | 0.00424 | 0.00236 | 0.00298 | 0.00534 |
| HU_HL | 0.01146 | 0.00268 | 0.01414 | 0.04537 | 0.00414 | 0.04951 |
| **HU_HU** | **0.00139** | 0.00197 | **0.00336** | **0.00215** | **0.00291** | **0.00506** |
| **NA_NA** | **0.00146** | **0.00106** | **0.00251** | **0.00248** | **0.00378** | **0.00626** |
| NA_HL | 0.01193 | 0.00157 | 0.01350 | 0.05113 | 0.00531 | 0.05644 |
| NA_HU | 0.00648 | 0.00119 | 0.00767 | 0.00911 | 0.00455 | 0.01367 |

Table 16. Summary Statistics for the Main Effect of Number of Attribute for DINA(-H)

| | | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| K=6 | HL_HL | 0.002754 | 0.000816 | 0.003571 | 0.003678 | 0.001393 | 0.005071 |
| | HU_HU | 0.002649 | 0.002917 | 0.005566 | 0.003817 | 0.002147 | 0.005964 |
| | NA_NA | 0.002665 | 0.001007 | 0.003672 | 0.004219 | 0.002828 | 0.007046 |
| | **Mean** | 0.002690 | 0.001580 | 0.004270 | 0.003905 | **0.002123** | 0.006027 |
| K=8 | HL_HL | 0.000118 | 0.000438 | 0.000556 | 0.000035 | 0.001582 | 0.001617 |
| | HU_HU | 0.000126 | 0.001020 | 0.001146 | 0.000493 | 0.003671 | 0.004164 |
| | NA_NA | 0.000252 | 0.001104 | 0.001356 | 0.000745 | 0.004737 | 0.005483 |
| | **Mean** | **0.000165** | **0.000854** | **0.001019** | **0.000424** | 0.003330 | **0.003754** |

Table 17. Summary Statistics for the Main Effect of Test Length for DINA(-H)

| | | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| J=12 | HL_HL | 0.000167 | 0.000697 | 0.000864 | 0.000041 | 0.001836 | 0.001877 |
| | HU_HU | 0.000180 | 0.003069 | 0.003249 | 0.000619 | 0.003740 | 0.004359 |
| | NA_NA | 0.000301 | 0.001548 | 0.001850 | 0.000970 | 0.004888 | 0.005858 |
| | **Mean** | **0.000216** | 0.001772 | **0.001988** | **0.000543** | 0.003488 | **0.004031** |
| J=30 | HL_HL | 0.002706 | 0.000557 | 0.003262 | 0.003672 | 0.001139 | 0.004811 |
| | HU_HU | 0.002595 | 0.000868 | 0.003463 | 0.003691 | 0.002078 | 0.005769 |
| | NA_NA | 0.002616 | 0.000562 | 0.003178 | 0.003995 | 0.002677 | 0.006671 |
| | **Mean** | 0.002639 | **0.000662** | 0.003301 | 0.003786 | **0.001965** | 0.005750 |

Table 18. Summary Statistics for the Main Effect of Sample Size for DINA(-H)

| | | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| N=300 | HL_HL | 0.00409 | 0.00134 | 0.00543 | 0.00555 | 0.00328 | 0.00882 |
| | HU_HU | 0.00399 | 0.00376 | 0.00776 | 0.00622 | 0.00640 | 0.01262 |
| | NA_NA | 0.00421 | 0.00235 | 0.00656 | 0.00697 | 0.00793 | 0.01491 |
| | **Mean** | 0.00410 | 0.00249 | 0.00658 | 0.00625 | 0.00587 | 0.01212 |
| N=1000 | HL_HL | 0.00010 | 0.00041 | 0.00050 | 0.00002 | 0.00090 | 0.00092 |
| | HU_HU | 0.00008 | 0.00162 | 0.00170 | 0.00011 | 0.00174 | 0.00185 |
| | NA_NA | 0.00007 | 0.00064 | 0.00071 | 0.00027 | 0.00255 | 0.00281 |
| | **Mean** | **0.00008** | 0.00089 | 0.00097 | 0.00013 | 0.00173 | 0.00186 |
| N=3000 | HL_HL | 0.00013 | 0.00013 | 0.00026 | 0.00001 | 0.00028 | 0.00029 |
| | HU_HU | 0.00009 | 0.00052 | 0.00061 | 0.00014 | 0.00058 | 0.00072 |
| | NA_NA | 0.00009 | 0.00017 | 0.00027 | 0.00021 | 0.00087 | 0.00107 |
| | **Mean** | 0.00010 | **0.00028** | **0.00038** | **0.00012** | **0.00058** | **0.00069** |

Table 19. Summary Statistics for the Interaction Effect of J by K for DINA(-H)

| | | | K=6 | | | K=8 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | J=12 | HL_HL | 0.00011 | 0.00101 | 0.00113 | 0.00022 | 0.00038 | 0.00060 |
| | | HU_HU | 0.00015 | 0.00468 | 0.00483 | 0.00021 | 0.00146 | 0.00167 |
| | | NA_NA | 0.00014 | 0.00146 | 0.00160 | 0.00047 | 0.00163 | 0.00210 |
| | | **Mean** | 0.00013 | 0.00239 | 0.00252 | 0.00030 | 0.00116 | 0.00146 |
| | J=30 | HL_HL | 0.00539 | 0.00062 | 0.00602 | 0.00002 | 0.00049 | 0.00051 |
| | | HU_HU | 0.00515 | 0.00115 | 0.00630 | 0.00004 | 0.00058 | 0.00062 |
| | | NA_NA | 0.00520 | 0.00055 | 0.00574 | 0.00004 | 0.00058 | 0.00061 |
| | | **Mean** | 0.00525 | 0.00077 | 0.00602 | **0.00003** | **0.00055** | **0.00058** |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | J=12 | HL_HL | 0.00004 | 0.00178 | 0.00181 | 0.00004 | 0.00190 | 0.00194 |
| | | HU_HU | 0.00041 | 0.00293 | 0.00334 | 0.00082 | 0.00455 | 0.00538 |
| | | NA_NA | 0.00090 | 0.00403 | 0.00494 | 0.00103 | 0.00574 | 0.00678 |
| | | **Mean** | 0.00045 | 0.00291 | 0.00336 | 0.00063 | 0.00406 | 0.00470 |
| | J=30 | HL_HL | 0.00731 | 0.00101 | 0.00833 | 0.00003 | 0.00127 | 0.00129 |
| | | HU_HU | 0.00722 | 0.00136 | 0.00858 | 0.00016 | 0.00279 | 0.00295 |
| | | NA_NA | 0.00753 | 0.00162 | 0.00915 | 0.00046 | 0.00373 | 0.00419 |
| | | **Mean** | 0.00735 | **0.00133** | 0.00869 | **0.00021** | 0.00260 | **0.00281** |

Table 20. Summary Statistics for the Interaction Effect of N by K for DINA(-H)

| | | | K=6 | | | K=8 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | N=300 | HL_HL | 0.00809 | 0.00173 | 0.00982 | 0.00008 | 0.00095 | 0.00104 |
| | | HU_HU | 0.00774 | 0.00526 | 0.01300 | 0.00025 | 0.00226 | 0.00251 |
| | | NA_NA | 0.00785 | 0.00222 | 0.01007 | 0.00057 | 0.00249 | 0.00306 |
| | | **Mean** | 0.00789 | 0.00307 | 0.01096 | 0.00030 | 0.00190 | 0.00220 |
| | N=1000 | HL_HL | 0.00008 | 0.00055 | 0.00063 | 0.00011 | 0.00027 | 0.00038 |
| | | HU_HU | 0.00012 | 0.00263 | 0.00275 | 0.00004 | 0.00060 | 0.00064 |
| | | NA_NA | 0.00006 | 0.00062 | 0.00068 | 0.00008 | 0.00066 | 0.00075 |
| | | **Mean** | 0.00009 | 0.00127 | 0.00135 | **0.00008** | 0.00051 | 0.00059 |
| | N=3000 | HL_HL | 0.00009 | 0.00017 | 0.00027 | 0.00016 | 0.00009 | 0.00025 |
| | | HU_HU | 0.00009 | 0.00086 | 0.00095 | 0.00009 | 0.00019 | 0.00028 |
| | | NA_NA | 0.00008 | 0.00019 | 0.00027 | 0.00010 | 0.00016 | 0.00027 |
| | | **Mean** | 0.00009 | 0.00041 | 0.00049 | 0.00012 | **0.00015** | **0.00027** |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | N=300 | HL_HL | 0.011010 | 0.003026 | 0.014036 | 0.000080 | 0.003530 | 0.003610 |
| | | HU_HU | 0.011138 | 0.004889 | 0.016027 | 0.001302 | 0.007912 | 0.009214 |
| | | NA_NA | 0.012215 | 0.006007 | 0.018222 | 0.001732 | 0.009861 | 0.011593 |
| | | **Mean** | 0.011454 | 0.004641 | 0.016095 | 0.001038 | 0.007101 | 0.008139 |
| | N=1000 | HL_HL | 0.000014 | 0.000879 | 0.000894 | 0.000016 | 0.000925 | 0.000941 |
| | | HU_HU | 0.000149 | 0.001165 | 0.001314 | 0.000067 | 0.002325 | 0.002392 |
| | | NA_NA | 0.000266 | 0.001911 | 0.002177 | 0.000267 | 0.003181 | 0.003448 |
| | | **Mean** | 0.000143 | 0.001318 | 0.001461 | 0.000117 | 0.002143 | 0.002260 |
| | N=3000 | HL_HL | 0.000008 | 0.000275 | 0.000283 | 0.000009 | 0.000291 | 0.000300 |
| | | HU_HU | 0.000164 | 0.000387 | 0.000551 | 0.000109 | 0.000777 | 0.000886 |
| | | NA_NA | 0.000176 | 0.000565 | 0.000741 | 0.000237 | 0.001171 | 0.001407 |
| | | **Mean** | **0.000116** | **0.000409** | **0.000525** | 0.000118 | 0.000746 | 0.000864 |

Table 21. Summary Statistics for the Interaction Effect of N by J for DINA(-H)

| | | | J=12 | | | J=30 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | N=300 | HL_HL | 0.00008 | 0.00155 | 0.00163 | 0.00809 | 0.00113 | 0.00922 |
| | | HU_HU | 0.00024 | 0.00590 | 0.00614 | 0.00774 | 0.00163 | 0.00937 |
| | | NA_NA | 0.00060 | 0.00351 | 0.00411 | 0.00782 | 0.00119 | 0.00901 |
| | | **Mean** | 0.00031 | 0.00365 | 0.00396 | 0.00789 | 0.00132 | 0.00920 |
| | N=1000 | HL_HL | 0.00018 | 0.00041 | 0.00059 | 0.00001 | 0.00041 | 0.00042 |
| | | HU_HU | 0.00013 | 0.00256 | 0.00269 | 0.00003 | 0.00067 | 0.00070 |
| | | NA_NA | 0.00013 | 0.00090 | 0.00103 | 0.00002 | 0.00038 | 0.00039 |
| | | **Mean** | 0.00015 | 0.00129 | 0.00144 | 0.00002 | 0.00049 | 0.00051 |
| | N=3000 | HL_HL | 0.00024 | 0.00013 | 0.00038 | 0.00001 | 0.00013 | 0.00014 |
| | | HU_HU | 0.00016 | 0.00075 | 0.00091 | 0.00001 | 0.00030 | 0.00031 |
| | | NA_NA | 0.00018 | 0.00023 | 0.00041 | 0.00001 | 0.00012 | 0.00013 |
| | | **Mean** | 0.00020 | 0.00037 | 0.00057 | **0.00001** | **0.00018** | **0.00019** |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | N=300 | HL_HL | 0.00010 | 0.00401 | 0.00411 | 0.01099 | 0.00254 | 0.01354 |
| | | HU_HU | 0.00145 | 0.00839 | 0.00985 | 0.01099 | 0.00441 | 0.01540 |
| | | NA_NA | 0.00212 | 0.01028 | 0.01240 | 0.01183 | 0.00559 | 0.01742 |
| | | **Mean** | 0.00122 | 0.00756 | 0.00878 | 0.01127 | 0.00418 | 0.01545 |
| | N=1000 | HL_HL | 0.00002 | 0.00115 | 0.00117 | 0.00001 | 0.00065 | 0.00066 |
| | | HU_HU | 0.00016 | 0.00210 | 0.00226 | 0.00006 | 0.00139 | 0.00145 |
| | | NA_NA | 0.00042 | 0.00331 | 0.00372 | 0.00012 | 0.00178 | 0.00190 |
| | | **Mean** | 0.00020 | 0.00219 | 0.00238 | 0.00006 | 0.00127 | 0.00134 |
| | N=3000 | HL_HL | 0.00001 | 0.00034 | 0.00035 | 0.00001 | 0.00022 | 0.00023 |
| | | HU_HU | 0.00025 | 0.00072 | 0.00097 | 0.00002 | 0.00044 | 0.00046 |
| | | NA_NA | 0.00037 | 0.00108 | 0.00145 | 0.00004 | 0.00066 | 0.00070 |
| | | **Mean** | 0.00021 | 0.00072 | 0.00093 | **0.00002** | **0.00044** | **0.00046** |

Table 22. Summary Statistics for the Three-Way Interaction Effect of N by J by K for DINA(-H)

| | | | K=6 | | | K=8 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | N=300 | J=12 | 0.000079 | 0.004693 | 0.004772 | 0.000533 | 0.002612 | 0.003145 |
| | | J=30 | 0.015706 | 0.001445 | 0.017151 | 0.000067 | 0.001191 | 0.001258 |
| | N=1000 | J=12 | 0.000158 | 0.001904 | 0.002062 | 0.000137 | 0.000678 | 0.000815 |
| | | J=30 | 0.000020 | 0.000627 | 0.000648 | 0.000020 | 0.000344 | 0.000363 |
| | N=3000 | J=12 | 0.000164 | 0.000560 | 0.000724 | 0.000226 | 0.000183 | 0.000409 |
| | | J=30 | 0.000011 | 0.000251 | 0.000263 | **0.000009** | **0.000116** | **0.000125** |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | N=300 | J=12 | 0.000874 | 0.006396 | 0.007270 | 0.001570 | 0.008727 | 0.010297 |
| | | J=30 | 0.022035 | 0.002885 | 0.024920 | 0.000506 | 0.005475 | 0.005981 |
| | N=1000 | J=12 | 0.000264 | 0.001804 | 0.002069 | 0.000129 | 0.002572 | 0.002701 |
| | | J=30 | 0.000022 | 0.000832 | 0.000854 | 0.000104 | 0.001715 | 0.001819 |
| | N=3000 | J=12 | 0.000216 | 0.000536 | 0.000752 | 0.000205 | 0.000894 | 0.001099 |
| | | J=30 | **0.000016** | **0.000282** | **0.000298** | 0.000031 | 0.000598 | 0.000629 |

Table 23. Results of Fit Indices of the Main Effect of Model Consistency for DINO(-H)

| Overall | MAIC | MBIC |
|---|---|---|
| HL_NO | 32872 | 33851 |
| **HL_HL** | **32576** | **32815** |
| HL_HU | 32721 | 33315 |
| HU_NO | 33278 | 34257 |
| HU_HL | 33445 | **33684** |
| **HU_HU** | **33128** | 33723 |
| **NO_NO** | **35127** | **36106** |
| NO_HL | 36218 | 36457 |
| NO_HU | 36065 | 36659 |

Table 24. Summary Statistics for the Main Effect of Model Consistency for DINO(-H)

| Overall | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|
| HL_NO | 0.00054 | 0.00172 | 0.00198 | 0.00134 | 0.00074 | 0.00207 |
| **HL_HL** | **0.00003** | **0.00143** | **0.00146** | **0.00004** | **0.00061** | **0.00065** |
| HL_HU | 0.00011 | 0.00147 | 0.00158 | 0.00144 | 0.00064 | 0.00208 |
| HU_NO | 0.00525 | **0.01150** | 0.01676 | 0.00034 | 0.00111 | 0.00145 |
| HU_HL | 0.02297 | 0.01368 | 0.03665 | 0.00732 | **0.00079** | 0.00811 |
| **HU_HU** | **0.00055** | 0.01448 | **0.01503** | **0.00022** | 0.00092 | **0.00113** |
| **NO_NO** | **0.00063** | **0.00403** | **0.00466** | **0.00022** | **0.00103** | **0.00124** |
| NO_HL | 0.04939 | 0.00672 | 0.05611 | 0.01022 | 0.00162 | 0.01184 |
| NO_HU | 0.01658 | 0.00756 | 0.02413 | 0.00759 | 0.00110 | 0.00869 |

Table 25. Summary Statistics for the Main Effect of Number of Attribute for DINO(-H)

| | | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| K=6 | HL_HL | 0.000021 | 0.001280 | 0.001302 | 0.000051 | 0.000800 | 0.000851 |
| | HU_HU | 0.000209 | 0.006030 | 0.006239 | 0.000289 | 0.001171 | 0.001460 |
| | NO_NO | 0.000330 | 0.002982 | 0.003311 | 0.000281 | 0.001273 | 0.001554 |
| | **Mean** | **0.000187** | **0.003431** | **0.003617** | 0.000207 | 0.001081 | 0.001288 |
| K=8 | HL_HL | 0.000035 | 0.001579 | 0.001614 | 0.000030 | 0.000427 | 0.000457 |
| | HU_HU | 0.000894 | 0.022922 | 0.023816 | 0.000148 | 0.000661 | 0.000810 |
| | NO_NO | 0.000938 | 0.005070 | 0.006008 | 0.000153 | 0.000779 | 0.000931 |
| | **Mean** | 0.00062 | 0.00985 | 0.01047 | **0.000110** | **0.000622** | **0.000733** |

Table 26. Summary Statistics for the Main Effect of Test Length for DINO(-H)

| | | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| J=12 | HL_HL | 0.000034 | 0.001682 | 0.001716 | 0.000030 | 0.000738 | 0.000768 |
| | HU_HU | 0.000504 | 0.013632 | 0.014136 | 0.000391 | 0.001242 | 0.001633 |
| | NO_NO | 0.000856 | 0.005301 | 0.006157 | 0.000406 | 0.001513 | 0.001919 |
| | **Mean** | 0.000465 | 0.006871 | 0.007336 | 0.000275 | 0.001164 | 0.001440 |
| J=30 | HL_HL | 0.000023 | 0.001177 | 0.001200 | 0.000051 | 0.000489 | 0.000540 |
| | HU_HU | 0.000599 | 0.015320 | 0.015919 | 0.000046 | 0.000590 | 0.000637 |
| | NO_NO | 0.000411 | 0.002751 | 0.003162 | 0.000028 | 0.000539 | 0.000567 |
| | **Mean** | **0.000344** | **0.006416** | **0.006760** | **0.000042** | **0.000539** | **0.000581** |

Table 27. Summary Statistics for the Main Effect of Sample Size for DINO(-H)

| | | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| N=300 | HL_HL | 0.00006 | 0.00304 | 0.00310 | 0.00008 | 0.00134 | 0.00142 |
| | HU_HU | 0.00114 | 0.03251 | 0.03364 | 0.00039 | 0.00206 | 0.00244 |
| | NO_NO | 0.00138 | 0.00864 | 0.01002 | 0.00046 | 0.00233 | 0.00279 |
| | **Mean** | 0.00086 | 0.01473 | 0.01559 | 0.00031 | 0.00191 | 0.00222 |
| N=1000 | HL_HL | 0.00002 | 0.00093 | 0.00095 | 0.00002 | 0.00037 | 0.00040 |
| | HU_HU | 0.00028 | 0.00781 | 0.00809 | 0.00013 | 0.00054 | 0.00067 |
| | NO_NO | 0.00026 | 0.00262 | 0.00288 | 0.00010 | 0.00058 | 0.00068 |
| | **Mean** | 0.00019 | 0.00378 | 0.00397 | **0.00008** | 0.00050 | 0.00058 |
| N=3000 | HL_HL | 0.00001 | 0.00032 | 0.00033 | 0.00002 | 0.00013 | 0.00014 |
| | HU_HU | 0.00024 | 0.00311 | 0.00335 | 0.00014 | 0.00015 | 0.00029 |
| | NO_NO | 0.00026 | 0.00082 | 0.00108 | 0.00009 | 0.00016 | 0.00025 |
| | **Mean** | **0.00017** | **0.00142** | **0.00159** | **0.00008** | **0.00015** | **0.00023** |

Table 28. Summary Statistics for the Interaction Effect of J by K for DINO(-H)

| | | | K=6 | | | K=8 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | J=12 | HL_HL | 0.00002 | 0.00166 | 0.00168 | 0.00004 | 0.00171 | 0.00175 |
| | | HU_HU | 0.00033 | 0.00864 | 0.00896 | 0.00068 | 0.01863 | 0.01931 |
| | | NO_NO | 0.00057 | 0.00419 | 0.00476 | 0.00114 | 0.00641 | 0.00755 |
| | | **Mean** | 0.00031 | 0.00483 | 0.00514 | 0.00062 | 0.00891 | 0.00954 |
| | J=30 | HL_HL | 0.00002 | 0.00090 | 0.00092 | 0.00003 | 0.00145 | 0.00148 |
| | | HU_HU | 0.00009 | 0.00342 | 0.00351 | 0.00111 | 0.02722 | 0.02832 |
| | | NO_NO | 0.00009 | 0.00177 | 0.00186 | 0.00073 | 0.00373 | 0.00446 |
| | | **Mean** | **0.00007** | **0.00203** | **0.00210** | 0.00062 | 0.01080 | 0.01142 |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | J=12 | HL_HL | 0.00005 | 0.00106 | 0.00111 | 0.00001 | 0.00041 | 0.00042 |
| | | HU_HU | 0.00055 | 0.00179 | 0.00234 | 0.00023 | 0.00069 | 0.00093 |
| | | NO_NO | 0.00055 | 0.00204 | 0.00259 | 0.00026 | 0.00099 | 0.00125 |
| | | **Mean** | 0.00038 | 0.00163 | 0.00201 | 0.00017 | 0.00070 | 0.00087 |
| | J=30 | HL_HL | 0.00006 | 0.00053 | 0.00059 | 0.00005 | 0.00044 | 0.00049 |
| | | HU_HU | 0.00003 | 0.00055 | 0.00058 | 0.00006 | 0.00063 | 0.00069 |
| | | NO_NO | 0.00001 | 0.00051 | 0.00052 | 0.00004 | 0.00057 | 0.00061 |
| | | **Mean** | **0.00003** | **0.00053** | **0.00056** | 0.00005 | 0.00055 | 0.00060 |

Table 29. Summary Statistics for the Interaction Effect of N by K for DINO(-H)

| | | | K=6 | | | K=8 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | N=300 | HL_HL | 0.00004 | 0.00273 | 0.00277 | 0.00008 | 0.00335 | 0.00343 |
| | | HU_HU | 0.00035 | 0.01225 | 0.01260 | 0.00192 | 0.05276 | 0.05469 |
| | | NO_NO | 0.00073 | 0.00664 | 0.00737 | 0.00202 | 0.01064 | 0.01267 |
| | | **Mean** | 0.00037 | 0.00721 | 0.00758 | 0.00134 | 0.02225 | 0.02360 |
| | N=1000 | HL_HL | 0.00002 | 0.00083 | 0.00085 | 0.00002 | 0.00102 | 0.00104 |
| | | HU_HU | 0.00013 | 0.00458 | 0.00471 | 0.00043 | 0.01103 | 0.01146 |
| | | NO_NO | 0.00010 | 0.00181 | 0.00191 | 0.00043 | 0.00342 | 0.00385 |
| | | **Mean** | **0.00008** | 0.00241 | 0.00249 | 0.00029 | 0.00516 | 0.00545 |
| | N=3000 | HL_HL | 0.00001 | 0.00028 | 0.00028 | 0.00001 | 0.00036 | 0.00037 |
| | | HU_HU | 0.00015 | 0.00126 | 0.00140 | 0.00032 | 0.00497 | 0.00530 |
| | | NO_NO | 0.00016 | 0.00049 | 0.00065 | 0.00036 | 0.00114 | 0.00151 |
| | | **Mean** | 0.00010 | **0.00068** | **0.00078** | 0.00023 | 0.00216 | 0.00239 |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | N=300 | HL_HL | 0.000108 | 0.001740 | 0.001848 | 0.000058 | 0.000940 | 0.000998 |
| | | HU_HU | 0.000540 | 0.002614 | 0.003153 | 0.000236 | 0.001498 | 0.001734 |
| | | NO_NO | 0.000671 | 0.002911 | 0.003582 | 0.000253 | 0.001751 | 0.002005 |
| | | **Mean** | 0.000439 | 0.002422 | 0.002861 | 0.000183 | 0.001396 | 0.001579 |
| | N=1000 | HL_HL | 0.000025 | 0.000494 | 0.000519 | 0.000019 | 0.000256 | 0.000275 |
| | | HU_HU | 0.000201 | 0.000697 | 0.000898 | 0.000061 | 0.000380 | 0.000441 |
| | | NO_NO | 0.000101 | 0.000719 | 0.000820 | 0.000098 | 0.000450 | 0.000547 |
| | | **Mean** | 0.000109 | 0.000637 | 0.000746 | **0.000059** | 0.000362 | 0.000421 |
| | N=3000 | HL_HL | 0.000020 | 0.000166 | 0.000186 | 0.000012 | 0.000085 | 0.000098 |
| | | HU_HU | 0.000125 | 0.000202 | 0.000328 | 0.000148 | 0.000106 | 0.000254 |
| | | NO_NO | 0.000072 | 0.000189 | 0.000261 | 0.000107 | 0.000135 | 0.000241 |
| | | **Mean** | 0.000073 | 0.000186 | 0.000258 | 0.000089 | **0.000109** | **0.000198** |

Table 30. Summary Statistics for the Interaction Effect of N by J for DINO(-H)

| | | | J=12 | | | J=30 | | |
|---|---|---|---|---|---|---|---|---|
| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
| G | N=300 | HL_HL | 0.00007 | 0.00362 | 0.00368 | 0.00005 | 0.00247 | 0.00252 |
| | | HU_HU | 0.00070 | 0.02576 | 0.02645 | 0.00158 | 0.03925 | 0.04083 |
| | | NO_NO | 0.00177 | 0.01147 | 0.01323 | 0.00099 | 0.00582 | 0.00681 |
| | | **Mean** | 0.00084 | 0.01361 | 0.01446 | 0.00087 | 0.01585 | 0.01672 |
| | N=1000 | HL_HL | 0.00002 | 0.00106 | 0.00108 | 0.00001 | 0.00079 | 0.00081 |
| | | HU_HU | 0.00041 | 0.01101 | 0.01142 | 0.00015 | 0.00461 | 0.00476 |
| | | NO_NO | 0.00040 | 0.00350 | 0.00390 | 0.00012 | 0.00174 | 0.00186 |
| | | **Mean** | 0.00028 | 0.00519 | 0.00547 | 0.00009 | 0.00238 | 0.00247 |
| | N=3000 | HL_HL | 0.00001 | 0.00037 | 0.00038 | 0.00001 | 0.00027 | 0.00027 |
| | | HU_HU | 0.00040 | 0.00413 | 0.00453 | 0.00007 | 0.00210 | 0.00217 |
| | | NO_NO | 0.00040 | 0.00094 | 0.00134 | 0.00012 | 0.00070 | 0.00082 |
| | | **Mean** | 0.00027 | 0.00181 | 0.00208 | **0.00007** | **0.00102** | **0.00109** |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | N=300 | HL_HL | 0.00005 | 0.00161 | 0.00167 | 0.00011 | 0.00107 | 0.00118 |
| | | HU_HU | 0.00066 | 0.00282 | 0.00348 | 0.00011 | 0.00130 | 0.00141 |
| | | NO_NO | 0.00085 | 0.00348 | 0.00434 | 0.00007 | 0.00118 | 0.00125 |
| | | **Mean** | 0.00052 | 0.00264 | 0.00316 | 0.00010 | 0.00118 | 0.00128 |
| | N=1000 | HL_HL | 0.00002 | 0.00045 | 0.00047 | 0.00003 | 0.00030 | 0.00033 |
| | | HU_HU | 0.00024 | 0.00072 | 0.00096 | 0.00002 | 0.00036 | 0.00038 |
| | | NO_NO | 0.00019 | 0.00084 | 0.00103 | 0.00001 | 0.00033 | 0.00034 |
| | | **Mean** | 0.00015 | 0.00067 | 0.00082 | 0.00002 | 0.00033 | 0.00035 |
| | N=3000 | HL_HL | 0.00002 | 0.00015 | 0.00017 | 0.00001 | 0.00010 | 0.00011 |
| | | HU_HU | 0.00027 | 0.00019 | 0.00046 | 0.00001 | 0.00011 | 0.00012 |
| | | NO_NO | 0.00017 | 0.00022 | 0.00039 | 0.00000 | 0.00011 | 0.00011 |
| | | **Mean** | 0.00015 | 0.00019 | 0.00034 | **0.00001** | **0.00011** | **0.00012** |

Table 31. Summary Statistics for the Three-Way Interaction Effect of N by J by K for DINO(-H)

| | | | ASB(g) | AVAR(g) | AMSE(g) | ASB(g) | AVAR(g) | AMSE(g) |
|---|---|---|---|---|---|---|---|---|
| | | | | K=6 | | | K=8 | |
| G | N=300 | J=12 | 0.000625 | 0.010269 | 0.010894 | 0.001061 | 0.016958 | 0.018020 |
| | | J=30 | 0.000123 | 0.004143 | 0.004266 | 0.001623 | 0.027549 | 0.029171 |
| | N=1000 | J=12 | 0.000106 | 0.003351 | 0.003457 | 0.000453 | 0.007026 | 0.007479 |
| | | J=30 | 0.000058 | 0.001468 | 0.001525 | 0.000132 | 0.003291 | 0.003423 |
| | N=3000 | J=12 | 0.000190 | 0.000865 | 0.001055 | 0.000354 | 0.002759 | 0.003113 |
| | | J=30 | **0.000019** | **0.000488** | **0.000507** | 0.000111 | 0.001557 | 0.001668 |
| | | | ASB(s) | AVAR(s) | AMSE(s) | ASB(s) | AVAR(s) | AMSE(s) |
| S | N=300 | J=12 | 0.000805 | 0.003692 | 0.004497 | 0.000242 | 0.001582 | 0.001825 |
| | | J=30 | 0.000074 | 0.001152 | 0.001226 | 0.000123 | 0.001210 | 0.001333 |
| | N=1000 | J=12 | 0.000200 | 0.000944 | 0.001144 | 0.000099 | 0.000393 | 0.000492 |
| | | J=30 | 0.000018 | 0.000329 | 0.000347 | 0.000020 | 0.000331 | 0.000351 |
| | N=3000 | J=12 | 0.000140 | 0.000258 | 0.000398 | 0.000167 | 0.000116 | 0.000283 |
| | | J=30 | **0.000005** | 0.000113 | 0.000119 | 0.000011 | **0.000101** | **0.000112** |

Table 32. Differences of Fit Indices between the DINA(-H) and DINO(-H) Models for the Main Effect of Model Consistency

| DINO-DINA | MAIC | MBIC |
|---|---|---|
| HL_HL - HL_HL | -705 | -704 |
| HU_HU - HU_HU | -2491 | -2490 |
| NO_NO - NA_NA | -1034 | -1034 |

Table 33. Differences of Summary Statistics between the DINA(-H) and DINO(-H) Models for the Main Effect of Model Consistency

| DINO-DINA | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|
| HL_HL - HL_HL | -0.00141 | 0.00080 | -0.00060 | -0.00182 | -0.00088 | -0.00269 |
| HU_HU - HU_HU | -0.00084 | 0.01251 | 0.01167 | -0.00193 | -0.00199 | -0.00393 |
| NO_NO - NA_NA | -0.00083 | 0.00297 | 0.00215 | -0.00226 | -0.00275 | -0.00502 |

Table 34. Differences of Summary Statistics between the DINA(-H) and DINO(-H) Models for the Main Effect of Number of Attribute

| | DINO-DINA | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| K=6 | HL_HL - HL_HL | -0.0027 | 0.0005 | -0.0023 | -0.0036 | -0.0006 | -0.0042 |
| | HU_HU - HU_HU | -0.0024 | 0.0031 | 0.0007 | -0.0035 | -0.0010 | -0.0045 |
| | NO_NO - NA_NA | -0.0023 | 0.0020 | -0.0004 | -0.0039 | -0.0016 | -0.0055 |
| | **Mean** | -0.0025 | 0.0019 | -0.0007 | **-0.0037** | -0.0010 | -0.0047 |
| K=8 | HL_HL - HL_HL | -0.0001 | 0.0011 | 0.0011 | 0.0000 | -0.0012 | -0.0012 |
| | HU_HU - HU_HU | 0.0008 | 0.0219 | 0.0227 | -0.0003 | -0.0030 | -0.0034 |
| | NO_NO - NA_NA | 0.0007 | 0.0040 | 0.0047 | -0.0006 | -0.0040 | -0.0046 |
| | **Mean** | 0.0005 | **0.0090** | **0.0095** | -0.0003 | -0.0027 | -0.0030 |

Table 35. Differences of Summary Statistics between the DINA(-H) and DINO(-H) Models for the Main Effect of Test Length

| | DINO-DINA | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| J=12 | HL_HL - HL_HL | -0.0001 | 0.0010 | 0.0009 | 0.0000 | -0.0011 | -0.0011 |
| | HU_HU - HU_HU | 0.0003 | 0.0106 | 0.0109 | -0.0002 | -0.0025 | -0.0027 |
| | NO_NO - NA_NA | 0.0006 | 0.0038 | 0.0043 | -0.0006 | -0.0034 | -0.0039 |
| | **Mean** | 0.0002 | 0.0051 | **0.0053** | -0.0003 | -0.0023 | -0.0026 |
| J=30 | HL_HL - HL_HL | -0.0027 | 0.0006 | -0.0021 | -0.0036 | -0.0007 | -0.0043 |
| | HU_HU - HU_HU | -0.0020 | 0.0145 | 0.0125 | -0.0036 | -0.0015 | -0.0051 |
| | NO_NO - NA_NA | -0.0022 | 0.0022 | 0.0000 | -0.0040 | -0.0021 | -0.0061 |
| | **Mean** | -0.0023 | **0.0058** | 0.0035 | **-0.0037** | -0.0014 | -0.0052 |

Table 36. Differences of Summary Statistics between the DINA(-H) and DINO(-H) Models for the Main Effect of Sample Size

| | DINO-DINA | ASB(g) | AVAR(g) | AMSE(g) | ASB(s) | AVAR(s) | AMSE(s) |
|---|---|---|---|---|---|---|---|
| N=300 | HL_HL - HL_HL | -0.00403 | 0.00170 | -0.00233 | -0.00547 | -0.00194 | -0.00740 |
| | HU_HU - HU_HU | -0.00285 | 0.02875 | 0.02588 | -0.00583 | -0.00434 | -0.01018 |
| | NO_NO - NA_NA | -0.00283 | 0.00629 | 0.00346 | -0.00651 | -0.00560 | -0.01212 |
| | **Mean** | -0.00324 | **0.01224** | 0.00901 | **-0.00594** | -0.00396 | **-0.00990** |
| N=1000 | HL_HL - HL_HL | -0.00008 | 0.00052 | 0.00045 | 0.00000 | -0.00053 | -0.00052 |
| | HU_HU - HU_HU | 0.00020 | 0.00619 | 0.00639 | 0.00002 | -0.00120 | -0.00118 |
| | NO_NO - NA_NA | 0.00019 | 0.00198 | 0.00217 | -0.00017 | -0.00197 | -0.00213 |
| | **Mean** | 0.00010 | 0.00290 | 0.00300 | -0.00005 | -0.00123 | -0.00128 |
| N=3000 | HL_HL - HL_HL | -0.00012 | 0.00019 | 0.00007 | 0.00001 | -0.00015 | -0.00015 |
| | HU_HU - HU_HU | 0.00015 | 0.00259 | 0.00274 | 0.00000 | -0.00043 | -0.00043 |
| | NO_NO - NA_NA | 0.00017 | 0.00065 | 0.00081 | -0.00012 | -0.00071 | -0.00082 |
| | **Mean** | 0.00007 | 0.00114 | 0.00121 | -0.00003 | -0.00043 | -0.00047 |

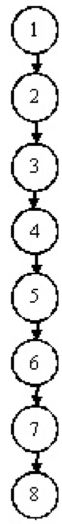


Figure 1: Linear Attribute Hierarchies of Eight Attributes

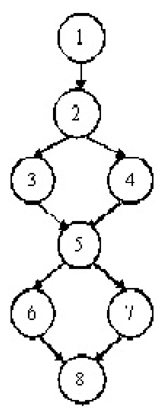


Figure 2: Convergent Attribute Hierarchies of Eight Attributes

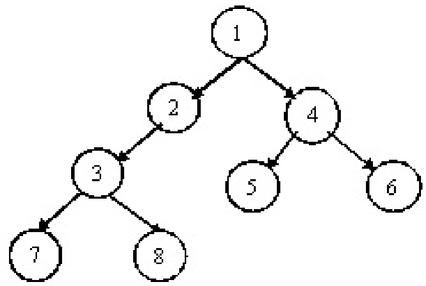


Figure 3: Divergent Attribute Hierarchies of Eight Attributes

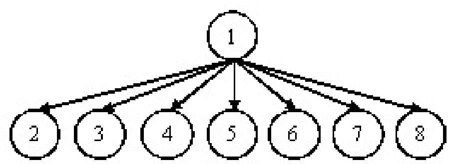


Figure 4: Unstructured Attribute Hierarchies of Eight Attributes

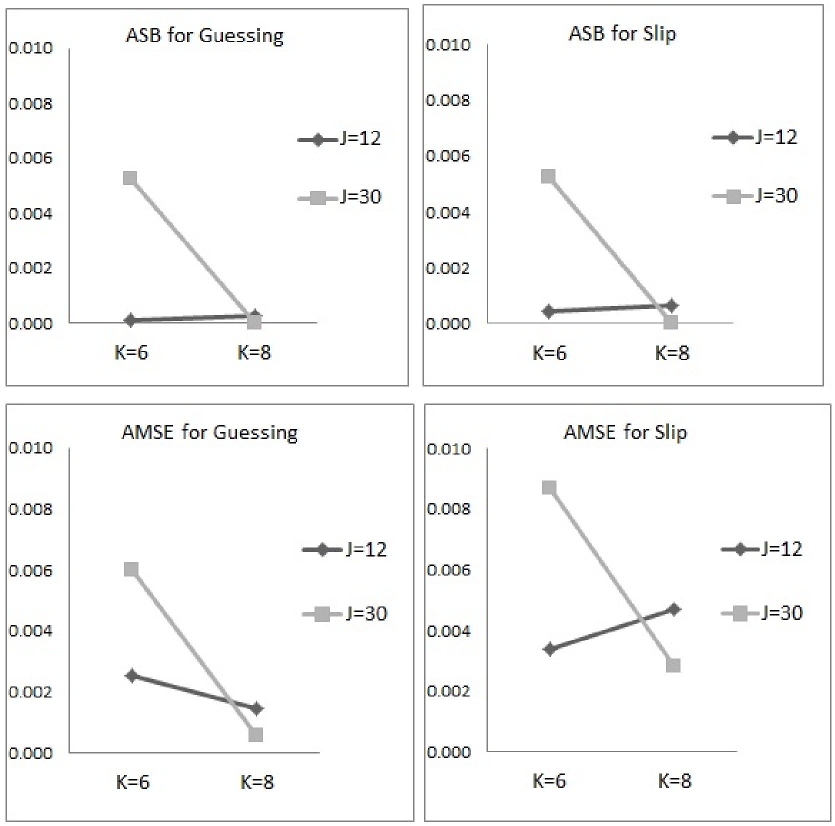


Figure 5: Summary statistics for the interaction effect of J by K for DINA(-H)

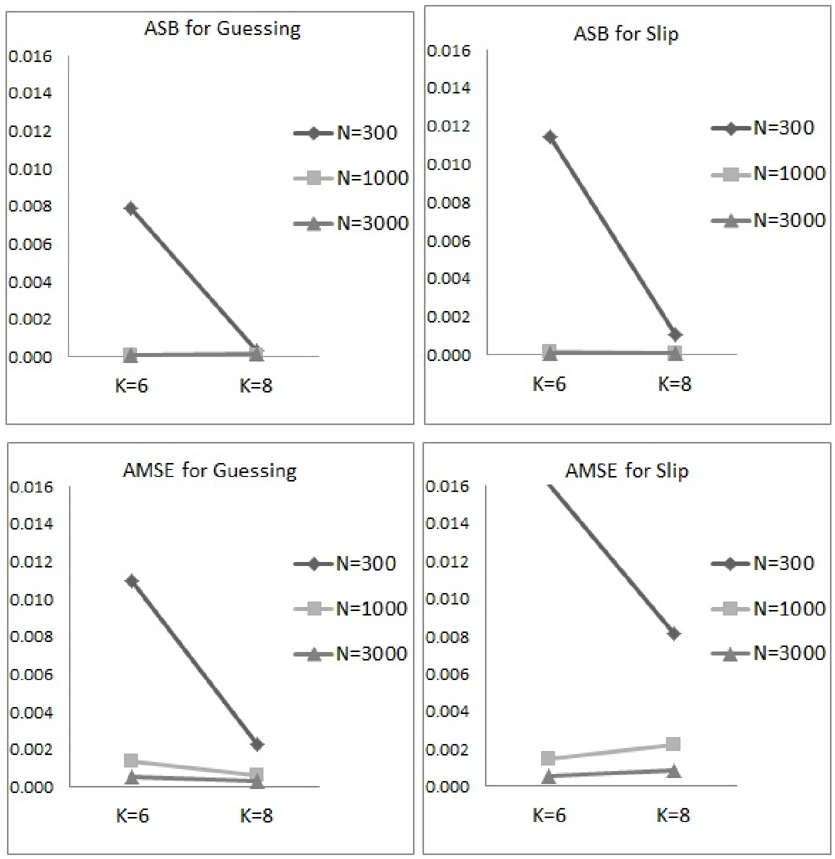


Figure 6: Summary statistics for the interaction effect of N by K for DINA(-H)

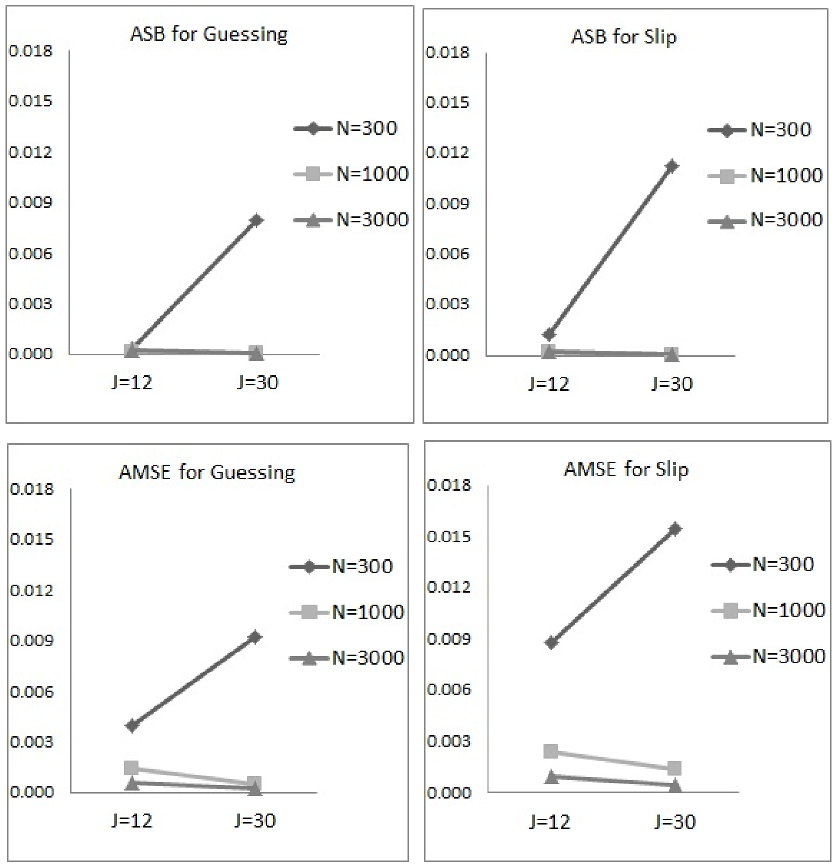


Figure 7: Summary statistics for the interaction effect of N by J for DINA(-H)

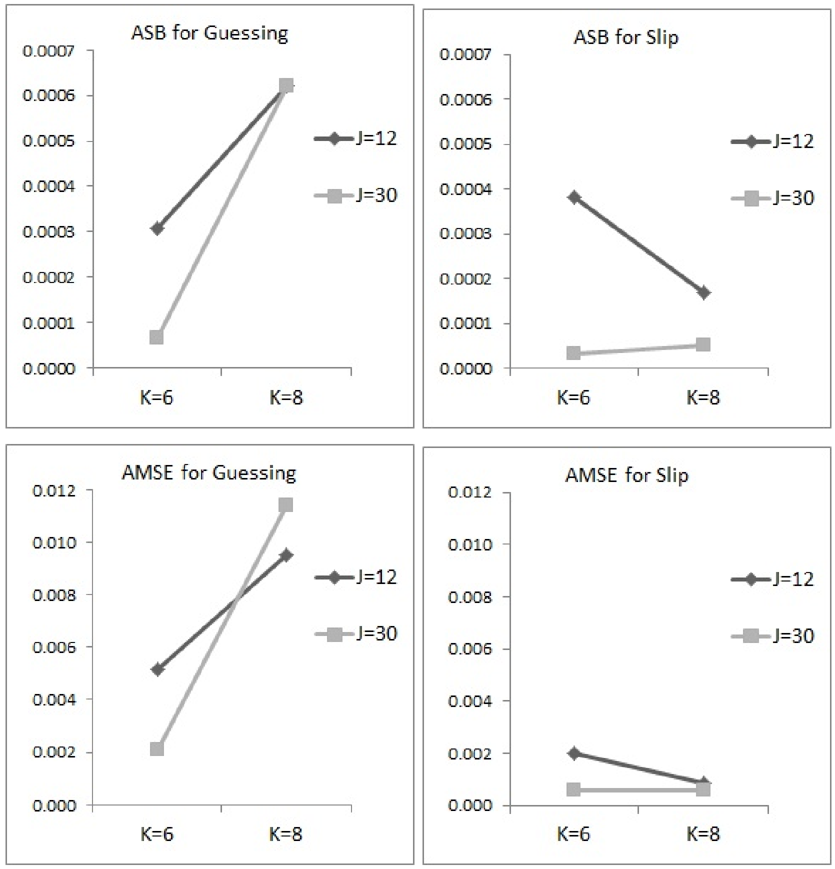


Figure 8: Summary statistics for the interaction effect of J by K for DINO(-H)

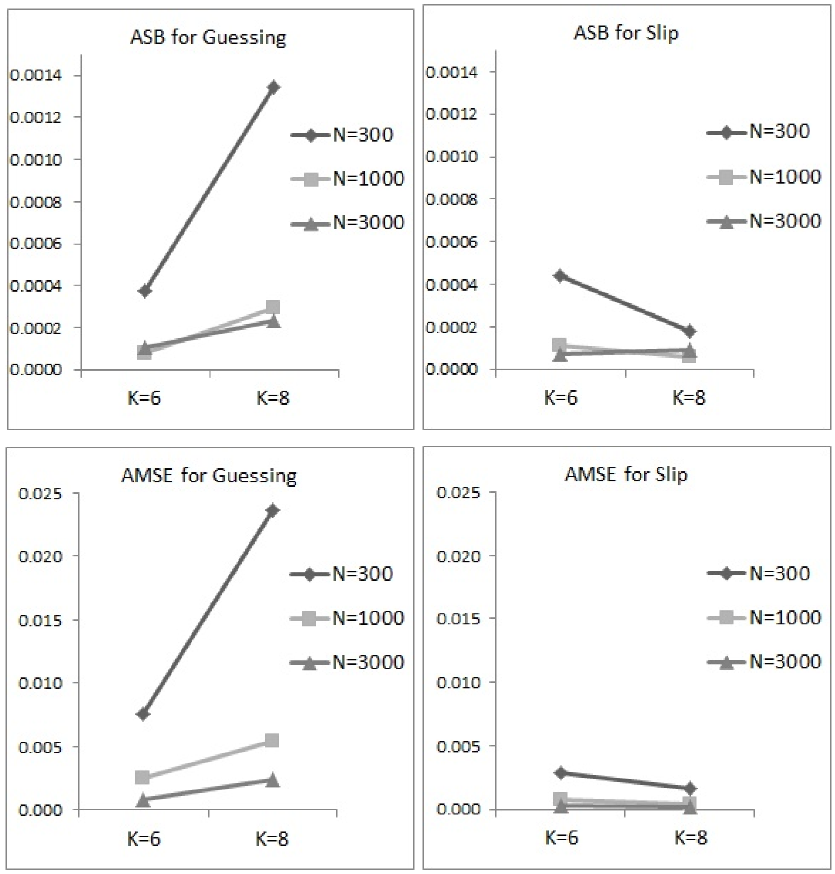


Figure 9: Summary statistics for the interaction effect of N by K for DINO(-H)

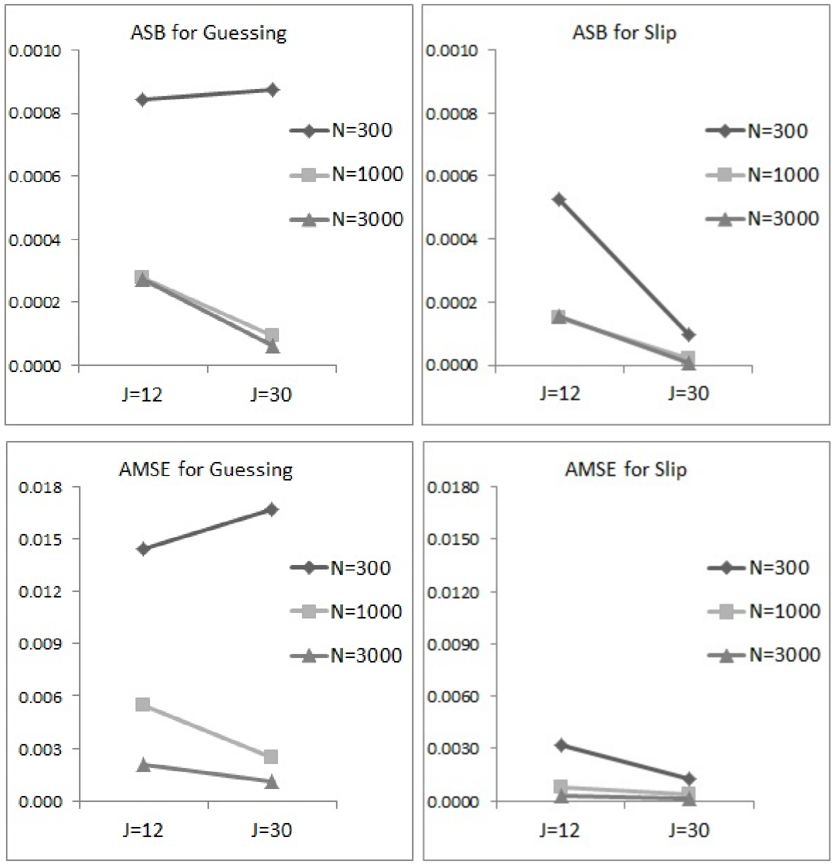


Figure 10: Summary statistics for the interaction effect of N by J for DINO(-H)