



# Cognitive Diagnosis Modeling Incorporating Item-Level Missing Data Mechanism

Na Shan<sup>1,2\*</sup> and Xiaofei Wang<sup>2,3</sup>

<sup>1</sup> School of Psychology, Northeast Normal University, Changchun, China, <sup>2</sup> Key Laboratory of Applied Statistics of the Ministry of Education, Northeast Normal University, Changchun, China, <sup>3</sup> School of Mathematics and Statistics, Northeast Normal University, Changchun, China

The aim of cognitive diagnosis is to classify respondents' mastery status of latent attributes from their responses on multiple items. Since respondents may answer some but not all items, item-level missing data often occur. Even if the primary interest is to provide diagnostic classification of respondents, misspecification of missing data mechanism may lead to biased conclusions. This paper proposes a joint cognitive diagnosis modeling of item responses and item-level missing data mechanism. A Bayesian Markov chain Monte Carlo (MCMC) method is developed for model parameter estimation. Our simulation studies examine the parameter recovery under different missing data mechanisms. The parameters could be recovered well with correct use of missing data mechanism for model fit, and missing that is not at random is less sensitive to incorrect use. The Program for International Student Assessment (PISA) 2015 computer-based mathematics data are applied to demonstrate the practical value of the proposed method.

**Keywords:** cognitive diagnosis, item-level, missing data, missing data mechanism, cognitive diagnosis model

## OPEN ACCESS

### Edited by:

Tao Xin,  
Beijing Normal University, China

### Reviewed by:

Pasquale Anselmi,  
University of Padua, Italy  
Wenchao Ma,  
University of Alabama, United States

### \*Correspondence:

Na Shan  
dann285@nenu.edu.cn

### Specialty section:

This article was submitted to  
Quantitative Psychology and  
Measurement,  
a section of the journal  
Frontiers in Psychology

**Received:** 22 May 2020

**Accepted:** 28 October 2020

**Published:** 30 November 2020

### Citation:

Shan N and Wang X (2020) Cognitive  
Diagnosis Modeling Incorporating  
Item-Level Missing Data Mechanism.  
*Front. Psychol.* 11:564707.  
doi: 10.3389/fpsyg.2020.564707

## 1. INTRODUCTION

Cognitive diagnosis has recently received increasing concern in psychological and educational assessment, which can provide fine-grained classifications and diagnostic feedback for respondents from their performance on test items (Leighton and Gierl, 2007; Rupp et al., 2010). It is a useful tool to identify students' mastery status of different latent skills based on their responses to test items and to evaluate patients' presence of mental disorders based on their responses to diagnostic questions. More specifically, cognitive diagnosis has been used to study fraction subtraction (de la Torre and Douglas, 2004), language proficiency (von Davier, 2008; Chiu and Köhn, 2016), psychological disorders (Templin and Henson, 2006; Peng et al., 2019), and so forth.

Various cognitive diagnosis models (CDMs), also called diagnosis classification models, have been developed, such as the deterministic inputs, noisy "and" gate (DINA) model (Macready and Dayton, 1977; Junker and Sijtsma, 2001) and the deterministic inputs, noisy "or" gate (DINO) model (Templin and Henson, 2006). Most CDMs are parametric and model the probability of item response as a function of latent attributes. The simplicity and interpretability make the parametric CDMs popular in practice. More general CDMs, such as the log-linear cognitive diagnosis model (LCDM; Henson et al., 2009) and the generalized DINA model (de la Torre, 2011), assume a more flexible relationship between the item responses and latent attributes. Moreover, higher-order latent trait models for cognitive diagnosis (de la Torre and Douglas, 2004) have been introduced to link the correlated latent traits by a general high-order ability.

Multiple items are often used for cognitive diagnosis. When respondents choose to answer some but not all items, item-level missing data occur (Chen et al., 2020). Respondents may refuse to answer items that they deem too difficult, quit the test early because it is too long, or just skip items because of carelessness. Missing data lead to loss of information and may result in biased conclusions (Glas and Pimentel, 2008; Köhler et al., 2015; Kuha et al., 2018). Many studies have employed a complete case analysis that only use subjects without missing data (e.g., Xu and Zhang, 2016; Chen et al., 2017; Zhan et al., 2019a). In this case, the subjects with missing data cannot receive any diagnostic feedback and, more importantly, may produce biased results when subjects with complete data are systematically different from those with missing data (Pan and Zhan, 2020).

In the literature, three missing data mechanisms should be distinguished: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) (Rubin, 1976; Little and Rubin, 2020). The MCAR holds if the probability of missingness is independent of both the observed and unobserved responses, whereas the MAR holds if the probability of missingness is independent of the unobserved responses given the observed responses. If either of these two conditions cannot be satisfied, i.e., the probability of missingness depends on the unobserved responses, the MNAR occurs. If the missing data mechanism is MCAR or MAR, unbiased estimation can be obtained from the observed data; if the missing data mechanism is MNAR, a model for the missing data mechanism should be included to obtain valid estimations of the primary parameters. Limited approaches have been proposed in CDMs incorporating missing data mechanisms. Ömür Sünbül (2018) considered MCAR and MAR in the DINA model. Recently, Ma et al. (2020) have used the sequential process model to accommodate omitted items due to MNAR, where the first internal task, i.e., making the decision to either skip the item or respond to it, is assumed to be affected by a latent categorical variable representing the response tendency. As stated by Heller et al. (2015), CDMs have connections to knowledge space theory (KST), which has been developed by Doignon and Falmagne (1985) (see also Doignon and Falmagne, 1999; Falmagne and Doignon, 2011), de Chiusole et al. (2015) and Anselmi et al. (2016) have developed models for the analysis of MCAR, MAR, and MNAR data in the framework of KST. In their work, the MCAR holds if the missing response pattern is independent of the individual's knowledge state (i.e., the collection of all items that an individual is capable of solving in a certain disciplinary domain) and of the observed responses; the MAR holds if the missing response pattern is conditionally independent of the knowledge state given the observed responses; and the MNAR holds if the missing response pattern depends on the knowledge state. In CDMs, the attribute profile (i.e., the collection of all attributes that an individual masters in a certain disciplinary domain) is similar to the knowledge state in KST. In our study, an additional latent categorical variable is introduced to indicate missingness propensity. This latent categorical variable affects the probability of missingness for each item and, in the meantime, may affect the latent attributes through the influence of the general high-order ability. The missing data mechanism can then

be incorporated into cognitive diagnosis. Similar ideas can be seen in Holman and Glas (2005), Rose et al. (2015), and Kuha et al. (2018).

In this paper, we propose a joint cognitive diagnosis modeling including a higher-order latent trait model for item responses and a missingness propensity model for item-level missing data mechanism. We take the DINA model as an example for illustration because of its popular use, and the latent traits are linked by a general high-order ability. For a flexible specification of the missingness propensity model, the latent missingness propensity is represented by a categorical variable. The MNAR holds if the distribution of the general high-order ability depends on the latent classes, whereas the MCAR holds if the distribution of the general high-order ability is independent of the latent classes.

The rest of this paper is organized as follows. Section 2 presents the proposed joint model for item responses and item-level missing data mechanism. The Bayesian approach is then developed for model parameter estimation using JAGS. In section 3, simulation studies are conducted to compare the performance of the parameter recovery under different missing data mechanisms. Real data analysis using the PISA 2015 computer-based mathematics data is given in section 4. Some concluding remarks are given in section 5.

## 2. JOINT MODELING INCORPORATING ITEM-LEVEL MISSING DATA MECHANISM

We consider  $N$  subjects taking a test of  $I$  items, and there are  $K$  latent attributes to be evaluated. Let  $Y_{ni}$  be the response for subject  $n$  ( $n = 1, \dots, N$ ) to item  $i$  ( $i = 1, \dots, I$ ). Let  $R_{ni}$  be a missingness indicator corresponding to  $Y_{ni}$ , where  $R_{ni} = 1$  if  $Y_{ni}$  is observed and  $R_{ni} = 0$  if  $Y_{ni}$  is missing.

### 2.1. The Missingness Propensity Model

Let  $\xi_n$  denote the latent missingness propensity for subject  $n$ .  $\xi_n$  is unobserved and has  $C$  categories, which we refer to as latent missingness classes. The missingness probability of  $(R_{n1}, \dots, R_{nI})$  is given by

$$p(R_{n1}, \dots, R_{nI}) = \sum_{c=1}^C \left\{ \prod_{i=1}^I p(R_{ni} | \xi_n = c) \right\} p(\xi_n = c), \quad (1)$$

where  $p(R_{n1}, \dots, R_{nI})$  is the joint probability of  $(R_{n1}, \dots, R_{nI})$ ,  $p(R_{ni} | \xi_n = c)$  is the conditional probability of  $R_{ni}$  given  $\xi_n = c$  and  $p(\xi_n = c)$  is the probability of  $\xi_n = c$ .

The latent missingness propensity  $\xi_n$  is specified as the categorical distribution

$$\xi_n \sim \text{Categorical}(\boldsymbol{\pi}), \quad (2)$$

where  $\boldsymbol{\pi} = (\pi_1, \dots, \pi_C)$  and  $\sum_{c=1}^C \pi_c = 1$ . The conditional probability of  $R_{ni}$  is specified as the logistic function

$$\text{logit}\{p(R_{ni} = 1 | \xi_n = c)\} = \tau_{0i} + \sum_{c=2}^C \tau_{ci} \xi_{(c)}, \quad (3)$$

where  $\tau_{0i}$  and  $\tau_{ci}(c = 2, \dots, C)$  are the intercept and slope parameters, and  $\xi_{(c)}(c = 2, \dots, C)$  are dummy variables for the missingness classes. A positive slope parameter  $\tau_{ci}$  is assumed, which means that the missingness probability reduces in the latter missingness class compared to the first class. Denote  $\boldsymbol{\tau}_0 = (\tau_{01}, \dots, \tau_{0I})$  and  $\boldsymbol{\tau}_c = (\tau_{c1}, \dots, \tau_{cI})$  for  $c = 2, \dots, C$ . The idea of introducing latent variables to model the non-response mechanism have been proposed previously by, for example, Lin et al. (2004) and Hafez et al. (2015). The missingness propensity model is identifiable when  $C \leq 2^I / (1 + I)$ .

### 2.2. The High-Order DINA Model

The DINA model describes the probability of item response as a function of latent attributes as follows:

$$p(Y_{ni} = 1) = g_i + (1 - s_i - g_i) \prod_{k=1}^K \alpha_{nk}^{q_{ik}}, \tag{4}$$

where  $p(Y_{ni} = 1)$  is the probability of a correct response for subject  $n$  to item  $i$ ;  $s_i$  and  $g_i$  are the slipping and guessing probability for item  $i$  respectively, and  $1 - s_i - g_i$  is the item discrimination index for item  $i$  (IDI<sub>*i*</sub>; de la Torre, 2008),  $\alpha_{nk}$  is the  $k$ th ( $k = 1, \dots, K$ ) latent attribute for subject  $n$ , with  $\alpha_{nk} = 1$  if subject  $n$  masters attribute  $k$  and  $\alpha_{nk} = 0$  otherwise. The  $Q$  matrix is a  $I \times K$  matrix with binary entries  $q_{ik}$  (Tatsuoka, 1983). For each  $i$  and  $k$ ,  $q_{ik} = 1$  indicates that attribute  $k$  is required to answer item  $i$  correctly and  $q_{ik} = 0$  otherwise.

Equation (4) can be reparameterized, and it is called the reparameterized DINA model (DeCarlo, 2011) as

$$\text{logit}(p(Y_{ni} = 1)) = \beta_i + \delta_i \prod_{k=1}^K \alpha_{nk}^{q_{ik}}, \tag{5}$$

where  $\beta_i = \text{logit}(g_i)$  and  $\delta_i = \text{logit}(1 - s_i) - \text{logit}(g_i)$  are called item intercept and interaction parameter, respectively.

As stated in the literature (de la Torre and Douglas, 2004; Zhan et al., 2018a), attributes in a test are often correlated, and a higher-order structure for the attributes can be formulated by

$$\text{logit}(p(\alpha_{nk} = 1)) = \gamma_k \theta_n - \lambda_k, \tag{6}$$

where  $p(\alpha_{nk} = 1)$  is the probability of subject  $n$ 's mastery of attribute  $k$ ;  $\theta_n$  is a general (higher-order) ability for subject  $n$ ; and  $\gamma_k$  and  $\lambda_k$  are the slope and intercept parameter for attribute  $k$ . Denote  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_K)$  and  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_K)$ . Following Zhan et al. (2018a), a positive slope parameter is assumed, which means the probability of mastery of attribute  $k$  increases as the general ability  $\theta_n$  grows. Including a higher-order structure for cognitive diagnosis can not only reduce the number of model parameters for correlated latent attributes but also obtain an assessment for subjects' overall ability.

The distribution of the general ability  $\theta_n$  may be affected by the latent missingness classes of  $\xi_n$ , and we suppose that

$$\theta_n | \xi_n = c \sim N(\mu_c, \sigma_c^2), \tag{7}$$

for  $n = 1, \dots, N$  and  $c = 1, \dots, C$ . If  $\mu_c$  or  $\sigma_c^2$  ( $c = 1, \dots, C$ ) may vary between different classes, the missing data mechanism

is MNAR, and we set  $\mu_1 = 0$  and  $\sigma_1^2 = 1$  for model identification. If  $\mu_c$  and  $\sigma_c^2$  remain unchanged for different  $c$ , the probability of missingness does not depend on the responses and MCAR holds, and this is where we set  $\theta_n \sim N(0, 1)$  for identification.

### 2.3. Bayesian Parameter Estimation

The parameters of the proposed model can be estimated using the Bayesian MCMC approach. JAGS (version 4.2.0; Plummer, 2015) and the R2jags package (Su and Yajima, 2020) in R (R Core Team, 2019) were used for estimation, and the JAGS code can be found in the **Supplementary Material**. The priors of the model parameters are given below. For the majority of the parameters, the conjugate priors are used. The priors and the hyper priors for the item parameters are assigned the same as those given in Zhan et al. (2018a), please find them for details. Moreover, the noninformative prior is used for Dirichlet distribution. The (truncated) normal distribution and the inverted gamma distribution priors are chosen to obtain dispersed values for each corresponding parameter.

$$(\pi_1, \dots, \pi_C) \sim \text{Dirichlet}(1, \dots, 1),$$

$$\mu_c \sim N(0, 4), \quad \sigma_c^2 \sim \text{InvGamma}(1, 1), \quad (c = 2, \dots, C),$$

$$\gamma_k \sim N(0, 4)I(\gamma_k > 0), \quad \lambda_k \sim N(0, 4), \quad (k = 1, \dots, K),$$

$$\tau_{0i} \sim N(0, 4), \quad \tau_{ci} \sim N(0, 4)I(\tau_{ci} > 0), \\ (i = 1, \dots, I; c = 2, \dots, C),$$

$$\begin{pmatrix} \beta_i \\ \delta_i \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_\beta \\ \mu_\delta \end{pmatrix}, \Sigma_{\text{item}} \right), \quad (i = 1, \dots, I).$$

The hyper priors are specified:

$$\mu_\beta \sim N(-2.197, 2), \quad \mu_\delta \sim N(4.394, 2)I(\mu_\delta > 0), \quad \Sigma_{\text{item}} \\ \sim \text{InvWishart}(I, 2),$$

where  $I$  is a  $2 \times 2$  identity matrix. In this case, the mean guessing and slipping probabilities are approximately equal to 0.1.

In this paper, the number of latent missingness classes  $C$  is taken as fixed. In fact,  $C$  can be selected by some information criterion, for example, deviance information criterion (DIC; Spiegelhalter et al., 2002), which can result in a statistically optimal number. In practice, it is efficient to determine  $C$  beforehand, using latent class analysis (Linzer and Lewis, 2011) just for the missingness indicators. In the following simulation studies, the number of latent missingness classes  $C$  is fixed to 2 for simplicity. For the values of  $C > 2$ , the results are similar and we report some results for  $C = 3$  in the **Supplementary Material**.

### 3. SIMULATION STUDY

Two simulation studies were conducted to evaluate the empirical performance of the proposed method. Simulation 1 aimed to examine the parameter recovery using the Bayesian MCMC





the case of low missingness;  $\mu_2 = 1.0$  and  $\sigma_2^2 = 0.5$ ;  $\gamma_k = 1.5$  for all attributes and  $\lambda = (-1.0, -0.5, 0.0, 0.5, 1.0)$ , indicating moderate correlations between attributes;  $\mu_{item} = (\mu_\beta, \mu_\delta) = (-2.197, 4.394)$  and

$$\Sigma_{item} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} 1.0 & -0.8 \\ -0.8 & 1.0 \end{pmatrix},$$

which was the mean vector and the covariance matrix for a bivariate normal distribution generating  $\beta_i$  and  $\delta_i$ . Other assignments for the model parameters, for example, equal probabilities for each latent class as those used in Ma et al. (2020), also make sense, and the results are similar to the above parameter settings.

In each of the eight conditions, models with MCAR and MNAR were fitted to the simulated data, respectively. Thirty replications were implemented for each fitted model. Pilot runs showed that the algorithm converged using 20,000 iterations with a burn-in phase of 10,000 iterations. The convergence of the chains was monitored by multivariate potential scale reduction factor, which were  $< 1.1$  (Gelman and Rubin, 1992). The bias and the root mean square error (RMSE) of the Bayesian estimates were computed to assess the parameter recovery. For evaluating the classification of each attribute and attribute profiles, the attribute correct classification rate (ACCR) and the pattern correct classification rate (PCCR) were computed. More formally, if the  $k$ th ( $k = 1, \dots, K$ ) latent attribute for subject  $n$  ( $n = 1, \dots, N$ ), denoted as  $\alpha_{k,n}$ , is estimated by  $\hat{\alpha}_{k,n}$ , ACCR for the  $k$ th latent attribute and PCCR can be expressed as  $ACCR = \sum_{n=1}^N I(\hat{\alpha}_{k,n} = \alpha_{k,n})/N$  and  $PCCR = \sum_{n=1}^N [\prod_{k=1}^K I(\hat{\alpha}_{k,n} = \alpha_{k,n})]/N$ , respectively. Here  $I(\hat{\alpha}_{k,n} = \alpha_{k,n})$  is an indicator function such that  $I(\hat{\alpha}_{k,n} = \alpha_{k,n}) = 1$  if  $\hat{\alpha}_{k,n} = \alpha_{k,n}$  and zero otherwise.

Figure 2 presents the recovery of the item mean vector and the item covariance matrix for the models with MCAR and MNAR in all eight conditions. First, the patterns of the item parameter recovery were similar between MCAR and MNAR, especially the RMSEs of MCAR and MNAR were close in each condition. The item mean vector can be well recovered, as the bias were small and the RMSEs were relatively smaller than those of the item covariance matrix. The bias of the item mean vector under MNAR were smaller than those under MCAR at a higher sample size  $N = 1,000$ . The RMSEs of the item mean vector decreased as test length increased. The bias and RMSEs of the item covariance matrix were relatively higher, and decreased as test length increased. The sample size had no consistent impact on the estimates of the item covariance matrix. Moreover, the missingness probabilities had little impact on the item parameter recovery.

Table 1 summarizes the item parameter recovery for models with MCAR and MNAR in high missingness conditions. The results for low missingness are similar and presented in Supplementary Figure 1. The mean absolute value (MAV) of the bias and RMSE are reported. In each condition, the MAVs of the bias and RMSE under MCAR and MNAR were close. Detailed information about the recovery of each item parameter with MCAR and MNAR was similar and not reported here.

TABLE 1 | Summary of the item parameters for high missingness conditions.

I	N	$\beta$				$\delta$			
		Bias		RMSE		Bias		RMSE	
		MAR	MNAR	MAR	MNAR	MAR	MNAR	MAR	MNAR
15	500	0.058	0.067	0.398	0.400	0.062	0.070	0.500	0.502
	1,000	0.056	0.047	0.355	0.355	0.063	0.057	0.407	0.407
30	500	0.053	0.054	0.319	0.318	0.061	0.059	0.406	0.405
	1,000	0.043	0.037	0.241	0.239	0.049	0.045	0.306	0.304

Table 2 summarizes the estimation of general ability parameter in high missingness conditions. The results for low missingness are similar and presented in Supplementary Figure 2. The MAV and Range of the bias and RMSE are reported. The correlation between the true and the estimated general abilities is also given. The MAVs of the bias under MCAR were much higher than those under MNAR, and the bias under MCAR were all negative, which can be seen from their Ranges. The MAVs of the RMSEs under MCAR were also higher than those under MNAR. The correlations under MNAR were higher than those under MAR. From the above results, we find that when data are generated with MNAR, incorrect use of missing data mechanism could lead to biased estimation of the general abilities.

Figures 3, 4 presents the recovery of higher-order parameters between the attributes and the general ability in each condition. For the attribute slope parameters, the bias was closer to zero and the RMSEs were relatively small. For the attribute intercept parameters, their recovery under MNAR were good with small bias and RMSEs. Under MCAR, the bias and RMSEs for attribute intercept parameters were large, with absolute values  $> 1.0$ , and all the bias were negative.

Figure 5 shows the correct classification rates for each attribute and attribute profiles in all conditions. All ACCRs were  $> 0.90$  and all PCCRs were  $> 0.70$ , which indicated a good recovery of the mastery status of attributes. ACCRs and PCCRs raised as test length increased and changed little as sample size increased. The correct classification rates of MCAR and MNAR were very close in each condition. These results may be caused by the fact that the impact of the missing data mechanism on the latent attributes is indirect through the general high-order ability, and we assume the same model for item response under both missing data mechanisms.

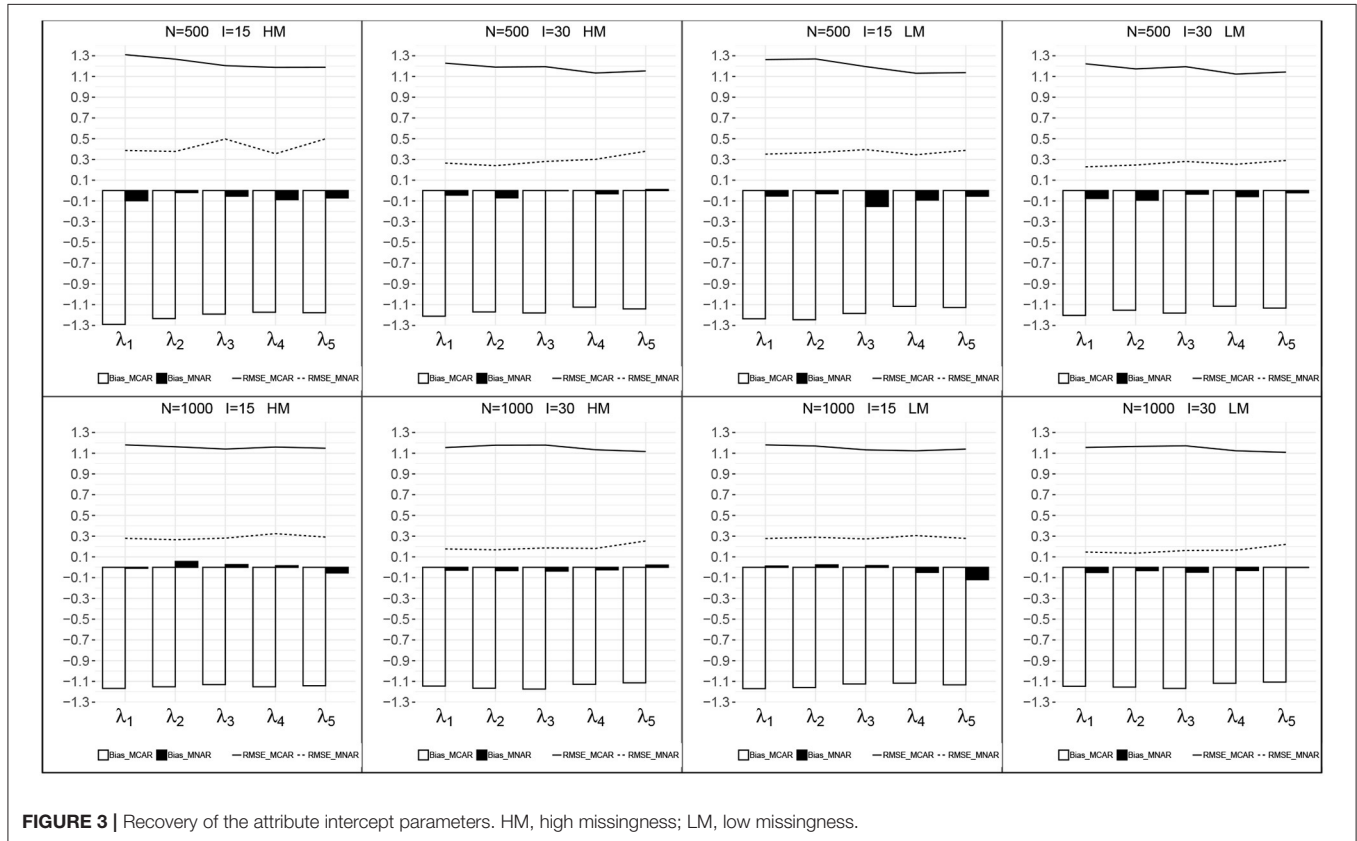
### 3.2. Simulation Study 2

The aim of Simulation 2 was to empirically examine the sensitivity of incorrect use of MNAR for model fit. In this study, the simulated data were generated with MCAR, i.e., the distribution of the general higher-order ability is independent of the latent missingness classes, and we set  $\theta_n \sim N(0, 1)$  for identification. The other settings were the same as those used for  $N = 500, J = 15$  and low missingness in Simulation 1, which is a weak condition studied in Simulation Study 1.

**TABLE 2** | Summary of the general ability for high missingness conditions.

I	N	Bias				RMSE				Cor	
		MAV		Range		MAV		Range		MCAR	MNAR
		MCAR	MNAR	MCAR	MNAR	MCAR	MNAR	MCAR	MNAR		
15	500	0.702	0.094	(-1.130, -0.311)	(-0.342, 0.387)	0.966	0.656	(0.680, 1.286)	(0.382, 1.012)	0.691	0.782
	1,000	0.699	0.095	(-1.036, -0.291)	(-0.283, 0.366)	0.965	0.647	(0.674, 1.335)	(0.347, 1.045)	0.690	0.728
30	500	0.707	0.090	(-1.066, -0.328)	(-0.283, 0.382)	0.943	0.610	(0.610, 1.272)	(0.403, 0.873)	0.730	0.766
	1,000	0.707	0.085	(-1.078, -0.344)	(-0.404, 0.323)	0.939	0.588	(0.649, 1.230)	(0.346, 0.899)	0.742	0.776

MAV, mean absolute value; Range = (minimum, maximum).



**FIGURE 3** | Recovery of the attribute intercept parameters. HM, high missingness; LM, low missingness.

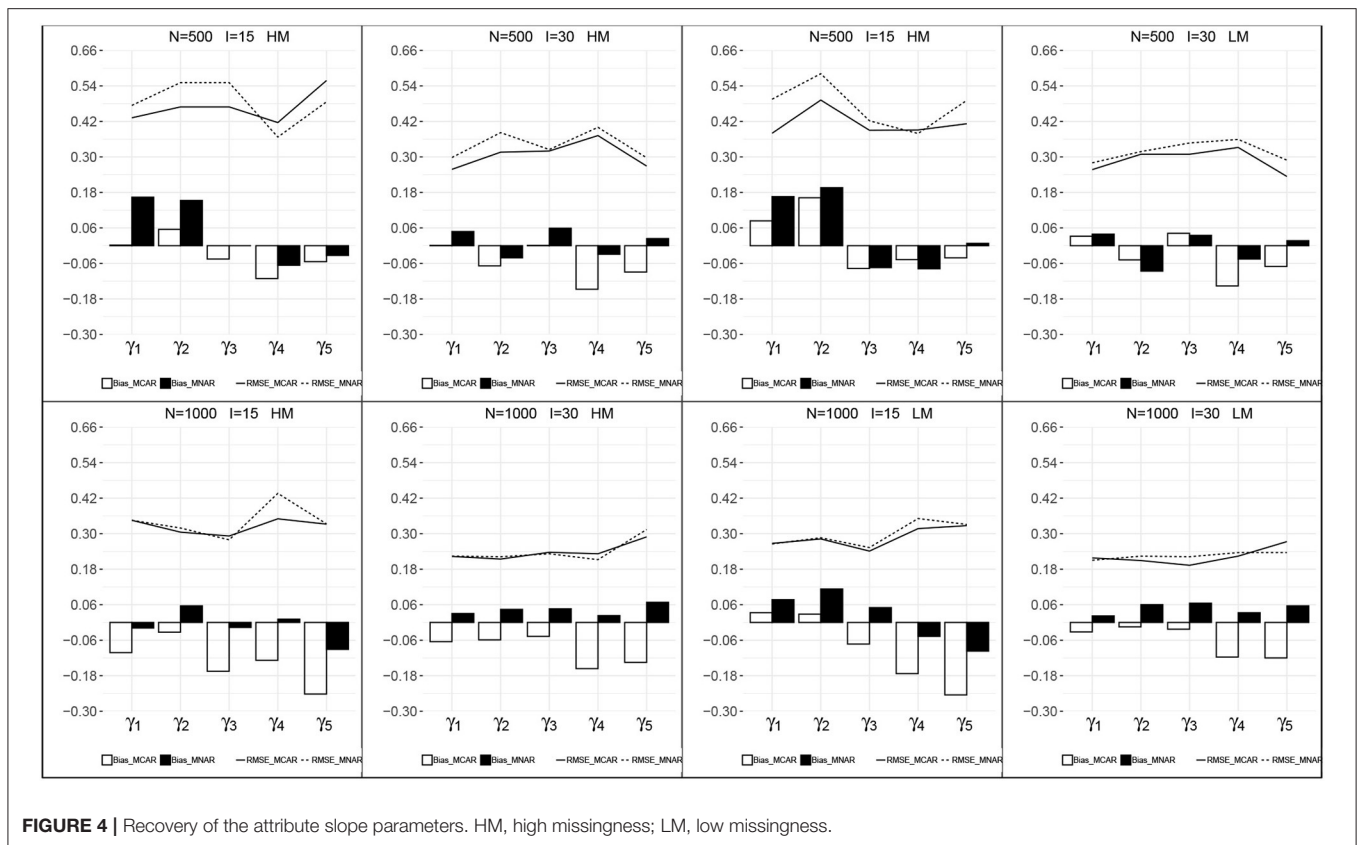
**Table 3** presents the bias and RMSE for the item mean vector, the item covariance matrix and the higher-order structure parameters. The recovery of the item parameter mean vector and covariance matrix were similar under MCAR and MNAR. For the higher-order structure parameters, unlike the results in Simulation Study 1, the recovery of the parameters under MNAR were as good as those under MCAR.

**Table 4** summarizes the recovery of the item and the general ability parameter, where the mean, standard deviation, minimum, and maximum of the bias and RMSE are reported. It also shows the correct classification rates for each attribute and attribute profiles. The results under MCAR and MNAR were similar, with mean bias close to zero and approximately equal correct classification rates for the recovery of each attribute and

attribute profile. From **Tables 3, 4**, we found that when the data were generated under MCAR, the missing data mechanism used in the model fit had little impact on the results for diagnostic classification in our framework.

### 4. REAL DATA ANALYSIS

To illustrate the application of the proposed method, the PISA 2015 computer-based mathematics data were used. The data include item scores and response times for each item, and we only select dichotomous item scores for illustration. Four attributes belonging to the mathematical content knowledge were evaluated, i.e., change and relationship ( $\alpha_1$ ), quantity ( $\alpha_2$ ),



**FIGURE 4 |** Recovery of the attribute slope parameters. HM, high missingness; LM, low missingness.

space and shape ( $\alpha_3$ ) and uncertainty and data ( $\alpha_4$ ).  $I = 9$  items with dichotomous responses were selected, with items IDs CM033Q01, CM474Q01, CM155Q01, CM155Q04, CM411Q01, CM411Q02, CM803Q01, CM442Q02, and CM034Q01. The Q matrix was shown in **Table 5**. Item responses with code 0 (no credit), code 1 (full credit), and code 9 (noresponse) were considered here. The responses “nonresponse” (code 9) were treated as missing data and might be due to a MNAR mechanism.  $N = 758$  test-takers from Albania with responses 0, 1, and 9 to each of the nine items were used for analysis. The missing rate (i.e., the proportion of “nonresponse”) for each item ranged from 0 to 14.78%. The models with missing data mechanisms MCAR and MNAR were both fitted to the data. The number of latent missingness classes  $C = 2$  was determined by latent class analysis. Then, the analysis was specified in the same way as the simulation study. The DIC was applied to compare model fit for models under different missing data mechanisms.

The DIC values under MCAR and MNAR were 11454.2 and 10406.3, respectively, which indicated that MNAR was preferred with a lower DIC value. We were only interested in the results concerning diagnostic classification. **Table 6** reports the estimated parameters and its corresponding standard deviations for the item mean vector, the item covariance matrix, and the attribute intercept and slope parameters. The results for the item mean vector and covariance matrix were similar under different missing data mechanisms, and the estimated  $\Sigma_{12}$  was  $-1.069$ , which indicated that items with a higher intercept corresponded

to a lower interaction.  $\mu_\beta$  was estimated to be  $-2.033$ , which means that the mean guessing probability was approximate 0.12. All estimated attribute intercept and slope parameters were positive. The estimations for the item covariance matrix and the attribute intercept and slope parameters were poor, which were consistent with previous studies (de la Torre and Douglas, 2004; Zhan et al., 2018a). **Table 7** reports the estimated item parameters. All  $\delta_i$  were positive, which means that all items satisfy  $g_i < 1 - s_i$ . Only  $\beta_i$  for CM033Q01 was positive, which means that the guessing probability of this item is higher than 0.5.

Though there were 16 possible attribute patterns for four attributes, 15 attribute patterns except (0101) were found in the estimated attributes under both MCAR and MNAR. **Figure 6** presents the top five most frequent attribute patterns in the real data under MCAR and MNAR. The most prevalent attribute pattern was (0000) and the second most prevalent pattern was (1111) under both MCAR and MNAR, where the corresponding proportions were slightly different. The third and the fourth most prevalent patterns under MCAR and MNAR were reverse. The above results indicate that the missing data mechanisms have some influences on the estimated attribute patterns.

## 5. CONCLUSIONS

When multiple items are used to classify subjects’ mastery status of latent attributes, it is almost inevitable that item-level

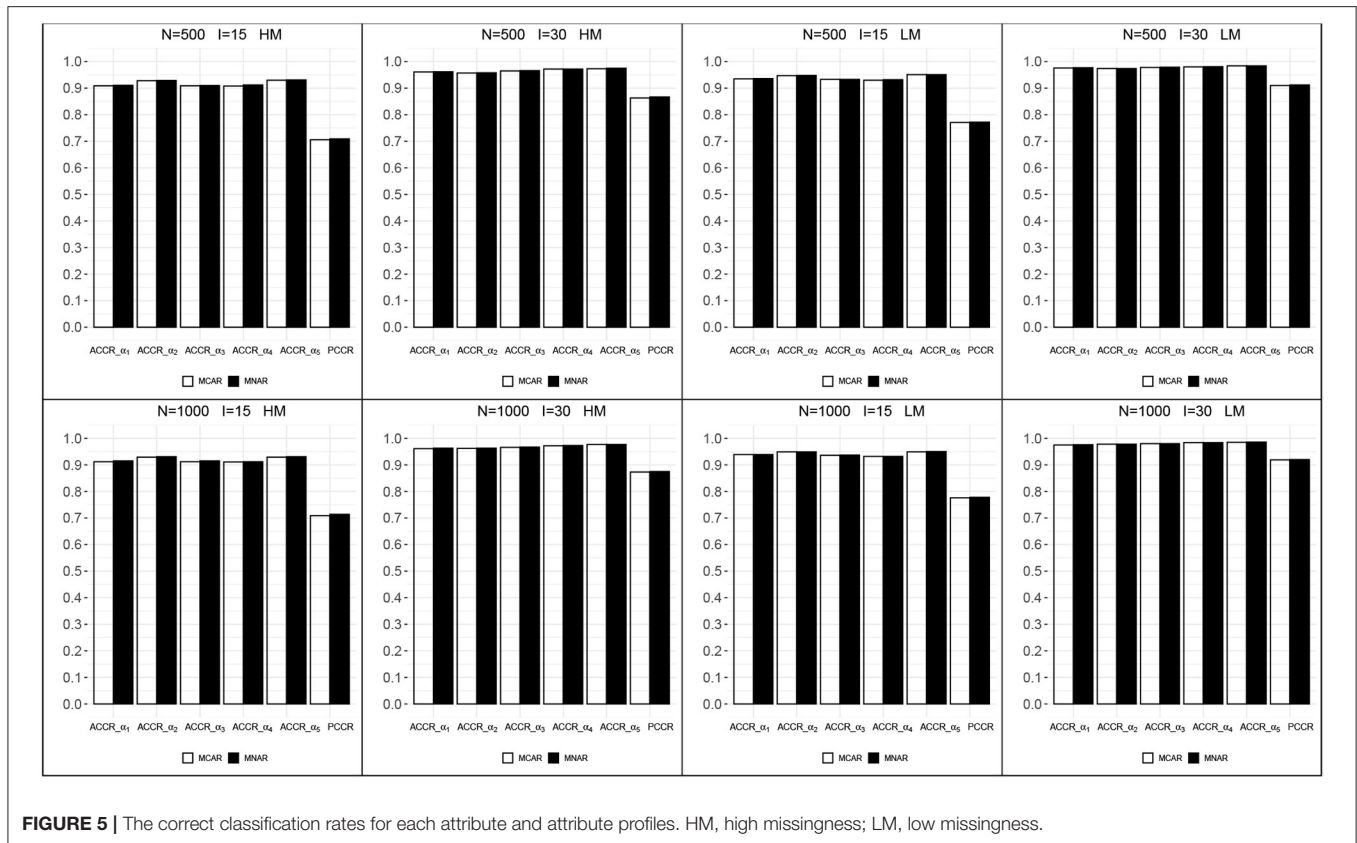


FIGURE 5 | The correct classification rates for each attribute and attribute profiles. HM, high missingness; LM, low missingness.

TABLE 3 | Recovery of the item mean vector, the item covariance matrix, the attribute slope, and intercept.

	Index	$\mu_\beta$	$\mu_\delta$	$\Sigma_{11}$	$\Sigma_{12}$	$\Sigma_{22}$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\gamma_1$	$\gamma_2$	$\gamma_3$	$\gamma_4$	$\gamma_5$
MAR	Bias	0.043	-0.005	0.080	0.028	0.158	0.001	-0.083	0.025	0.046	0.055	0.092	0.096	0.136	0.095	0.114
	RMSE	0.265	0.263	0.394	0.357	0.454	0.271	0.221	0.232	0.278	0.230	0.487	0.552	0.438	0.372	0.367
MNAR	Bias	0.045	-0.008	0.084	0.020	0.168	-0.052	-0.122	-0.028	0.019	0.024	0.070	0.062	0.104	0.067	0.095
	RMSE	0.263	0.261	0.399	0.368	0.469	0.346	0.280	0.299	0.342	0.317	0.494	0.624	0.462	0.384	0.386

missing data will occur. It is possible that the missing data mechanism is related to the item responses, and without very strong assumptions, item-level non-response can be thought to depend on some latent variables. Motivated by this idea, we have proposed a joint modeling method incorporating item responses and missing data mechanism for cognitive diagnosis. A latent categorical variable is employed to describe the latent missingness propensity, which can avoid distributional assumptions and result in a more flexible model. Then, the latent missingness classes are linked to each item missingness indicator by the logistic function. Applying the hierarchical modeling framework, the general higher-order ability's distribution is affected by the latent missingness class in the case of MNAR and is independent of the latent missingness class in the case of MCAR.

A Bayesian MCMC method is used to estimate the model parameters under the missing data mechanisms MCAR and MNAR. The simulation study demonstrates that when the data are generated under MNAR, the estimated general ability and

the attribute intercept parameters have higher bias if an incorrect missing data mechanism is used for model fit; when the data are generated under MCAR, the results between different missing data mechanisms do not have much difference. Similar results about the impact of the missing data mechanisms have been found by de Chiusole et al. (2015) in the framework of KST. The PISA 2015 computer-based mathematics data are used to explore the magnitude and direction of item and person parameters, and the results support the MNAR in the real data analysis.

Our proposed method can be further investigated in several aspects. First, the proposed model has good performance with DINA model used for illustration, and the joint modeling method can be extended to other types of CDMs for further studies. Second, this study assumes that each latent attribute is a binary variable. When polytomous attributes are involved in CDMs (Zhan et al., 2019b), modified higher-order CDMs could be utilized in the framework of our model. Third, multiple sources of data about a subject behavior, for example, response time and other process data, can be combined to build up a more



**TABLE 4 |** Summary of the item parameter, general ability, and attributes.

Index			$\beta$	$\delta$	$\theta$		$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$
MAR	Bias	Mean	0.036	0.054	0.100	ACCR	0.863	0.891	0.876	0.885	0.920
		SD	0.040	0.062	0.126	PCCR	0.595				
		Min.	-0.030	-0.077	-0.388						
		Max.	0.110	0.138	0.464						
	RMSE	Mean	0.343	0.510	0.683						
		SD	0.177	0.131	0.097						
		Min.	0.175	0.401	0.431						
		Max.	0.853	0.909	1.060						
Cor.			0.722								
MNAR	Bias	Mean	0.028	0.027	0.111	ACCR	0.864	0.891	0.875	0.885	0.921
		SD	0.045	0.067	0.129	PCCR	0.594				
		Min.	-0.041	-0.091	-0.425						
		Max.	0.112	0.141	0.476						
	RMSE	Mean	0.340	0.509	0.716						
		SD	0.171	0.124	0.097						
		Min.	0.171	0.399	0.462						
		Max.	0.823	0.871	1.049						
Cor.			0.720								

SD, standard deviation; Min., minimum; Max., maximum; Cor., correlation between true and estimated values.

**TABLE 5 |** The Q matrix in the real data.

Attribute	CM033Q01	CM474Q01	CM155Q01	CM155Q04	CM411Q01	CM411Q02	CM803Q01	CM442Q02	CM034Q01
$\alpha_1$	0	0	1	1	0	0	0	0	0
$\alpha_2$	1	0	0	0	0	0	0	0	1
$\alpha_3$	0	1	0	0	1	0	0	1	0
$\alpha_4$	0	0	0	0	0	1	1	0	0

**TABLE 6 |** Estimates and standard errors of the parameters for the real data.

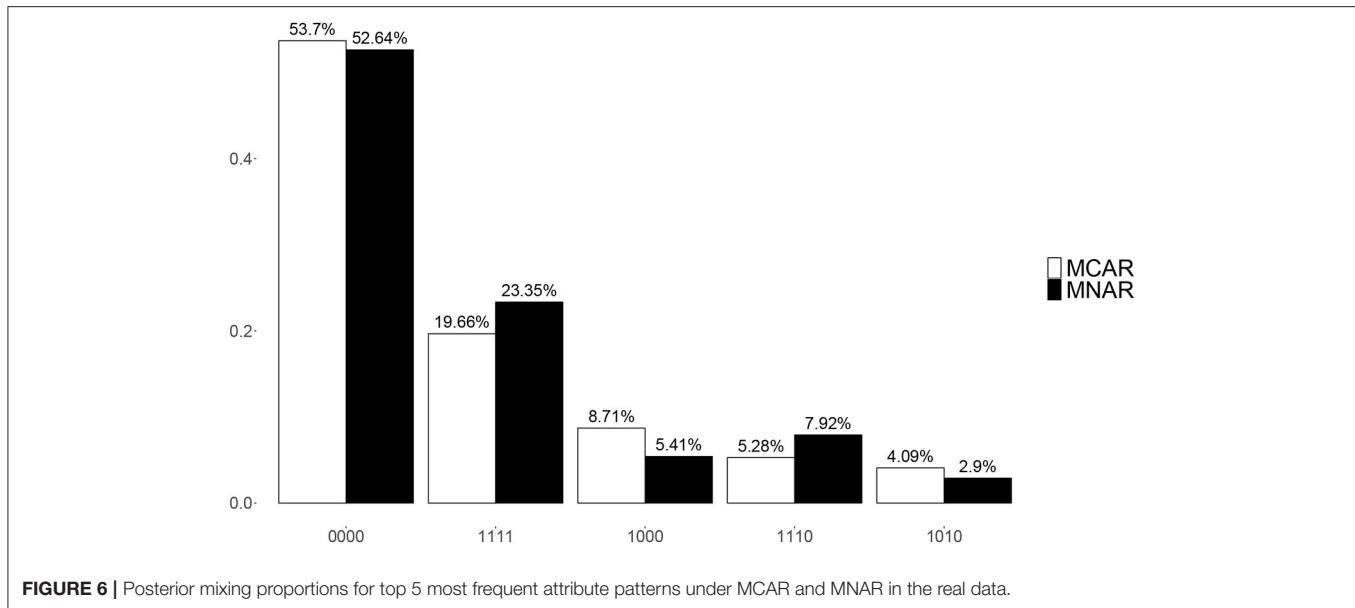
	DIC	Index	$\mu_\beta$	$\mu_\delta$	$\Sigma_{11}$	$\Sigma_{12}$	$\Sigma_{22}$	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\gamma_1$	$\gamma_2$	$\gamma_3$	$\gamma_4$
MCAR	11454.2	Est.	-2.033	2.548	3.383	-1.069	1.451	0.677	2.051	1.421	1.888	3.843	3.795	4.208	3.259
		SD	0.587	0.426	2.341	1.411	1.195	0.688	0.845	0.710	0.866	1.082	0.891	0.990	0.912
MNAR	10406.3	Est.	-2.073	2.419	3.313	-1.105	1.269	2.828	2.653	3.055	3.852	2.261	1.787	2.099	1.675
		SD	0.572	0.391	2.280	1.305	1.087	0.966	0.901	0.994	1.317	0.862	0.778	0.683	0.702

Est., Estimated values; SD, standard deviation of the posterior distribution.

**TABLE 7 |** Estimates and standard errors of the parameters for the real data.

	Par.	033Q01	474Q01	155Q01	155Q04	411Q01	411Q02	803Q01	442Q02	034Q01
MCAR	$\beta_i$	0.155 (0.139)	-0.501 (0.156)	-0.665 (0.289)	-1.445 (0.185)	-1.737 (0.256)	-1.355 (0.166)	-4.227 (0.811)	-4.587 (0.846)	-2.945 (0.412)
	$\delta_i$	2.447 (0.588)	1.589 (0.256)	3.342 (0.678)	1.588 (0.262)	2.508 (0.333)	0.788 (0.304)	3.458 (0.824)	3.219 (0.869)	2.775 (0.456)
MNAR	$\beta_i$	0.019 (0.151)	-0.579 (0.145)	-0.676 (0.245)	-1.452 (0.164)	-1.761 (0.244)	-1.321 (0.139)	-4.073 (0.795)	-4.543 (0.740)	-3.347 (0.575)
	$\delta_i$	2.081 (0.432)	1.655 (0.233)	3.068 (0.500)	1.588 (0.228)	2.316 (0.295)	0.733 (0.277)	3.308 (0.809)	3.042 (0.747)	2.881 (0.570)

Par., estimated parameter; the first two letters for item IDs (i.e., "CM") are omitted to save space.



general model, which could provide a more comprehensive reflection of individual behavior. Finally, the number of latent missingness classes can be varied. For selecting it, Akaike information criterion (AIC), Bayesian information criterion (BIC) or DIC, can be utilized to compare different models under various choices.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <http://www.oecd.org/pisa>.

## AUTHOR CONTRIBUTIONS

NS provided original idea of the paper. XW provided the technical support. Both authors contributed to the article and approved the submitted version.

## REFERENCES

- Anselmi, P., Robusto, E., Stefanutti, L., and de Chiusole, D. (2016). An upgrading procedure for adaptive assessment of knowledge. *Psychometrika* 81, 461–482. doi: 10.1007/s11336-016-9498-9
- Chen, L., Savalei, V., and Rhemtulla, M. (in press). Two-stage maximum likelihood approach for item-level missing data in regression. *Behav. Res. Methods*. doi: 10.3758/s13428-020-01355-x
- Chen, Y., Li, X., Liu, J., and Ying, Z. (2017). Regularized latent class analysis with application in cognitive diagnosis. *Psychometrika* 82, 660–692. doi: 10.1007/s11336-016-9545-6
- Chiu, C. Y., and Köhn, H. F. (2016). The reduced RUM as a logit model: parameterization and constraints. *Psychometrika* 81, 350–370. doi: 10.1007/s11336-015-9460-2
- de Chiusole, D., Stefanutti, L., Anselmi, P., and Robusto, E. (2015). Modeling missing data in knowledge space theory. *Psychol. Methods* 20, 506–522. doi: 10.1037/met0000050
- de la Torre, J. (2008). An empirically based method of Q-matrix validation for the DINA model: development and applications. *J. Educ. Meas.* 45, 343–362. doi: 10.1111/j.1745-3984.2008.00069.x
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika* 76, 179–199. doi: 10.1007/s11336-011-9207-7
- de la Torre, J., and Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika* 69, 333–353. doi: 10.1007/BF02295640
- DeCarlo, L. T. (2011). On the analysis of fraction subtraction data: the DINA model, classification, latent class sizes, and the q-matrix. *Appl. Psych. Meas.* 35, 8–26. doi: 10.1177/0146621610377081
- Doignon, J. P., and Falmagne, J. C. (1985). Spaces for the assessment of knowledge. *Int. J. Man Mach. Stud.*, 23, 175–196. doi: 10.1016/S0020-7373(85)80031-6
- Doignon, J. P., and Falmagne, J. C. (1999). *Knowledge Spaces*. Berlin; Heidelberg; New York, NY: Springer-Verlag.
- Falmagne, J. C., and Doignon, J. P. (2011). *Learning Spaces: Interdisciplinary Applied Mathematics*. Berlin; Heidelberg: Springer.
- Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. doi: 10.1214/ss/1177011136

## FUNDING

The research of NS was supported in part by the National Natural Science Foundation of China (grant number 11871013) and the Project of the Educational Department of Jilin Province of China (grant number 2016315). The research of XW was supported in part by the Fundamental Research Funds for the Central Universities (grant number 2412019FZ030) and Jilin Provincial Science and Technology Development Plan funded Project (grant number 20180520026JH).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.564707/full#supplementary-material>

- Glas, C. A. W., and Pimentel, J. L. (2008). Modeling nonignorable missing data in speeded tests. *Educ. Psychol. Meas.* 68, 907–922. doi: 10.1177/0013164408315262
- Hafez, M. S., Moustaki, I., and Kuha, J. (2015) Analysis of multivariate longitudinal data subject to nonrandom dropout. *Struct. Equat. Model.* 22, 193–201. doi: 10.1080/10705511.2014.936086
- Heller, J., Stefanutti, L., Anselmi, P., and Robusto, E. (2015). On the link between cognitive diagnostic models and knowledge space theory. *Psychometrika* 80, 995–1019. doi: 10.1007/s11336-015-9457-x
- Henson, R. A., Templin, J. L., and Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika* 74, 191–210. doi: 10.1007/s11336-008-9089-5
- Holman, R., and Glas, C. A. W. (2005). Modelling non-ignorable missing-data mechanisms with item response theory models. *Br. J. Math. Stat. Psychol.* 58, 1–17. doi: 10.1348/000711005X47168
- Junker, B. W., and Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Appl. Psychol. Meas.* 25, 258–272. doi: 10.1177/01466210122032064
- Köhler, C., Pohl, S., and Carstensen, C. H. (2015). Taking the missing propensity into account when estimating competence scores: evaluation of item response theory models for nonignorable omissions. *Educ. Psychol. Meas.* 75, 850–874. doi: 10.1177/0013164414561785
- Kuha, J., Katsikatsou, M., and Moustaki, I. (2018). Latent variable modelling with non-ignorable item non-response: multigroup response propensity models for cross-national analysis. *J. R. Stat. Soc. A* 181, 1169–1192. doi: 10.1111/rssa.12350
- Leighton, J., and Gierl, M. (2007). *Cognitive Diagnostic Assessment for Education: Theory and Applications*. Cambridge, UK: Cambridge University Press.
- Lin, H., McCulloch, C. E., and Rosenheck, R. A. (2004). Latent pattern mixture models for informative intermittent missing data in longitudinal studies. *Biometrics* 60, 295–305. doi: 10.1111/j.0006-341X.2004.00173.x
- Linzer, D. A., and Lewis, J. B. (2011). poLCA: an R package for polytomous variable latent class analysis. *J. Stat. Softw.* 42, 1–29. doi: 10.18637/jss.v042.i10
- Little, R. J., and Rubin, D. B. (2020). *Statistical Analysis With Missing Data, 3rd Edn*. New York, NY: John Wiley and Sons, Inc.
- Ma, W., Jiang, Z., and Schumacker, R. E. (2020). “Modeling omitted items in cognitive diagnosis models [Roundtable Session],” in *AERA Annual Meeting* (San Francisco, CA).
- Macready, G. B., and Dayton, C. M. (1977). The use of probabilistic models in the assessment of mastery. *J. Educ. Behav. Stat.* 2, 99–120. doi: 10.3102/10769986002002099
- Ömür Sünbül, S. (2018). The impact of different missing data handling methods on DINA model. *Int. J. Eval. Res. Educ.* 7, 77–86. doi: 10.11591/ijere.v1i1.11682
- Pan, Y., and Zhan, P. (2020). The impact of sample attrition on longitudinal learning diagnosis: a prologue. *Front. Psychol.* 11:1051. doi: 10.3389/fpsyg.2020.01051
- Peng, S., Wang, D., Gao, X., Cai, Y., and Tu, D. (2019). The CDA-BPD: retrofitting a traditional borderline personality questionnaire under the cognitive diagnosis model framework. *J. Pac. Rim. Psychol.* 13:e22. doi: 10.1017/prp.2019.14
- Plummer, M. (2015). *JAGS Version 4.0.0 User Manual*. Lyon. Retrieved from: <https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/>.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rose, N., von Davier, M., and Nagengast, B. (2015). Commonalities and differences in IRT-based methods for nonignorable item nonresponses. *Psychol. Test Assess. Model.* 57, 472–498.
- Rubin, D. B. (1976). Inference and missing data (with discussion). *Biometrika* 63, 581–592. doi: 10.1093/biomet/63.3.581
- Rupp, A., Templin, J., and Henson, R. (2010) *Diagnostic Measurement: Theory, Methods, and Applications*. New York, NY: Guilford Press.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* 64, 583–639. doi: 10.1111/1467-9868.00353
- Su, Y. S., and Yajima, M. (2020). *R2jags: Using R to Run 'JAGS'*. R package version 0.6-1. Retrieved from: <https://CRAN.R-project.org/package=R2jags>.
- Tatsuoka, K. K. (1983). Rule space: an approach for dealing with misconceptions based on item response theory. *J. Educ. Meas.* 20, 345–354. doi: 10.1111/j.1745-3984.1983.tb00212.x
- Templin, J. L., and Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychol. Methods* 11, 287–305. doi: 10.1037/1082-989X.11.3.287
- von Davier, M. (2008). A general diagnostic model applied to language testing data. *Br. J. Math. Stat. Psychol.* 61, 287–301. doi: 10.1348/000711007X193957
- Xu, G., and Zhang, S. (2016). Identifiability of diagnostic classification models. *Psychometrika* 81, 625–649. doi: 10.1007/s11336-015-9471-z
- Zhan, P., Jiao, H., and Liao, D. (2018a). Cognitive diagnosis modelling incorporating item response times. *Br. J. Math. Stat. Psychol.* 71, 262–286. doi: 10.1111/bmsp.12114
- Zhan, P., Jiao, H., Liao, D., and Li, F. (2019a). A longitudinal higher-order diagnostic classification model. *J. Educ. Behav. Stat.* 44, 251–281. doi: 10.3102/1076998619827593
- Zhan, P., Jiao, H., Liao, M., and Bian, Y. (2018b). Bayesian DINA modeling incorporating within-item characteristic dependency. *Appl. Psych. Meas.* 43, 143–158. doi: 10.1177/0146621618781594
- Zhan, P., Wang, W. C., and Li, X. (2019b). A partial mastery, higher-order latent structural model for polytomous attributes in cognitive diagnostic assessments. *J. Classif.* 37, 328–351. doi: 10.1007/s00357-019-09323-7

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Shan and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.