Journal of Mathematics and Computing Science, 2021,
Volume 7, No 2, 41-48

Bias in Person Parameter Estimates: The Maximum Likelihood Approach to Rasch Rating Scale Model Against Skewed Distributions

Nurul Hafizah Azizan[1]*, Zamalia Mahmud[2] and Adzhar Rambli[3]

[123]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan Kampus Kota Bharu, 15050 Kota Bharu, Kelantan, Malaysia

*hafizahaz@uitm.edu.my

Abstract: The conventional approach of parameter estimation technique, such as maximum likelihood estimation (MLE), can be negatively affected by the skewed distributions of the data. Consequently, estimates of the parameters in the model produced by the MLE in this condition are more likely to be biased. This article explores the biases in the Rasch rating scale person estimates while using the MLE approach against skewed distributions. The Markov Chain Monte Carlo (MCMC) simulation analysis was carried out with 1000 iterations based on 126 simulation conditions. These simulation conditions were formed using three criteria, which are the number of sample sizes, the number of items, and the type of distributions (i.e., standard normal distribution and skew-normal distribution). The bias in estimation was calculated based on the mean squared difference between the estimated values and actual values of the person parameter. Overall, the findings obtained from the simulation analysis proved that, in skewed distributions, the MLE approach is prone to produce biased person estimates, and the results are getting worse in small sample sizes. Thus, the MLE approach is strongly not recommended when estimating person parameters in Rasch rating scale model (RRSM) under skewed distributions, especially if the sample size is too small.

Keywords: Bias, Maximum likelihood estimation (MLE), Rasch rating scale model (RRSM),

1 Introduction

Normality assumption is an essential requirement when estimating the unknown parameters in Rasch measurement model (RMM) and item response theory (IRT). Unfortunately, this assumption may be violated in most real studies. Indeed, violating this assumption can lead to biased parameter estimates. It is well known that normality assumption has also become a crucial prerequisite for several statistical analyses. Generally, the MLE approach is among the classical parameter estimation techniques widely used for estimating unknown parameters not only in RMM and IRT but also in several statistical models over the years. However, the problem with MLE approach is that it does not work well with non-normal distributions. Under this condition, the MLE approach is often more likely to produce biased parameter estimates with large errors (e.g., root mean square error). Consequently, it will lead to misleading inferences of the analyses.

As shown in previous studies, the maximum likelihood estimator is prone to produce a significantly large standard error when estimating both item and person parameters of the two-parameter logistic model under an extremely skewed simulated dataset [1]. Likewise, a simulation analysis conducted with the dichotomous Rasch model has also revealed the same outcome, where the root mean square error was larger in the simulated dataset drawn from the skewed distributions than those from the normal distribution [2]. In line with these findings, a more recent study by Do [3] has also discovered that sample size and distribution of the data had some effect on the estimation of the person parameter of the four-parameter unidimensional binary IRT model, where, on average, the maximum likelihood estimator yielded biased person estimates with large root mean square error due to the violation of the normality assumption and the results also become worse in the case of skewed distributions.

In addition, the empirical findings by Finch and Edwards [2] have demonstrated that the degree of biases in estimates was also affected by the degree of skewness. According to the findings of this study, the greater degree of skewness, the greater the bias of the estimates would be. In a recent study, researchers also argue that a highly skewed dataset is more likely to result in biased parameter estimates of the IRT models [4]. Both studies revealed that the performance of the maximum likelihood estimator is also significantly influenced by the level of skewness. It is clearly shown that skewed distributions have become a serious issue that should not be taken lightly by psychometricians and researchers. Thus, the estimation of the unknown parameters with severely skewed distributions should be carried out with more caution.

Although the poor performance of the MLE approach that is linked with the severity of skewed distribution has been successfully demonstrated with the dichotomous Rasch model and other IRT models (i.e., two-parameter logistic model and four-parameter logistic model), unfortunately, the empirical findings with the Rasch rating scale model (RRSM) are still lacking. Therefore, this article focuses on the performance of the MLE approach when estimating person parameters in RRSM. In particular, the main goal of this article is to address the impact of the skewness level on the bias in the estimate of person parameters produced by the MLE approach. Findings from the study will contribute to the existing literature on how the degree of skewness affects the performance of the MLE approach while estimating person parameters in RRSM.

2 Methodology

A Rasch Rating Scale Model

In this article, the RRSM proposed by Andrich [5] is re-expressed as suggested by Andersen [6]. Let the person, the item, and the threshold parameters denoted as ρ_i, δ_j and, τ_k, respectively. Hence, the probability of the person i with ability ρ_i to obtain a score of k ($k = 0, ..., m$) for the item j having difficulty δ_j and threshold τ_k is given as in equation (1).

$$P\left(\gamma_i, \delta_j + \tau_k\right) = \frac{exp \sum_{k=0}^{m} \left[\left(\delta_j + \tau_k\right)\right]}{\sum_{r=0}^{m} exp \sum_{k=0}^{m} \left[\left(\delta_j + \tau_k\right)\right]} \quad (1)$$

where based on this model, item j has 0 up to m categories, and k is the count of the number of successfully completed categories for that item. In this model, the count of gaps within text j that person i filled in correctly is represented by the subscript k. Mathematically, the likelihood function for the model given in equation (1) is derived as an equation (2).

$$L = \frac{exp \sum_{i}^{N} \sum_{j}^{n} \sum_{k=0}^{x_{ni}} \left[\left(\delta_j + \tau_k\right)\right]}{\prod_{i}^{N} \prod_{j}^{n} \left[\sum_{k=0}^{m} exp \sum_{j=0}^{k} \left[\left(\delta_j + \tau_k\right)\right]\right]} \quad (2)$$

As highlighted by Engelhard Jr [7], traditionally, the unknown parameters in the Rasch model are estimated by using the classical approach known as maximum likelihood estimation (MLE) techniques (i.e. joint maximum likelihood estimation, conditional maximum likelihood estimation and marginal maximum likelihood estimation). Although the MLE approach is among the most widely used to estimate unknown parameters in several statistical models, it is, however, when the assumption of the normality is violated, the estimation of these unknown parameters yielded by the MLE approach tends to be biased [2]. A previous study conducted with the dichotomous Rasch model revealed that the marginal maximum likelihood estimation had caused biases in estimates of unknown parameters under non-normal distribution, and the results become considerably more severe in datasets with skewed distributions [8]. Thus, in this article, the performance of the MLE approach when estimating person parameters in RRSM is addressed. Specifically, this article only focuses on the bias in estimating person parameters produced by MLE under skewed distributions.

B Data Generation, Parameter Estimation, Comparison and Simulation Analyses

The issue addressed in this article was investigated using Markov Chain Monte Carlo (MCMC) simulation analysis. Although there are several software (i.e. SAS, S-Plus, STATA, and python) that can be used to run the simulation analysis, however, the R programming software was found to be more convenient as it also provides several valuable packages to fit the Rasch model (i.e. TAM, mixRasch, sirt, eRm and psychotools). In the first stage of the MCMC simulation analysis, the simulated or artificial survey data was generated according to the RRSM with a 6-point Likert scale. This simulated survey data was executed randomly based on the standard normal and skew-normal distributions proposed in previous studies [9], [10]. The rnorm() function and the sn() function were used to generate data from the standard normal distribution and skew-normal distribution, respectively. In this study, several skewness values (i.e. 1, 3, 5, -1, -3, -5) were considered. The simulated survey data was also generated based on another two criteria, which are three sample sizes (*N* = 100, 50, 30) and six numbers of items (*n* = 5, 10, 15, 20, 25, 30). From that, a total of 126 simulation conditions (7 distributions of the data x 3 sample sizes x 6 number of items) are formed, as summarized in Table 1.

Table 1: Summary of Simulation Conditions for Standard Normal (Mean = 0, SD = 1) and Skewness (SN = 1, 3, 5, -1, -3, -5).

n	N	n	N	n	N	n	N	n	N	n	N
5	100	10	100	15	100	20	100	25	100	30	100
5	50	10	50	15	50	20	50	25	50	30	50
5	30	10	30	15	30	20	30	25	30	30	30

n = Test length/ Number of items

N = Number of respondents/Sample sizes

After generating the simulated survey data, the maximum likelihood estimates of the person parameter were then obtained using the TAM package. In particular, the performance of the MLE approach in this study was investigated based on bias statistics; which was calculated using the mean squared difference between the estimated values and actual values of the person parameter as given in the equation (3):

$$Bias = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\rho}_i - \rho_i\right)^2 \tag{3}$$

In this equation, $\hat{\rho}_i$ and ρ_i denote the estimate and the true values of the person parameter, respectively. Whereas the number of sample sizes is represented by the *N*. As a general rule of thumb, the larger the values of this bias statistics imply the worse the parameter estimates. The simulation process is then repeated up to 1000 iterations for each of the 126 simulation conditions to fulfil the convergence criteria needed by the MCMC analysis. This is to ensure that a stationary distribution of the estimates is satisfied before any inferences are drawn.

3 Results and Analyses

This study aims to investigate the impact of violation of the normality assumption (i.e. skewed distributions) on the bias in estimates of person parameters of Rasch rating scale model (RRSM), which is estimated using the MLE approach. In particular, various skewness values were used to examine the severity of the effect of skewed distributions on bias in person estimates. At the same time, the standard normal distribution was also employed as a benchmark for comparison purposes. As mentioned in the previous section, findings of the simulation analysis were then compared based on the bias statistics given in the equation (3). This statistics shows which simulation condition produced more biased person estimates based on the rule of thumb that is previously explained.

Before proceeding to the comparison analysis, MCMC convergence diagnostic was performed to verify whether Markov Chain with the 1000 iterations has achieved stationarity of Monte Carlo estimates of person parameter. Stationary condition is one of the most fundamental criteria required by MCMC simulation analysis to ensure that stable estimates have been reached before any inferences are drawn. Trace plot appears to be the most frequently used graphical technique for MCMC convergence diagnostics. Figure 1 depicts a combination of four different trace plots for the convergence diagnostic of the MCMC simulation analysis performed in this study. Clearly, the trace plots here show that the person estimates produced by the MCMC simulation analysis have already converged to the stationary distribution with 1000 iterations.

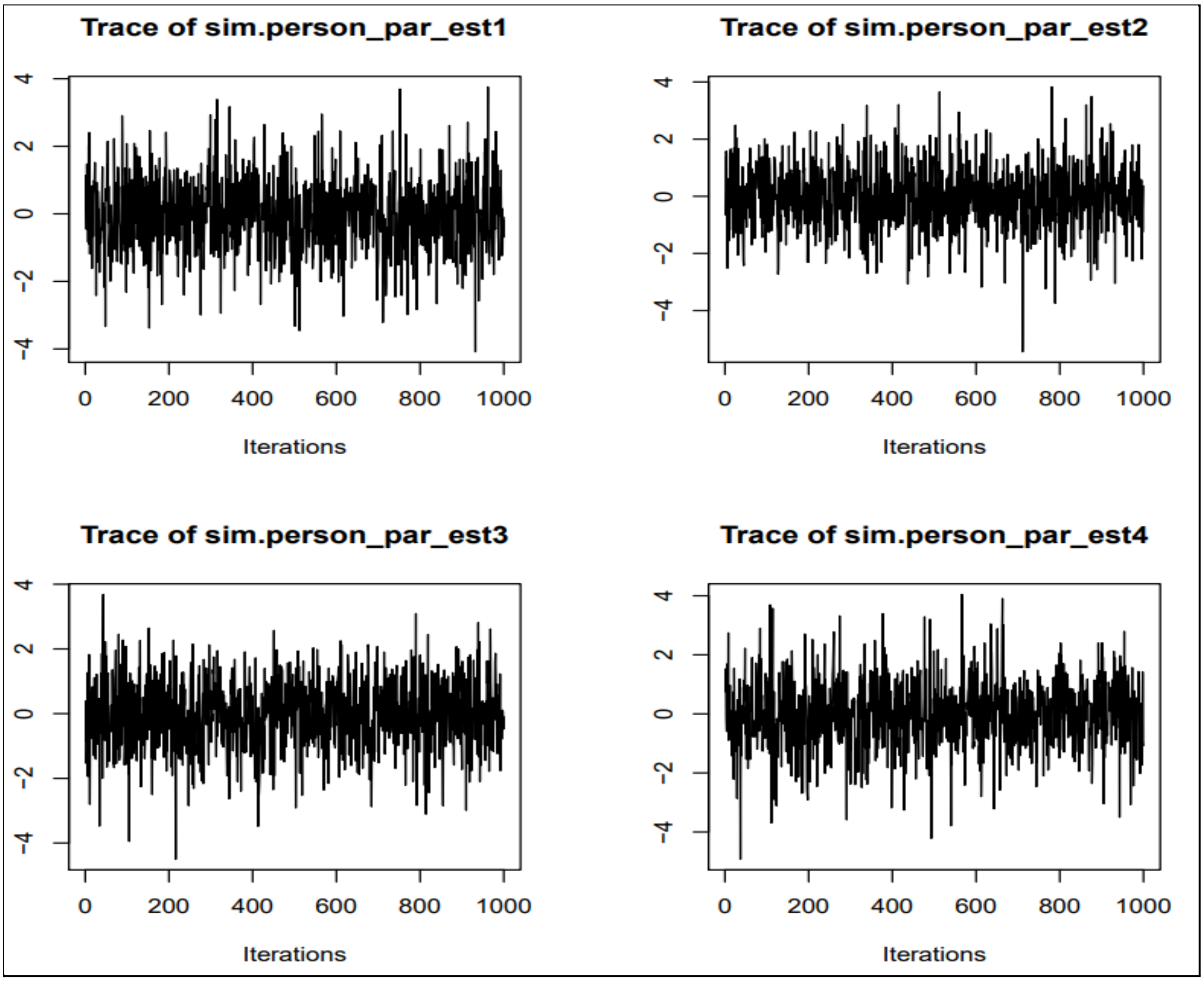

Figure 1: Trace Plots for the Convergence Diagnostic

Table 2 summarizes the bias statistics for each of the 126 simulation conditions. Based on these findings, it can be seen that, as compared to the simulated survey that is generated from the standard normal distribution (*n*=5, *N*=100, Bias=0.2434), the bias statistics were slightly higher in the simulated survey data which were generated with skewness=+1 (*n*=5, *N*=100, Bias=0.2654), and skewness=-1 (*n*=5, *N*=100, Bias=0.2666). Also, the bias statistics were becoming increasingly larger in the cases of skewness=+3 (*n*=5, *N*=100, Bias=0.3854), and skewness=-3 (*n*=5, *N*=100, Bias=0.4064). Moreover, the bias statistics continued to increase in distribution that was highly skewed to the right (skewness=5, *n*=5, *N*=100, Bias=0.3974), and in distribution that was highly skewed to the left (skewness=-5, *n*=5, *N*=100, Bias=0.4085). Besides that, bias statistics for the lower sample sizes were also found to be significantly higher compared to the larger sample size (skewness=5, *n*=5, *N*=100, Bias=0.3974; skewness=-5, *n*=5, *N*=100, Bias=0.4085) for both positively skewed datasets (skewness=5, *n*=5, *N*=50, Bias=0.4209; skewness=5, *n*=5, *N*=30, Bias=0.4571) and negatively skewed dataset (skewness=-5, *n*=5, *N*=50, Bias=0.4321; skewness=-5, *n*=5, *N*=30, Bias=0.4682). As expected, the findings clearly reveal that the maximum likelihood estimator yielded more biased estimates of the person parameter in skewed distributions, especially in smaller sample sizes. These findings suggest that the bias in estimation was significantly more severe in highly skewed simulated survey data. Moreover, another important finding to note here is that bias statistics were also significantly higher for simulated survey data generated with a smaller number of items (skewness=5, *n*=5, *N*=100, Bias=0.3974; skewness=5, *n*=10, *N*=100, Bias=0.3127) compared to those generated with a larger number of items (skewness=5, *n*=25, *N*=100, Bias=0.2163; skewness=5, *n*=30, *N*=100, Bias=0.2312). Thus, it reveals that the number of items also play a vital role that can affect the performance of the maximum likelihood estimator when estimating the unknown parameters in the model. These results hold across all simulation conditions. Overall, the findings from this study are in line with the results obtained by other researchers that are conducted with a two-parameter logistic model [1] and dichotomous Rasch model [2].

Table 2: Comparison of Estimation Performance for the Person Measure using Maximum Likelihood Approach

N	**Normal**	**SkewNormal (1)**	**SkewNormal (3)**	**SkewNormal (5)**	**SkewNormal (-1)**	**SkewNormal (-3)**	**SkewNormal (-5)**
				***n*=5**			
100	0.2434	0.2654	0.3854	0.3974	0.2666	0.4064	0.4085
50	0.2669	0.2889	0.4089	0.4209	0.2901	0.4299	0.4321
30	0.3031	0.3251	0.4451	0.4571	0.3263	0.4661	0.4682
				***n*=10**			
100	0.1267	0.2487	0.3007	0.3127	0.2499	0.3217	0.3238
50	0.1398	0.2618	0.3138	0.3258	0.263	0.3348	0.3369
30	0.1690	0.2917	0.3436	0.3551	0.2922	0.364	0.3661
				***n*=15**			
100	0.0899	0.2119	0.2639	0.2759	0.2131	0.2849	0.2879
50	0.1081	0.2301	0.2821	0.2941	0.2313	0.3031	0.3052
30	0.1279	0.2499	0.3019	0.3139	0.2511	0.3229	0.3254
				***n*=20**			
100	0.0763	0.1883	0.2403	0.2523	0.1895	0.2613	0.2634
50	0.0932	0.2052	0.2572	0.2692	0.2064	0.2782	0.2803
30	0.1152	0.2272	0.2792	0.2912	0.2284	0.3002	0.3023
				***n*=25**			
100	0.0703	0.1823	0.2043	0.2163	0.1835	0.2253	0.2274
50	0.0858	0.1978	0.2198	0.2318	0.1995	0.2408	0.2429
30	0.1095	0.2215	0.2435	0.2555	0.2227	0.2645	0.2666
				***n*=30**			
100	0.0652	0.1772	0.1992	0.2312	0.1784	0.2202	0.2423
50	0.0826	0.1946	0.2166	0.2286	0.1958	0.2376	0.2397
30	0.0932	0.2052	0.2272	0.2392	0.2064	0.2482	0.2503

n = Test length/ Number of items; *N* = Number of respondents/ Number of sample sizes

Specifically, the findings suggest that the skewed distribution affects the performance of the maximum likelihood estimator in estimating person parameters of RRSM. Thus, in other words, it shows that normality assumption becomes an important pre-requirement for using the MLE approach while estimating person parameters of RRSM. However, this assumption may be hard to be fulfilled. In fact, in most cases, the assumption of normality is violated, especially when the sample size in the study is too small. In addition, the estimation of person parameters has also become severely biased in skewed distributions with extremely small sample sizes. Hence, this is one of the biggest concerns among researchers that requires a better solution. To deal with biases in estimating person parameters under skewed distribution, one might need to explore another parameter estimation technique. In particular, these findings could act as a vital alarm on the need for an alternative technique to the MLE approach that is much more appropriate while treating the skewed datasets.

4 Conclusion and Recommendation

Normality assumption is one of the essential criteria to be fulfilled by most statistical models. When the assumption of normality does not hold, the estimation of the parameters in the model tends to be either underestimated or overestimated. This becomes a serious issue as it will lead to misleading inferences of the analysis. In this article, the effect of the skewed distributions on the performance of the MLE approach in estimating person parameters of RRSM is presented. In conclusion, the findings of the MCMC simulation analysis indicate that the maximum likelihood estimator has produced biased person estimates in skewed distributions. It can be seen that, as skewness values increase, the bias measure also increases accordingly. In fact, the results also become more obvious in small sample sizes. This proved that the MLE approach is not a superior parameter estimation technique in dealing with the distributions that are skewed either to the left- or to the right-hand side, especially when sample sizes are too small.

It should be noted here that this finding is only limited to the estimation of person parameters in RRSM. Hence, the effect of skewed distributions on the performance of the MLE approach also needs to be further explored with other parameters; the item and threshold parameters. The findings from this analysis will provide better insight regarding the effect of the skewed distributions on the performance of the MLE while estimating the parameters in RRSM. Besides that, future research can also consider using other parameter estimation techniques (e.g. Bayesian estimation) that are more powerful to treat skewed distributions. In fact, previous studies have already proved that the Bayesian estimation is one of the preferred approaches to be used in dealing with non-normal cases [2], [11], [12] particularly in small sample sizes [13]–[19].

Acknowledgements

We would like to thank the Ministry of Education (MOE) of Malaysia for the financial support for this research. This is a three-year project which is fully funded under Fundamental Research Grant Scheme (FRGS/1/2019/STG06/UiTM//02/8).

References

[1] Y. Xu, X., & Jia, "The sensitivity of parameter estimates to the latent ability distribution.," *ETS Res. Rep. Ser.*, vol. 2011, no. 2, p. i-17, 2011.

[2] H. Finch and J. M. Edwards, "Rasch Model Parameter Estimation in the Presence of a Nonnormal Latent Trait Using a Nonparametric Bayesian Approach," *Educ. Psychol. Meas.*, vol. 76, no. 4, pp. 662–684, 2016.

[3] H. Do, "Parameter Recovery for the Four-Parameter Unidimensional Binary IRT Model: A Comparison of Marginal Maximum Likelihood and Markov Chain Monte Carlo Approaches.," Ohio University, 2021.

[4] R. P. Smits, N., Öğreden, O., Garnier-Villarreal, M., Terwee, C. B., & Chalmers, "A study of alternative approaches to non-normal latent trait distributions in item response theory models used for health outcome measurement.," *Stat. Methods Med. Res.*, vol. 29, no. 4, pp. 1030–1048, 2020.

[5] D. Andrich, "A rating formulation for ordered response categories," *Psychometrika*, vol. 43, no. 4, pp. 561–573, 1978.

[6] E. B. Andersen, "Polytomous Rash Models and their Estimation," in *In Rasch Models*, Springer, New York, NY, 1995, pp. 271–291.

[7] G. Engelhard Jr, *Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences.* Routledge, 2013.

[8] A. L. Zwinderman, A. H., & Van den Wollenberg, "Robustness of Marginal Maximum Likelihood Estimation in the Rasch Model.," *Appl. Psychol. Meas.*, vol. 14, no. 1, pp. 73–81, 1990.

[9] A. Azzalini, "A class of distributions which includes the normal ones.," *Scand. J. Stat.*, pp. 171–178, 1985.

[10] A. Azzalini, "Further results on a class of distributions which includes the normal ones.," *Statistica*, vol. 46, no. 2, pp. 199–208, 1986.

[11] D. F. Azevedo, C. L., Bolfarine, H., & Andrade, "Bayesian inference for a skew-normal IRT model under the centred parameterization.," *Comput. Stat. Data Anal.*, vol. 55, no. 1, pp. 353–365, 2011.

[12] J. Zhang, X., Wang, C., Weiss, D. J., & Tao, "Bayesian Inference for IRT Models with Non-Normal Latent Trait Distributions.," *Multivariate Behav. Res.*, pp. 1–21, 2020.

[13] H. Swaminathan, R. K. Hambleton, S. G. Sireci, D. Xing, and S. M. Rizavi, "Small sample estimation in dichotomous item response models: Effect of priors based on judgmental information on the accuracy of item parameter estimates," *Appl. Psychol. Meas.*, vol. 27, no. 1, pp. 27–51, 2003.

[14] H. Swaminathan and J. A. Gifford, "Bayesian Estimation in the Three-Parameter Logistic Model," *Psychometrika*, vol. 51, no. 4, 1986.

[15] H. Swaminathan and J. A. Gifford, "Bayesian Estimation in the Rasch Model," *J. Educ. Stat.*, vol. 7, no. 3, pp. 175–191, 1982.

[16] S.-H. Kim, A. S. Cohen, F. B. Baker, M. J. Subkoviak, and T. Leonard, "An investigation of hieratchical Bayes procedures in item response theory," 1992.

[17] J. G. Burgos, "Bayesian methods in psychological research: The case of IRT," *Int. J. Psychol. Res.*, vol. 3, no. 1, pp. 164–176, 2010.

[18] H. Finch and B. F. French, "A Comparison of Estimation Techniques for IRT Models With Small Samples," *Appl. Meas. Educ.*, vol. 32, no. 2, pp. 77–96, 2019.

[19] H. Swaminathan and J. A. Gifford, "Bayesain Estimation in the Two-Parameter Logistic Model," *Psychometrika*, vol. 50, no. 3, pp. 349–364, 1985.