Rasch measurement model: a review of Bayesian estimation for estimating the person and item parameters

Nurul Hafizah Binti Azizan, Zamalia Binti Mahmud, Adzhar Bin Rambli

Statistics, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia

hafizahaz@uitm.edu.my, zamalia@tmsk.uitm.edu.my, adzhar_rambli@tmsk.uitm.edu.my

Abstract. This paper focuses on the methods used for estimating the parameters in Rasch Measurement Model (RMM). These include the MLE and Bayesian Estimation (BE) techniques. The accuracy and precision of the parameter estimates based on these two MLE and BE were discussed and compared. A questionnaire is a well-known measurement instrument used by most of the researchers. It is a powerful tool for collecting data in survey research. It should be noted that the quality of a measurement instrument used plays a key role in ensuring the quality of data obtained in the survey. Therefore, it has become essential for the researchers to carefully design their questionnaire so that the quality of the data obtained can be preserved. Then, it is also vital for the researchers to assess the quality of the data obtained before it can be successfully used for further analysis. Review of the literature shows that RMM is a psychometric approach widely used as an assessment tool of many measurement instruments developed in various fields of study. At present, the Maximum Likelihood Estimation (MLE) techniques were used to estimate the parameters in the RMM. In order to obtain more precise and accurate parameter estimates, a certain number of sample size and normality assumption are usually required. However, in a small sample, MLE could produce bias, imprecise and less accurate estimates with bigger standard error. A proper selection of the parameter estimation techniques to deal with small sample and non-normality of data is required to obtain more precise and accurate parameter estimates. From the review, it reveals that BE has successfully dealt with the issues of small sample and non-normality of the data. It produced a more accurate parameter estimate with smallest Mean Squared Error (MSE), particularly in a small sample compared to MLE.

1. Introduction

A questionnaire is the most commonly used as a data collection tool in survey research. Before further analysis can be done, the quality of the constructed questionnaire should be properly evaluated using an appropriate technique. This is to ensure that the measurement instrument used for the survey meets the psychometrics requirements of both validity and reliability. It should be noted that the utility of the measurement instrument will depend on its quality, as it can be useful if certain psychometric and
practical requirements are met [1]. The Rasch Measurement Model (RMM) is among the most well-known psychometric model used to evaluate the measurement instrument for the survey. It is a tool for validating the measurement instrument which was developed by Georg Rasch in 1960. The first introduced model was with a dichotomous response; well known as Dichotomous Rasch Model. It has been successfully applied in the educational and learning environment. Then the application of the RMM has been widely spread in other various disciplines due to its simplicity and efficiency for validating the questionnaire used in survey research.

There are two parameters of interest in the RMM which are denoted as $\beta_v$ and $\delta_i$. These parameters refer to the ability and difficulty respectively. As remarked by previous researcher [2], these two parameters can be estimated either with non-iterative (e.g., graphical method and PROX) or iterative (e.g., Joint Maximum Likelihood Estimation (JMLE), Conditional Maximum Likelihood Estimation (JMLE) and Marginal Maximum Likelihood Estimation (MMLE)) approaches. Among these techniques, the MLE techniques are the most commonly used. The requirement of a certain number of sample size and the normality assumption are usually needed to obtain more precise and accurate parameter estimates. With a small sample, the Rasch measurement model could produce less precise estimates with bigger standard error, less powerful fit analysis and less robust estimates [3]. For more accurate and precise estimates, it is suggested collecting more relevant data to be included in the study [4]. However, most of the studies cannot fulfill these requirements due to certain constraints faced in the study. Therefore, with these circumstances, it is vital for the researcher to choose another appropriate technique as an alternative, so that the more precise and accurate parameter estimates could be obtained. The use of BE as a parameter estimation has received considerable attention not only in RMM but also in other statistical techniques and modelling for various applications. This is because BE has produced better estimates when dealing with a small sample and has been successfully utilized in various applications as compared to MLE. This article begins with a brief introduction to the RMM and its parameter estimations used. Next, the reasons for using BE have been discussed. Then, BE procedures used in estimating the parameters in the Dichotomous Rasch Model as proposed by previous researcher [5] are described. A comparison of accuracy and precision of the parameter estimates between MLE and BE approaches as found by previous studies is then provided.

2. Rasch measurement model

Rasch Measurement Model (RMM) among the most widely used psychometric approach in various disciplines. It facilitates as an assessment tool for many measurement instruments. It is a probabilistic statistical model introduced by Georg Rasch in 1960 which is also known as one-parameter logistic (1PL) model, since it depends on only a single item parameter called difficulty. The aim of RMM is to predict the probability of success or correct response of a person on an item, which depends on the person’s ability ($\beta_v$) and the item difficulty ($\delta_i$). As pointed out by previous researchers [6], the difference ($\beta_v - \delta_i$) yields the values between negative infinity and positive infinity, while the probability can only have the values between zero and one. To deal with this circumstance, an exponent of the natural constant $e = 2.71828$ has been introduced to the difference ($\beta_v - \delta_i$). Thus, in the Dichotomous Rasch Model, the probability of person $v$ with ability $\beta_v$ giving response $x_{vl}$ to item $i$ with difficulty $\delta_i$ will be obtained as given in equation (1) (6,7);

$$P(X_{vl} = x_{vl} | \beta_v, \delta_i) = \frac{\exp[x_{vl}(\beta_v - \delta_i)]}{1 + \exp(\beta_v - \delta_i)} ; X_{vl} \in \{0|1\}$$

where the notation $X_{vl} \in \{0|1\}$ implies that the element of $X_{vl}$ will be either 0 or 1 ($x_{vl} = 0$ for a failure or incorrect response and $x_{vl} = 1$ for a success or correct response). Only the formula given in equation (1) has a shape that suits perfectly the response curve given in figure 1, which makes a special characteristic of RMM called “specific objectivity” [6]. They also mentioned that this condition has
allowed the $\beta_v$ and $\delta_l$ to be estimated independently of one another. As this model only allows 0 and 1 responses, thus $X_{vl}$ is specified to follow Bernoulli distribution, $X_{vl} \sim \text{Bernoulli}(p_{vl})$. A probability of success/correct response $P(X_{vl} = 1|\beta_v, \delta_l)$ and the probability of a failure/incorrect response $P(X_{vl} = 0|\beta_v, \delta_l)$ for the basic Rasch model with binary or dichotomous data can be written as in equation (2) and equation (3).

$$\pi = P(X_{vl} = 1|\beta_v, \delta_l) = \frac{\exp(\beta_v - \delta_l)}{1 + \exp(\beta_v - \delta_l)}$$

$$1 - \pi = P(X_{vl} = 0|\beta_v, \delta_l) = \frac{1}{1 + \exp(\beta_v - \delta_l)}$$

where

$P(X_{vl} = 1|\beta_v, \delta_l)$ is the probability of a correct response for person $p$ on item $i$.

$P(X_{vl} = 0|\beta_v, \delta_l)$ is the probability of an incorrect response for person $p$ on item $i$.

$\beta_v$ is the ability of person $p$ on the latent variable scale.

$\delta_l$ is the difficulty of item $i$ on the latent variable scale.

The crucial requirement of the Rasch measurement model is “specific objectivity”, where the authors [7] in their book entitled Rating Scale Analysis defined specific objectivity as follow

“Specific objectivity means that the model can be written in a form in which its parameters are linear in the argument of an exponential expression so that they can be sufficiently estimated and conditioned out of the estimation of other parameters.”

Another important feature of the Rasch measurement model is the concept of “invariant measurement”. This is a key concept of the theory behind the development of the RMM, which encounters the limitation of the traditional approach to measurement [2]. According to another authors [8], invariant measurement can be defined as follow:

“The invariant requirement is that the values (measures) attributed to variables by any measurement system should be independent of the particular measurement instrument used (as long as the instrument is appropriate to the purpose), that is, the values should be invariant. Moreover, the calibrations of the measurement instrument should also remain invariant across any of its intended purposes.”

These concepts of “specific objectivity” and “invariant measurement” indicates that RMM is a unique and more practical approach as compared to the other classical technique of the measurement theory. The relationship between $\beta_v$ and $\delta_l$ is represented based on the response curve shown in Figure 1 [6]. The curve reveals that the probability of success will be equal to 0.5, $P(X_{vl} = 1) = 0.5$, if the person’s ability and item difficulty are equal $\beta_v = \delta_l$. In the case of person’s ability that is more than item difficulty $\beta_v > \delta_l$, the probability of success will be more than 0.5, $P(X_{vl} = 1) > 0.5$. In contrast, with the condition of the person’s ability that is less than item difficulty $\beta_v < \delta_l$, the probability of success will be less than 0.5, $P(X_{vl} = 1) < 0.5$. The relationship between the probabilities of success on an item for people with varying ability can be plotted as in the response curve shown in figure 1.
2.1. Parameter estimation techniques for estimating the person and item measures in the Rasch Model

In RMM framework, there are two parameters of interest to be estimated. These parameters refer to the person-level latent trait ($\beta_v$) and item-location ($\delta_l$) parameters. Generally, the method of MLE has been applied in RMM for estimating these two parameters in the past decade. The use of MLE as a parameter estimation is aimed to estimate the parameter which can maximize the likelihood function or probability of observations. In the literature, it is found that the three well-known Maximum Likelihood techniques used to estimate either the person and item parameters included MMLE [9–11], CMLE [12–14] and JMLE [15–17].

However, the requirement of large number of samples and normality assumption of the data in as required by MLE are mostly hard to fulfill due to certain limitations and constraints. Thus, with these circumstances, a more suitable method is needed. Alternatively, the use of Bayesian technique as a parameter estimation has become increasingly popular in several statistical techniques and modelling for various applications [18–21]. This is because it is proven that BE is more plausible to analyse small sample data [2]. BE has also been tested in the psychometric models such as Dichotomous Rasch Model [5], Two-Parameter Logistic model [23] and Three-Parameter Logistic model [24]. The results consistently showed that with a constraint in the number of samples, BE has outperform MLE as it produced the smallest MSE.

2.2. Reasons for using Bayesian estimation

The use of Bayesian technique as a parameter estimation has become increasingly popular not only in Rasch measurement model but also in other statistical techniques and modelling for various applications [25–27]. A steady increase of the application of BE as a standard estimation tool in psychometric models and other applications are due to wide range of reasons [28]. From the review, it was found that, among of the reason is due to computational difficulty faced with MLE [29], convergence issues with MLE [30], when dealing with a complex model [31], in handling missing data [32] and outlier problem [33]. Moreover, BE can be used to improve the performance of parameter estimation in small sample (25) and in violating assumption of other estimation technique [34], including normality assumption [35]. BE also has been preferred to be used for producing more accurate parameter estimates compared to conventional methods such as MLE [36]. Especially when sample size involved is too small, as small
samples might cause parameter estimates produced by MLE reduced its accuracy with inflated standard error [37].

2.3. Bayesian estimation procedures for estimating person and item measures in Rasch measurement model

BE procedures described in this article are based on work done by the previous authors [5]. Recall that the probability of person \( v \) with ability \( \beta_v \) giving response \( x_{vl} \) to item \( i \) with difficulty \( \delta_l \):

\[
P(X_{vl} = x_{vl}|\beta_v, \delta_l) = \frac{\exp[x_{vl}(\beta_v - \delta_l)]}{1 + \exp(\beta_v - \delta_l)}
\]  

The likelihood \( \Lambda \) of the data matrix \( (X_{vl}) \) is obtained based on the product of the equation (4). This is the probability of the whole data matrix \( (X_{vl}) \) given its parameters \( (\beta_v) \) and \( (\delta_l) \), which written as follows:

\[
\Lambda = P\left((X_{vl})| (\beta_v), (\delta_l)\right) = \prod_v \prod_l \left\{ \frac{\exp[x_{vl}(\beta_v - \delta_l)]}{1 + \exp(\beta_v - \delta_l)} \right\}
\]

and the numerator is simplified as:

\[
\prod_v \prod_l \{\exp[x_{vl}(\beta_v - \delta_l)]\} = \exp \sum_v \sum_l x_{vl}(\beta_v - \delta_l)
\]

\[
\sum_v \sum_l x_{vl}(\beta_v - \delta_l) = \sum_v \sum_l x_{vl}\beta_v - \sum_v \sum_l x_{vl}\delta_l
\]

Then, the two components on the right hand-side been written as follow:

\[
\sum_v \sum_l x_{vl}\beta_v = \sum_v r_v\beta_v \quad \text{where} \quad \sum_l x_{vl} = r_v
\]

\[
\sum_v \sum_l x_{vl}\delta_l = \sum_l s_l\delta_l \quad \text{where} \quad \sum_v x_{vl} = s_l
\]

Thus, the likelihood \( \Lambda \) can be expressed as:

\[
\Lambda = L(\beta_v, \delta_l|x) = \left\{ \frac{\exp(\sum_v r_v\beta_v - \sum_l s_l\delta_l)}{\prod_v \prod_l [1 + \exp(\beta_v - \delta_l)]} \right\}
\]

The application of the Bayesian approach to estimate the parameters involved in the models required for the researchers to specify the prior information. The choice of suitable prior information plays a significant role in the Bayesian estimation framework, as it has an impact on the accuracy of Bayesian inference. It is also highlighted that the use of prior information in Bayesian framework, certainly
increases the accuracy of the parameter estimates [23, 24]. In addition, with a wrong choice of prior, improper posterior distribution could be produced [38, 39], which lead to poor or inaccurate Bayesian inference [40]. Generally, there are two types of priors. Namely noninformative and informative priors. Noninformative is one that provide just little or no information related to the parameter, while the informative prior is a type of prior that has a considerable amount of information about the parameter of interest [40]. Therefore, the choice of prior (either informative or noninformative) to be used depends on the availability of the information about the parameters [41]. Based on the study conducted by previous researcher [5], there are two stages involved to obtain the posterior distribution. At the first stage, they have specified the prior to the person $\beta_v$ and item $\delta_l$ parameters, that are independently and identically normally distributed each, with respected mean $\mu$ and variance $\phi$. They are $\beta_v \sim \mathcal{N}(\mu_\beta, \phi_\beta)$ and $\delta_l \sim \mathcal{N}(\mu_\delta, \phi_\delta)$. At the second stage, the prior of the parameters $\mu$ and variance $\phi$ have been assumed to follow uniform distribution, $p(\mu, \phi) \propto p(\phi)$ and the inverse chi-square distribution, $p(\phi|\nu, \lambda) \propto \phi^{-(\nu+1)/2} \exp(-\lambda/2\phi)$, for the respected person $\beta_v$ and item $\delta_l$ parameters. Then, a certain mathematical process (i.e., derivation and integration) including Newton-Raphson iterative procedure (which are not discussed here) were involved to obtained the joint posterior of $\beta_v$ and $\delta_l$ as:

$$p(\beta_v, \delta_l, \mu_\beta, \phi_\beta, \mu_\delta, \phi_\delta|x, v_\beta, \lambda_\beta, v_\delta, \lambda_\delta)$$

$$\propto L(\beta_v, \delta_l|x) \prod_{v=1}^{N} p(\beta_v|\mu_\beta, \phi_\beta) \prod_{l=1}^{L} p(\delta_l|\mu_\delta, \phi_\delta) p(\phi_\beta|v_\beta, \lambda_\beta) p(\phi_\delta|v_\delta, \lambda_\delta)$$

where $L(\beta_v, \delta_l|x)$ is a likelihood for the model as given in equation (8). Then, the joint posterior density of $\beta_v$ and $\delta_l$ is obtained as:

$$p(\beta_v, \delta_l|x, v_\beta, \lambda_\beta, v_\delta, \lambda_\delta) = \left\{ \exp \left( \sum_{v=1}^{N} r_v \beta_v \right) \left[ \lambda_\beta + \sum_{v=1}^{N} (\beta_v - \beta) \right]^{\frac{(N+v_\beta-1)}{2}} \right\}$$

$$\cdot \left\{ \exp \left( - \sum_{l=1}^{L} s_l \delta_l \right) \left[ \lambda_\delta + \sum_{l=1}^{L} (\delta_l - \bar{\delta}) \right]^{\frac{(L+v_\delta-1)}{2}} \right\}$$

$$\cdot \left\{ \prod_{v=1}^{N} \prod_{l=1}^{L} \left[ 1 + \exp(\beta_v - \delta_l) \right] \right\}^{-1}$$

The notation $\beta$ and $\delta$ in the equation (10) imply mean for the person and item measures respectively. In this paper, the authors used joint posterior modes for the $\beta_v$ and $\delta_l$ estimates. These two estimates have been obtained by letting derivatives of $\log p(\beta_v, \delta_l|x, v_\beta, \lambda_\beta, v_\delta, \lambda_\delta)$ equal to zero. As the equation is nonlinear, hence the Newton-Raphson iterative procedure has been employed to solve the equation. Then, at the final, the estimated $\beta_v$ and $\delta_l$ at the $k$th iteration were obtained as follow:

$$\beta_v^{(k+1)} = \beta_v^{(k)} - f(\beta_v^{(k)})/f'(\beta_v^{(k)})$$

$$\delta_l^{(k+1)} = \delta_l^{(k)} - h(\delta_l^{(k)})/h'(\delta_l^{(k)})$$
The respected initial values of $\beta_v$ and $\delta_l$ are given as:

$$\beta_v^{(0)} = \delta + (1 + \omega^2 / 2.89)^{1/2} \log \left( r_v / (L - r_v) \right)$$  \hspace{1cm} (13)

and

$$\delta_l^{(0)} = \log \left( (N - s_l) / s_l \right)$$  \hspace{1cm} (14)

In this case, $s_l = 1/2$ and $N = -1/2$. Then, the values of $\beta_v$ are estimated. Next, with these estimated values of $\beta_v$, the revise estimates of $\delta_l$ were obtained. This process is repeated until the convergence criterion is reached.

2.3.1. A comparison of accuracy and precision of the parameter estimates between MLE and BE

The application of BE in the Dichotomous Rasch model was first introduced in 1982 [5]. Through simulation proses as demonstrated in this study, BE produced more accurate item parameter estimates with much smaller MSE as compared to MLE approach. The authors claimed that this successful result was produced due to incorporating prior information of the item parameters. Other study also claimed that as compared to the MLE, BE is more plausible to analyse small sample data [42]. The findings from this study have proven that the use prior information has increased the statistical power of analysis, although with only 20 samples involved in the study. In addition, it is also showed that MLE has insufficient coverage and power with very small sample, where it can be solved by including prior information as required by BE [25]. In this study they found that BE outperform the MLE, as smaller data set can be analysed without losing power while the precision and accuracy of the estimates can still be retained. Besides, in the presence of non-normal latent traits, MLE consistently produced less accurate estimation with much higher Mean Squared Error (MSE) compared to BE [35]. All the findings discussed earlier have shown that the choice of prior is an important contributor for more accurate and precise Bayesian estimates. It should be noted that a specification of the prior is one of the key element in the Bayesian framework [43], where appropriate selection of prior is an important issue to be addressed by the researchers when they decided to use BE [44].

3. Conclusion

From the review, it revealed the use of BE has been widespread not only in the psychometric models (i.e., Dichotomous Rasch model, Two-Parameter Logistic and Three-Parameter Logistic models), but also across fields and many other disciplines. The BE is gaining in popularity and becoming an alternative choice among the researchers might be due to various reasons. These include computational difficulty faced with MLE, convergence issues occurred with MLE and when dealing with a complex model, in handling missing data and outlier problem. BE also been used in order to improve the performance of parameter estimation in small sample and when the assumptions of other estimation technique have been violated (i.e., normality assumption). In addition, it also been employed for producing more accurate parameter estimates compared to conventional or frequentist approaches such as MLE. In the BE framework, it is important to note that the choice of prior information to be used play a key role in the accuracy of the Bayesian inference. This is because the wrong choice of prior information may introduce bias in the estimates, hence leads to poor or inaccurate Bayesian inference. Therefore, if BE is an alternative approach to be considered, it is important for the researchers to emphasize on the selection of prior information to be used especially when dealing with small sample, so that more accurate estimates could be produced. The review also indicated that, in general, as compared to the MLE, BE empirically yielded more accurate and precise estimates with the lowest MSE.
in such conditions of a small sample and non-normality distribution of data. Thus, the review provides some evidence that the BE may be a preferred estimation tool over the frequentist approach such as MLE, not only in the presence of non-normality problem and small sample but also other conditions as mentioned earlier.

References

[1] Shultz K S, Whitney D J and Zickar M J 2013 Measurement theory in action: Case studies and exercises (Routledge)
[2] Engelhard Jr G 2013 Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences (Routledge)
[3] Linacre J M 1994 Sample Size and Item Calibration or Person Measure Stability Rasch Meas Trans 7(4):328
[4] Linacre J M 1999 Estimation Methods for Rasch Measures J Outcome Meas 3:381–405
[5] Swaminathan H and Gifford J A 1982 Bayesian Estimation in the Rasch Model. J Educ Stat 7(3):175–91
[6] Wright B D and Stone M H 1979 Best Test Design (MESA press)
[7] Wright B D and Masters G N 1982 Rating Scale Analysis (MESA press)
[8] Bond T and Fox C M 2015 Applying the Rasch model: Fundamental measurement in the human sciences (Routledge)
[9] Glas C A 1988 The derivation of some tests for the Rasch model from the multinomial distribution Psychometrika 53(4):525–46
[10] Thissen D 1982 Marginal maximum likelihood estimation for the one-parameter logistic model Psychometrika 47(2):175–86
[11] Bock R D and Aitkin M 1981 Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm Psychometrika 46(4):443–59
[12] Agresti A 1993 Computing conditional maximum likelihood estimates for generalized Rasch models using simple loglinear models with diagonals parameters Scand J Stat 63–71
[13] Adams R J, Wilson M and Wang W C 1997 The multidimensional random coefficients multinomial logit model Appl Psychol Meas. 21(1):1–23
[14] Rost J 1990 Rasch models in latent classes: An integration of two approaches to item analysis. Appl Psychol Meas. 14(3):271–82
[15] Rost J and Carstensen C H 2002 Multidimensional Rasch measurement via item component models and faceted designs Appl Psychol Meas. 26(1):42–56
[16] Willse J T 2011 Mixture Rasch models with joint maximum likelihood estimation Educ Psychol Meas. 71(1):5–19
[17] Doran H, Bates D, Bliese and Dowling M 2007 Estimating the multilevel Rasch model: With the lme4 package J Stat Softw. 20(2):1–18
[18] Brav A 2000 Inference in Long-Horizon Event Studies: A Bayesian Approach with Application to Initial Public Offerings J Finance. 55(5):1979–2016
[19] Kochi I, Hubbell B and Kramer R 2006 An empirical Bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis Environ Resour Econ. 34(3):385–406
[20] Fletcher P C and Frith C D 2009 Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia Nat Rev Neurosci. 10(1):48
[21] Spiegelhalter D J and Best N G 2003 Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling Stat Med. 22(23):3687–709
[22] Zhang J and Shields M D 2018 The effect of prior probabilities on quantification and propagation of imprecise probabilities resulting from small datasets Comput Methods Appl Mech Eng. 334:483–506
[23] Swaminathan H and Gifford J A 1985 Bayesian Estimation in the Two-Parameter Logistic Model. *Psychometrika.* 50(3):349–64.

[24] Swaminathan H and Gifford J A 1986 Bayesian Estimation in the Three-Parameter Logistic Model. *Psychometrika.* 51(4)

[25] Van De Schoot R, Broere J J, Perryck K H, Zondervan-Zwijnenburg M and Van Loey N E 2015 Analyzing small data sets using Bayesian estimation: The case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *Eur J Psychotraumatol.* 6(1):25216

[26] Wang M and Wang W 2017 Bias-corrected maximum likelihood estimation of the parameters of the weighted Lindley distribution *Commun Stat Comput.* 46(1):530–45

[27] Boulet S, Ursino M, Thall P, Jannot A S and Zohar S 2019 Bayesian variable selection based on clinical relevance weights in small sample studies—Application to colon cancer. *Stat Med.*

[28] Van de Schoot R, Winter S D, Ryan O, Zondervan-Zwijnenburg M and Depaoli S 2017 A systematic review of Bayesian articles in psychology: The last 25 years *Psychol Methods.* 22(2):217

[29] Ryoo J H, Molfese V J, Brown E T, Karp K S, Welch G W and Bovaird J A 2015 Examining factor structures on the Test of Early Mathematics Ability—3: A longitudinal approach *Learn Individ Differ.* 41:21–9

[30] Lee J, Choi J Y and Cho Y 2011 A forecast simulation analysis of the next-generation DVD market based on consumer preference data *Int J Consum Stud.* 35(4):448–57

[31] Wang C and Nydick S W 2015 Comparing two algorithms for calibrating the restricted non-compensatory multidimensional IRT model *Appl Psychol Meas.* 39(2):119–34

[32] Öztürk N K and Karabatsos G 2017 A Bayesian robust IRT outlier-detection model. *Appl Psychol Meas.* 41(3):195–208

[33] Desmet P and Feinberg F M 2003 Ask and ye shall receive: The effect of the appeals scale on consumers’ donation behavior *J Econ Psychol.* 24(3):349–76

[34] Desmet P, Desmet L, Gutscher D and Feinberg F M 2007 Investigating a weakly informative prior for item scale hyperparameters in hierarchical 3PNO IRT models *Front Psychol.* 8:123

[35] Finlayson J and Edwards J M 2016 Rasch Model Parameter Estimation in the Presence of a Nonnormal Latent Trait Using a Nonparametric Bayesian Approach *Edu Psychol Meas.* 76(4):662–84

[36] Wilson S J, Barrineau M J, Butner J and Berg C A 2014 Shared possible selves, other-focus, and perceived wellness of couples with prostate cancer *J Fam Psychol.* 28(5):684

[37] Finch H and French B F 2019 A Comparison of Estimation Techniques for IRT Models With Small Samples *Appl Meas Educ.* 32(2):77–96

[38] Syversveen A R 1998 Noninformative bayesian priors. interpretation and problems with construction and applications [Internet]. Vol. 3, *Preprint statistics* p. 1–11. Available from: https://www.ime.unicamp.br/~veronica/M1402/Randi21998.pdf

[39] Lindqvist B H and Taroalson G 2018 On the proper treatment of improper distributions. *J Stat Plan Inference.* 195:93–104

[40] Box G E and Tiao G C 2011 Bayesian inferences in statistical analysis. John Wiley & Sons

[41] Sheng Y 2017 Investigating a weakly informative prior for item scale hyperparameters in hierarchical 3PNO IRT models *Front Psychol.* 8:123

[42] Zhang Z, Hamagami F, Lijuan Wang L, Nesselroade J R and Grimm K J 2007 Bayesian analysis of longitudinal data using growth curve models *Int J Behav Dev.* 31(4):374–83

[43] Marcoulides K M 2018 Careful with those priors: A note on Bayesian estimation in two-parameter logit model response theory models *Meas Interdiscip Res Perspect.* 16(2):92–9.

[44] Natesan P, Nandakumar R, Minka T and Rubright J D 2016 Bayesian prior choice in IRT estimation using MCMC and variational Bayes *Front Psychol.* 7:1422