The JML estimation in the equi-distant unfolding model for polytomous responses

Guanzhong Luo [1,2]

[1] Jiangxi Normal University, Nanchang, China
[2] South China Normal University, Guangzhou, China
gz.luo@iclound.com

Abstract. This paper operationalizes the general unfolding model for polytomous responses under the condition of equal-distances between the successive thresholds of an item. Using the joint maximum likelihood estimation (JML), two algorithms are proposed for parameter estimation, each with a different technique to reduce the bias in the estimates produced by the inconsistency of the JML. One algorithm is a direct extension of the algorithm described in [15] in which a correction is applied on the mean latitude of acceptance parameter. Simulation studies show the efficiency of the algorithms in various situations. An analysis of real data is included for illustrative purposes.

Keywords: polytomous responses, unfolding models, joint maximum likelihood estimation, weighted likelihood estimation.

1 Introduction

Thurstone [1] introduced his rigorous method of scaling for the measurement of attitude. This method involved two stages. In the first stage, the statements are located by judges on a continuum which is often called a scale. As the result, each of the statements is assigned a scale value which may vary from negative, through neutral, to positive on the scale. In the second stage, the measurement of attitude is obtained by the *mid-point principle*, which takes the mean or median of the scale values of the statements that a person agrees to as the attitude value of the person on the scale. Although Thurstone's scaling principles were widely accepted as rigorous, the most popular method for the measurement of attitude in the last 70 years was not Thurstone's, but Likert's [2] *summated rating scaling*. In fact, after reversing the responses on the negative items, Likert's scaling uses a cumulative process characterized by monotonic response functions. The total scores of persons across items are used as the measures of attitudes. In contrast, the mid-point principle, which is used in the second stage of the Thursone's scaling, implies a single-peaked response function which characterizes what is often termed an *unfolding* process. These two response processes are not compatible [3] [4].

In the last decade or so, developments in computer technology and psychometric models have made the application of single-peaked response functions more practical, thus eliminating the need for the first stage of the Thurstone's procedure, even when

Q. Zhang (ed.), *Proceedings of the Pacific-Rim Objective Measurement Symposium (PROMS 2023)*, Atlantis Highlights in Social Sciences, Education and Humanities 23,
https://doi.org/10.2991/978-94-6463-494-5_3

single-peaked response processes are employed. The analyses of the responses immediately give the locations of both the persons and the statements, and therefore retain the features of Thurstone scaling. Various commercial packages for analyzing dichotomous responses with unfolding models are available (e.g. [5] [6]). Specific probabilistic unfolding models for polytomous responses have also been proposed ([3]; [7]; [8]; [9] and [10]), with accompanying software.

Luo [11] introduced a general class of probabilistic unfolding models for polytomous responses with the rationale of the *rating formulation* [12]. The feature of this class is that in addition to the item parameter, each item has a series of thresholds. These thresholds define intervals within which the most likely response is the corresponding category.

In [13], the form of the general class of unfolding models is expressed as

$$\Pr\{X_{ni} = k\} = \frac{\prod_{l=1}^{k} \Psi_l(\rho_{il}) \prod_{l=k+1}^{m_i} \Psi_l(\beta_n - \delta_i)}{\lambda_{ni}}, \quad k = 0, \ldots, m_i; \tag{1}$$

where β_n is the location parameter for person *n*, δ_i is the location parameter of item i, and ρ_{ik} (≥ 0) is the k-th threshold for item i. The functions $\{\Psi_k\}$ was termed the *operational functions,* and λ_{ni} is the normalization factor

$$\lambda_{ni} = \sum_{k=0}^{m_i} \prod_{l=1}^{k} \Psi_l(\rho_{il}) \prod_{l=k+1}^{m_i} \Psi_l(\beta_n - \delta_i). \tag{2}$$

An equivalent expression of Equation (1) is

$$\pi_{ik}(\beta_n - \delta_{ik}, \rho_{ik}) = \frac{\Pr\{X_{ni} = K\}}{\Pr\{X_{ni} = K-1\} + \Pr\{X_{ni} = K\}} = \frac{\Psi_k(\rho_{ik})}{\Psi_k(\rho_{ik}) + \Psi_k\{\beta_n - \delta_{ik}\}},$$

$$k = 1, \ldots, m_i. \tag{3}$$

The set of probabilities $\{\pi_{ik}, k = 1, \ldots, m\}$ have the form of a probabilistic unfolding model for dichotomous responses, thought these dichotomous responses are *latent* and not observed. The probabilistic functions of (1) and the curves of corresponding $\{\pi_{ik}, k = 1, \ldots, m\}$ in (3) are shown in Figure 1.

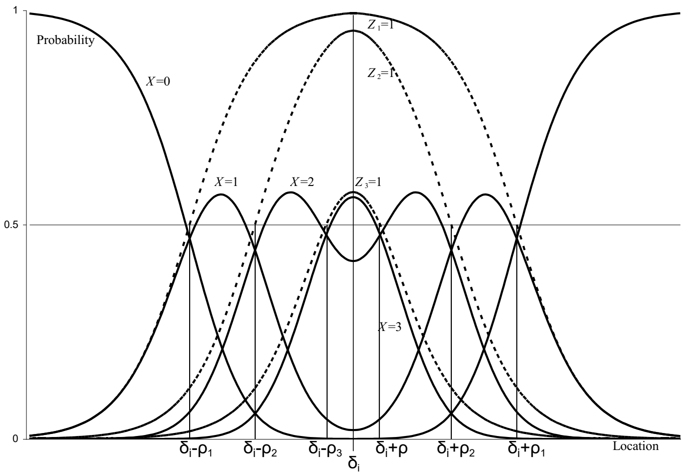


Fig. 1. Probabilistic functions of the general form for unfolding models for polytomous response (in solid lines) with the probabilistic functions of the latent dichotomous variables (In broken lines).

Figure 1 shows that if thresholds are in their serial order, $\rho_{i1} > \rho_{i2} > \ldots > \rho_{ik} > \ldots > \rho_{im}$, then ($\delta_i - \rho_{im}, \delta_i + \rho_{im}$) is the interval within which the response category m has the greatest probability. That is, for a person n with location $\beta_n \in (\delta_i - \rho_{im}, \delta_i + \rho_{im})$, the response ($x_{ni} = m$) is the most likely. For $k < m$, $(\delta_i - \rho_{i(k-1)}, \delta_i - \rho_{ik})$ and $(\delta_i + \rho_{ik}, \delta_i + \rho_{i(k-1)})$ are the two intervals in which the response category k has the greatest probability. That is, for a person n with location $\beta_n \in (\delta_i - \rho_{i(k-1)}, \delta_i - \rho_{ik}) \cup (\delta_i + \rho_{ik}, \delta_i + \rho_{i(k-1)})$, the response ($x_{ni} = k$) is the most likely.

A simpler case of model (1) is that all operational functions involved are identical (denoted as Ψ without a subscript) and the adjacent thresholds of an item have an equal distance:

$$\rho_{il} - \rho_{i(l+1)} = \zeta_i, \qquad l = 1,\ldots,m_i - 1; \tag{4}$$

Or

$$\rho_{il} = (m_i + 1 - l)\zeta_i, \qquad l = 1,\ldots,m_i; \tag{5}$$

where $\zeta_i > 0$.

It is anticipated that the investigation of the *equi-distant unfolding model* above would lead to further understanding of the general class of probabilistic unfolding models for polytomous responses. In addition, development of parameter estimation procedures for equi-distant unfolding model itself would be useful for analyzing the data where a single operational function and equi-distant thresholds are justified. Similar investigation in the family of Rasch models were studied by [14].

The development of parameter estimation procedures of the model above is to have three main steps. First, it extends the joint maximum likelihood (JML) algorithm for the general dichotomous unfolding model [15] to the polytomous equi-distant unfolding model described above to derive the solution equations. Second, with the same set of solution equations, it operationalizes two alternative procedures for minimizing the bias of the estimates, which is known to exist due to the inconsistency of the direct application of the JML. Third, because the solution equations derived are general, it examines the efficiencies of these two procedures with respect to two specific models with different operational functions, namely the hyperbolic cosine model (HCM), in which $\Psi(t) = \cosh(t)$, and the simple square logistic model (SSLM), in which $\Psi(t) = \exp(t^2)$. This examination is carried out with a series of simulated data sets and a real data set.

The rest of this paper is structured as follows. Section 2 begins with the expression of the equi-distant unfolding model, followed by a rationale for using the JML approach among other estimation approaches available. The solution equations with the JML approach are then derived, and two algorithms are proposed for parameter estimation. Though both of the algorithms use a two-stage approach with the same set of solution equations, the techniques for controlling bias in the estimates are different. One algorithm is a direct extension of the algorithm proposed in [15] and [16] , in which a correction procedure is applied to the estimate of the mean latitude of acceptance parameter to reduce the bias caused by the inconsistency of the JML, while the other prevents the accumulation of the bias in the estimation cycles by means of the weighted likelihood estimation of person locations ([17] , [18] and [19]). The results of simulation studies for the comparison of these two algorithms are presented in Appendix E. Appendix F provides an illustrative example with real data, in which the results from the two operational functions and algorithms are compared. A discussion with consideration of some issues for further investigation is also provided. As the focus of this paper is on efficient estimation procedures, the issue of the test fit, which is at least equally t.important for practical purposes, is left for further development.

2 Estimation procedure

2.1 The equi-distant unfolding model.

To emphasize that items may have different maximum scores and different distances between the thresholds, the maximum score and the distance between the adjacent thresholds of item i are denoted as m_i and ζ_i respectively. The condition of the equi-distant unfolding model can be expressed in terms of ζ_i as

$$\rho_{il} - \rho_{i(l+1)} = \zeta_i, \qquad l = 1,\ldots,m_i - 1; \tag{6}$$

$$\rho_{il} = (m_i + 1 - l)\zeta_i, \qquad l = 1,\ldots,m_i. \tag{7}$$

The parameter $\zeta_i > 0$ is termed the *item unit* as it specifies the distance between the adjacent thresholds. Then the equi-distant unfolding model can be written as

$$\begin{aligned}\Pr\{X_{ni} = k\} &= \frac{\prod_{l=1}^{k}\Psi((m_i+1-l)\zeta_i)\prod_{l=k+1}^{m_i}\Psi(\beta_n - \delta_i)}{\lambda_{ni}} \\ &= \frac{[\Psi(\beta_n - \delta_i)]^{m_i-k}\prod_{l=1}^{k}\Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}}, \quad k = 0,\ldots,m_i;\end{aligned} \tag{8}$$

where

$$\lambda_{ni} = \sum_{k=0}^{m_i}\{[\Psi(\beta_n - \delta_i)]^{m_i-k}\prod_{l=1}^{k}\Psi((m_i+1-l)\zeta_i)\}; \tag{9}$$

2.2 The JML verses MML estimation procedures.

Historically, the JML was used very early in the study of latent trait theory [20] (Birnbaum, 1968) with cumulative response models ([21]and [22]). The advantage of this procedure is that it is straightforward to implement with quick convergence and does not require any assumptions about the person distribution. The well known weakness of the JML is its inconsistency with a relatively small number of items [23].

The marginal maximum likelihood (MML) is another common procedure for the parameter estimation in Item Response Theory ([24]; [25] and [26]). In the context of unfolding, [9] used the MML with the EM algorithm in estimating the item parameters of the generalized graded unfolding model (GGUM), with person parameters in turn derived from an expected a posterior technique. On the other hand, [15] used the JML to estimate the parameters for the general unfolding models for dichotomous responses. In order to compensate for the effect of inconsistency, a correction procedure on the mean of the item units was proposed. The generalization of this procedure to the case of polytomous responses of (4) is one of the methods studied in this paper.

2.3 The JML Solution Equations.

For model (8), the joint likelihood function takes the form as

$$L = \prod_{n=1}^{N}\prod_{i=1}^{I}\Pr\{X_{ni} = x_{ni}\} = \prod_{n=1}^{N}\prod_{i=1}^{I}\frac{[\Psi(\beta_n - \delta_i)]^{m_i - x_{ni}}\prod_{l=1}^{x_{ni}}\Psi((m_i + 1 - l)\zeta_i)}{\lambda_{ni}}. \tag{10}$$

Then

$$\log L = \sum_{n=1}^{N}\sum_{i=1}^{I}\left\{(m_i - x_{ni})\log\Psi(\beta_n - \delta_i) + [\sum_{l=1}^{x_{ni}}\log\Psi((m_i + 1 - l)\zeta_i)] - \log\lambda_{ni}\right\}. \tag{11}$$

The solution equations derived in Appendix A are

$$\varphi_{\zeta_i} : \sum_{n=1}^{N}\left\{\sum_{l=1}^{x_{ni}}\frac{\partial\log\Psi((m_i + 1 - l)\zeta_i)}{\partial\zeta_i} - \sum_{k=0}^{m_i}\Pr\{X_{ni} = k\}\sum_{l=1}^{k}\frac{\partial\log\Psi((m_i + 1 - l)\zeta_i)}{\partial\zeta_i}\right\} = 0, \tag{12a}$$

$$\varphi_{\delta_i}: \sum_{n=1}^{N}\left\{-x_{ni}+E[X_{ni}]\right\}\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i}=0, \tag{12b}$$

$$\varphi_{\beta_n}: \sum_{i=1}^{I}\left\{-x_{ni}+E[X_{ni}]\right\}\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}=0. \tag{12c}$$

The conventional constraint on the item locations is applied as

$$\sum_{i=1}^{I}\delta_i=0 \tag{13}$$

When the maximum scores are $m_i=1$, the solution equations of (12) specialize to those derived in [15] for the general dichotomous unfolding models. The details of the solution equations for special cases are in Appendix B.

The derivation above assumes that there are no missing responses. That is, all persons give responses to all items. However, if person n does not give a response to item i, it can be handled by omitting the corresponding term involving x_{ni} in the log likelihood function of (11), but keeping the terms involving x_{nj}, $j \neq i$ when x_{nj} is present. Subsequently, the solution equation functions for the parameters of item i exclude person n, and the solution equation functions for the location parameter of person n excludes item i.

A conventional two-stage JML [20] (Birnbaum, 1968) algorithm can be used directly to solve the solution equations (12). [27] summarized the principle of two-stage JML estimation in the context of item response theory. In our situation, the principle of two-stage procedure can be applied to form an estimation cycle of $\{(\zeta_i), (\delta_i), (\beta_n)\}$. For example, when solving equation (12a) to estimate ζ_i within each cycle, the item location for the item δ_i and all the person locations (β_n) are fixed temporarily at their current (provisional) values estimated in the previous cycle. After solving (12b), the constraint of (13) is applied to adjust the estimates of (δ_i). In solving these equations sequentially using the procedure above, the expected values of the second order derivatives for the corresponding parameters are employed. These are derived in Appendix C:

$$E(\frac{\partial \varphi_{\zeta_i}}{\partial \zeta_i}) = -\sum_{n=1}^{N} \left\{ \left[\sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} [\sum_{l=1}^{K} \Delta(i,l)]^2 \right] - \left[\sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} [\sum_{l=1}^{k} \Delta(i,l)] \right]^2 \right\}$$

(14)

$$E(\frac{\partial \varphi_{\delta_i}}{\partial \delta_i}) = -\sum_{n=1}^{N} [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}]^2 \cdot \left\{ \left[\sum_{k=0}^{m_i} k^2 \Pr\{X_{ni} = k\} \right] - \left[\sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\} \right]^2 \right\}$$

(15)

$$E(\frac{\partial \varphi_{\beta_n}}{\partial \beta_n}) = \sum_{i=1}^{I} [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \cdot \left\{ \left[\sum_{k=0}^{m_i} k^2 \Pr\{X_{ni} = k\} \right] - \left[\sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\} \right]^2 \right\}$$

(16)

Where

$$\Delta(i,l) \equiv \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} \tag{17}$$

2.4 Information functions and standard errors.

In dichotomous cases of model (1) where $m_i = 1$, the information function for item i with respect to person n is given by

$$\begin{aligned} I_{ni} &= E[\frac{\partial^2 \log L}{\partial \beta_n^2}] \\ &= \pi_{ni}(1 - \pi_{ni})[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \\ &= (E(x_{ni}^2) - [E(x_{ni})]^2)[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2; \end{aligned} \tag{18}$$

Where

$$\pi_{ni} = \frac{\Psi(\rho_i)}{\Psi(\rho_i) + \Psi(\beta_n - \delta_i)} \tag{19}$$

Similarly, from (15), the information for item i on person *n* when $m_i \geq 1$ is

$$I_{ni} = \{E(x_{ni}^2) - [E(x_{ni})]^2\}[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}]^2. \tag{20}$$

Therefore, the information function for item i is

$$I_{\delta_i}(\beta_n) = \sum_{n=1}^{N} I_{ni} = \sum_{n=1}^{N} \{E(x_{ni}^2) - [E(x_{ni})]^2\}[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}]^2 ; \tag{21}$$

and the standard error for item i over all persons is

$$\sigma_{\delta_i} = 1/\sqrt{\sum_{n=1}^{N} I_{ni}} \tag{22}$$

Symmetrically, the standard error for person n over all items is

$$\sigma_{\beta_n} = 1/\sqrt{\sum_{i=1}^{I} I_{ni}} \tag{23}$$

The summation of the information function on all items $\sum_{i=1}^{I} I_{\delta_i}(\beta_n)$ is termed the *scale information function*.

2.5 Iteration Procedure A: with a correction on the estimate of mean unit.

The control of the inconsistency of the item parameters in procedure A involves a correction on the *mean unit* parameter for all items. Let ζ be the mean of the item unit parameters for all items:

$$\zeta = \sum_{i=1}^{I} \zeta_i / I \tag{24}$$

When the value of ζ is unknown, it can be estimated by solving (12) with the subsequent condition that all item unit parameters $\{\zeta_i\}$ have the mean value of ζ. A correction procedure to reduce the inconsistency on ζ is also employed:

$$\zeta_c = (1 - 2 / \sum_{i=1}^{I} m_i)\zeta \tag{25}$$

The heuristically simple correction formula could alleviate the effect of inconsistency greatly, as shown in the simulation studies in the next section.

In summary, the estimation procedure includes the following three steps:

Step 1. The value of mean unit parameter ζ is estimated by solving (12) with the constraint that the values of all item units equal the mean unit ζ;

$$\zeta_i = \zeta ,\ i = 1, ..., I. \tag{26}$$

Step 2. A correction factor (derived below) is applied on ζ, the estimate of ζ, giving the corrected value ζ_c.

Step 3. The parameters are re-estimated by solving (12) with the constraint (26) item unit parameter is released.

2.6 Iteration Procedure B: with the WLE for person locations.

Procedure B focuses on preventing the accumulation of the bias by using the WLE [17]. [18] gives an approximation of the bias function of the MLE as

$$Bias[MLE(\beta_n,\beta_n)] \cong E[\beta_n - \beta_n \mid \beta_n] \cong -\frac{1}{2[I(\beta_n)]^2}\sum_{i=1}^{I}\sum_{k=0}^{m_I}\frac{\frac{\partial P_{ni}(k)}{\partial \beta_n}\cdot\frac{\partial^2 P_{ni}(k)}{\partial \beta_n{}^2}}{P_{ni}(k)} \quad (28)$$

To prevent the bias, [19] Wang & Wang, (2001) proposed the solution equation as

$$\frac{\partial \log L}{\partial \beta_n} - I(\beta_n)Bias(\beta_n) = 0 \quad (29)$$

With the equi-distant unfolding model of (8), the MLE bias function is

$$Bias[MLE(\beta_n,\beta_n)] = -\frac{\sum_{i=1}^{I}\Delta(\beta_n-\delta_i)\left\{[\Delta(\beta_n-\delta_i)]^2 C_{ni}^3(\beta_n) + \frac{\partial \Delta(\beta_n-\delta_i)}{\partial \beta_n}V_{ni}^2(\beta_n)\right\}}{2[\sum_{i=1}^{I}[\Delta(\beta_n-\delta_i)]^2\cdot V_{ni}^2(\beta_n)]^2}; \quad (30)$$

Where

$$\Delta(\beta_n-\delta_i) = \frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial(\beta_n-\delta_i)}; \quad (31)$$

$$C_{ni}^3(\beta_n) \equiv \sum_{k=0}^{m_I}[k - E_{ni}(\beta_n)]^3 \Pr\{X_{ni} = k\}\}; \quad (32)$$

$$V_{ni}^2(\beta_n) \equiv \sum_{k=0}^{m_I}[k - E_{ni}(\beta_n)]^2 \cdot \Pr\{X_{ni} = k\}. \quad (33)$$

The derivation of (30) is in Appendix D. Hence, the alternative iteration Procedure B replaces the solution equation for person locations (12c) with (29), and no correction is placed on the item units.

2.7 Initial values.

At the beginning of the estimation in both procedures, the initial values of all parameters need to be specified. First, the values of (ζ_i) are set to zero. Second, t $\delta_i^{(0)}$ are calculated using (12b) when all β_n 's and ζ_i 's are set to 0. That is,

$$\frac{\partial \log \Psi(\delta_i)}{\partial \delta_i} \sum_{n=1}^{N} \left\{ -x_{ni} + \sum_{k=0}^{m_i} k \frac{\Psi(\delta_i)^{m_i-k}\Psi(0)^k}{\sum_{l=0}^{m_i} \Psi(\delta_i)^{m_i-l}\Psi(0)^l} \right\} = 0, \tag{34}$$

$$\sum_{k=0}^{m_i} k \frac{\Psi(\delta_i)^{m_i-k}\Psi(0)^k}{\sum_{l=0}^{m_i} \Psi(\delta_i)^{m_i-l}\Psi(0)^l} = \frac{1}{N}\sum_{n=1}^{N} x_{ni}, \tag{35}$$

$$\sum_{k=0}^{m_i} k\Psi(\delta_i)^{m_i-k}\Psi(0)^k = \left(\frac{1}{N}\sum_{n=1}^{N} x_{ni}\right)\left(\sum_{l=0}^{m_i} \Psi(\delta_i)^{m_i-l}\Psi(0)^l\right), \tag{36}$$

$$\sum_{k=0}^{m_i} \left(k - \frac{1}{N}\sum_{n=1}^{N} x_{ni}\right)\Psi(\delta_i)^{m_i-k}\Psi(0)^k = 0. \tag{37}$$

Equation (37) is a polynomial equation of $\Psi(\delta_i)$ and may be solved routinely. In particular, when the operational function is an exponential square, $\Psi(t) = \exp(t^2)$; and $m_i = 1$, (37) can be simply calculated by

$$\delta_i^{(0)} = \sqrt{\log\left(\frac{1-\mu_i}{\mu_i}\right)}; \tag{38}$$

Where

$$\mu_i = \frac{1}{N}\sum_{n=1}^{N} x_{ni} \tag{39}$$

When m_i=2,

$$\sum_{k=0}^{m_i}(k-\mu_i)[\exp(\delta_i^2)]^{m_i-k} \\ = -[\exp(\delta_i^2)]^2\mu_i + (1-\mu_i)[\exp(\delta_i^2)] + (2-\mu_i) = 0; \tag{40}$$

$$\begin{aligned}\exp(\delta_i^2) &= \frac{-(1-\mu_i) \pm \sqrt{(1-\mu_i)^2 - 4[-\mu_i](2-\mu_i)}}{2\cdot[-\mu_i]} \\ &= \frac{(1-\mu_i) \pm \sqrt{(1-\mu_i)^2 + 4\mu_i(2-\mu_i)}}{2\mu_i} \\ &= \frac{1-\mu_i}{2\mu_i} \pm \frac{\sqrt{(1-\mu_i)^2 + 4\mu_i(2-\mu_i)}}{2\mu_i} \\ &\approx \frac{2-\mu_i}{2\mu_i}.\end{aligned} \tag{41}$$

In general, the initial value can be obtained by

$$\delta_i^{(0)} = \pm\sqrt{\log(\frac{m_i-\mu_i}{m_i\mu_i})} \tag{42}$$

The sign (plus or minus) can be assigned conceptually or by the *sign analysis* proposed in [16].

In addition, the initial value for the person location β_n is calculated using the *mid-point principle* ([1] and [28]), but applied to polytomous responses according to

$$\beta_n^{(0)} = \frac{1}{r_n}\sum_{i=1}^{I} x_{ni}\delta_i^{(0)} \tag{43}$$

where

$$r_n = \sum_{i=1}^{I} x_{ni} \tag{44}$$

3 Summary and conclusion

The algorithm presented in this paper provides another example of the operationalization of unfolding models for statements with polytomous response formats. Though more investigation on parameter estimation would help (a comparison study on the estimation result of JML and MML, for example), the studies presented in this paper show that this operationalization efficiently recovers the parameter values of the model with satisfactory accuracy. *RateFold*, a program developed with this algorithm is available from the author of this paper.

Appendix A. Derivation of solution equation

A.1. Solution Equations for ζ_i.

From (11), the partial derivative of $\log L$ with respect to ζ_i is

$$\frac{\partial \log L}{\partial \zeta_i} = \sum_{n=1}^{N} \left\{ [\sum_{l=1}^{x_{ni}} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i}] - \frac{\partial \log \lambda_{ni}}{\partial \zeta_i} \right\}. \tag{A1}$$

Because

$$\frac{\partial \log \lambda_{ni}}{\partial \zeta_i} = \frac{1}{\lambda_{ni}} \frac{\partial}{\partial \zeta_i} \{\sum_{k=0}^{m_i} [\Psi(\beta_n - \delta_i)]^{m_i - k} \prod_{l=1}^{k} \Psi((m_i + 1 - l)\zeta_i)\}$$

$$= \frac{1}{\lambda_{ni}} \{\sum_{k=0}^{m_i} [\Psi(\beta_n - \delta_i)]^{m_i - k} \frac{\partial}{\partial \zeta_i} \prod_{l=1}^{k} \Psi((m_i + 1 - l)\zeta_i)\}$$

$$= \frac{1}{\lambda_{ni}} \{\sum_{k=0}^{m_i} [\Psi(\beta_n - \delta_i)]^{m_i - k}][\prod_{l=1}^{k} \Psi((m_i + 1 - l)\zeta_i)\}][\sum_{l=1}^{k} \frac{\frac{\partial}{\partial \zeta_i} \Psi((m_i + 1 - l)\zeta_i)}{\Psi((m_i + 1 - l)\zeta_i)}]\}$$

$$= \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} \sum_{l=1}^{k} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i}, \tag{A2}$$

Then the solution equations for ζ_i is

$$\frac{\partial \log L}{\partial \zeta_i} = \sum_{n=1}^{N} \left\{ \sum_{l=1}^{x_{ni}} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} - \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} \sum_{l=1}^{k} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} \right\} = 0. \tag{A3}$$

A.2. Solution Equations for δ_i.

From (11), the partial derivative of $\log L$ with respect to δ_i is

$$\frac{\partial \log L}{\partial \delta_i} = \sum_{n=1}^{N} \left\{ (m_i - x_{ni}) \frac{\partial \log \Psi(\delta_i - \beta_n)}{\partial \delta_i} - \frac{\partial \log \lambda_{ni}}{\partial \delta_i} \right\}. \tag{A4}$$

Because

$$\frac{\partial \log \lambda_{ni}}{\partial \delta_i} = \frac{1}{\lambda_{ni}} \frac{\partial}{\partial \delta_i} [\sum_{k=0}^{m_i} \{[\Psi(\beta_n - \delta_i)]^{m_i - k} \prod_{l=1}^{k} \Psi((m+1-l)\zeta_i)\}]$$

$$= \frac{1}{\lambda_{ni}} \sum_{k=0}^{m_i} \{[\frac{\partial}{\partial \delta_i} [\Psi(\beta_n - \delta_i)]^{m_i - k}] \prod_{l=1}^{k} \Psi((m+1-l)\zeta_i)\}$$

$$= \frac{1}{\lambda_{ni}} \sum_{k=0}^{m_i} \{(m_i - k)[\Psi(\beta_n - \delta_i)]^{m_i - k - 1} [\frac{\partial}{\partial \delta_i} [\Psi(\beta_n - \delta_i)] \prod_{l=1}^{k} \Psi((m+1-l)\zeta_i)\}$$

$$= \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} (m_i - k) \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i},$$

(A5)

(A4) can be simplified to

$$\frac{\partial \log L}{\partial \delta_i} = \sum_{n=1}^{N} \left\{ (m_i - x_{ni}) - \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} (m_i - k) \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}$$

$$= \sum_{n=1}^{N} \left\{ (m_i - x_{ni}) - m_i \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} + \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\} \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}.$$

(A6)

Because

$$\sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} = 1; \qquad \text{(A7)}$$

it leads to the solution equations for δ_i as

$$\frac{\partial \log L}{\partial \delta_i} = \sum_{n=1}^{N} \left\{ -x_{ni} + \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\} \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}$$

$$= \sum_{n=1}^{N} \left\{ -x_{ni} + E[X_{ni}] \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}. \qquad \text{(A8)}$$

A.3. Solution Equations for β_n .

From (11), the partial derivative of $\log L$ with respect to β_n is

$$\frac{\partial \log L}{\partial \beta_n} = \sum_{i=1}^{I}\left\{(m_i - x_{ni})\frac{\partial \log \Psi(\delta_i - \beta_n)}{\partial \beta_n} - \frac{\partial \log \lambda_{ni}}{\partial \beta_n}\right\}. \tag{A9}$$

Because that

$$\begin{aligned}\frac{\partial \log \lambda_{ni}}{\partial \beta_n} &= \frac{1}{\lambda_{ni}}\frac{\partial}{\partial \beta_n}[\sum_{k=0}^{m_i}\{[\Psi(\beta_n - \delta_i)]^{m_i-k}\prod_{l=1}^{k}\Psi((m+1-l)\zeta_i)\}] \\ &= \frac{1}{\lambda_{ni}}\sum_{k=0}^{m_i}\{[\frac{\partial}{\partial \beta_n}[\Psi(\beta_n - \delta_i)]^{m_i-k}]\prod_{l=1}^{k}\Psi((m+1-l)\zeta_i)\} \\ &= \frac{1}{\lambda_{ni}}\sum_{k=0}^{m_i}\{(m_i-k)[\Psi(\beta_n - \delta_i)]^{m_i-k-1}[\frac{\partial}{\partial \beta_n}[\Psi(\beta_n - \delta_i)]]\prod_{l=1}^{k}\Psi((m+1-l)\zeta_i)\} \\ &= \sum_{k=0}^{m_i}\Pr\{X_{ni} = k\}(m_i - k)\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n};\end{aligned} \tag{A10}$$

(A9) can be written as

$$\begin{aligned}\frac{\partial \log L}{\partial \beta_n} &= \sum_{i=1}^{I}\left\{(m_i - x_{ni}) - \sum_{k=0}^{m_i}\Pr\{X_{ni} = k\}(m_i - k)\right\}\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \\ &= \sum_{i=1}^{I}\left\{(m_i - x_{ni}) - m_i\sum_{k=0}^{m_i}\Pr\{X_{ni} = k\} + \sum_{k=0}^{m_i}k\Pr\{X_{ni} = k\}\right\}\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}.\end{aligned} \tag{A11}$$

(A9) can be further simplified to

$$\begin{aligned}\frac{\partial \log L}{\partial \beta_n} &= \sum_{i=1}^{I}\left\{-x_{ni} + \sum_{k=0}^{m_i}k\Pr\{X_{ni} = k\}\right\}\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \\ &= \sum_{i=1}^{I}\left\{-x_{ni} + E[X_{ni}]\right\}\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}.\end{aligned} \tag{A12}$$

In summary, the solution equations are

$$\varphi_{\zeta_i} : \sum_{n=1}^{N} \left\{ \sum_{l=1}^{x_{ni}} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} - \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} \sum_{l=1}^{k} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} \right\} = 0, \quad \text{(A13a)}$$

$$\varphi_{\delta_i} : \sum_{n=1}^{N} \left\{ -x_{ni} + E[X_{ni}] \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i} = 0, \quad \text{(A13b)}$$

$$\varphi_{\beta_n} : \sum_{i=1}^{I} \left\{ -x_{ni} + E[X_{ni}] \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} = 0; \quad \text{(A13c)}$$

with the conventional constraint on item locations

$$\sum_{i=1}^{I} \delta_i = 0. \quad \text{(A13d)}$$

In addition, let

$$\Delta(i,l) \equiv \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} \quad \text{(A14)}$$

(A13a) then can be expressed in terms of $\Delta(i,l)$ as

$$\varphi_{\zeta_i} : \sum_{n=1}^{N} \left\{ \sum_{l=1}^{x_{ni}} \Delta(i,l) - \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} \sum_{l=1}^{k} \Delta(i,l) \right\} = 0. \quad \text{(A15)}$$

(A14) and (A15) will be used in Appendix B.

Appendix B. Solution equations for special cases

B.1. Special case (I)-**SSLMP**:

$$\Psi(t) = \exp(t^2) \tag{B1}$$

In this case

$$\begin{aligned}\Delta(i,l) &\equiv \frac{\partial \log \Psi((m_i+1-l)\zeta_i)}{\partial \zeta_i} = \frac{\partial \log \exp((m_i+1-l)^2\zeta_i^2)}{\partial \zeta_i} \\ &= (m_i+1-l)^2 \frac{\partial \zeta_i^2}{\partial \zeta_i} \\ &= (m_i+1-l)^2 \cdot 2\zeta_i;\end{aligned} \tag{B2}$$

$$\begin{aligned}\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} &= \frac{\partial(\beta_n-\delta_i)^2}{\partial \delta_i} \\ &= -2(\beta_n-\delta_i);\end{aligned} \tag{B3}$$

And

$$\begin{aligned}\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n} &= \frac{\partial(\beta_n-\delta_i)^2}{\partial \beta_n} \\ &= 2(\beta_n-\delta_i).\end{aligned} \tag{B4}$$

The solution equations are

$$\begin{aligned}\frac{\partial \log L}{\partial \zeta_i} &= \sum_{n=1}^{N}\left\{\sum_{l=1}^{x_{ni}}(m_i+1-l)^2 \cdot 2\zeta_i - \sum_{k=0}^{m_i}[\Pr\{X_{ni}=k\}\sum_{l=1}^{k}(m_i+1-l)^2 \cdot 2\zeta_i]\right\} \\ &= 2\zeta_i \sum_{n=1}^{N}\left\{\sum_{l=1}^{x_{ni}}(m_i+1-l)^2 - \sum_{k=0}^{m_i}[\Pr\{X_{ni}=k\}\sum_{l=1}^{k}(m_i+1-l)^2]\right\} \\ &= 2\zeta_i \sum_{n=1}^{N}\left\{\sum_{l=1}^{x_{ni}}(m_i+1-l)^2 - \sum_{k=0}^{m_i}[\Pr\{X_{ni}=k\}\sum_{l=1}^{k}(m_i+1-l)^2]\right\};\end{aligned}$$

(B5a)

$$\begin{aligned}\frac{\partial \log L}{\partial \delta_i} &= \sum_{n=1}^{N}\left\{-x_{ni} + \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right\} - 2(\beta_n - \delta_i) \\ &= 2\sum_{n=1}^{N}\left\{x_{ni} - \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right\}(\beta_n - \delta_i);\end{aligned} \tag{B5b}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \beta_n} &= \sum_{i=1}^{I}\left\{-x_{ni} + \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right\} 2(\beta_n - \delta_i) \\ &= 2\sum_{i=1}^{I}\left\{-x_{ni} + \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right\}(\beta_n - \delta_i).\end{aligned} \tag{B5c}$$

Because that $\zeta_i > 0$, the solution equations can be simplified to

$$\varphi_{\zeta_i} : \sum_{n=1}^{N}\left\{\sum_{l=1}^{x_{ni}}(m_i + 1 - l)^2 - \sum_{k=0}^{m_i}[\Pr\{X_{ni} = k\}\sum_{l=1}^{k}(m_i + 1 - l)^2]\right\} = 0 \tag{B6a}$$

$$\varphi_{\delta_i} : \sum_{n=1}^{N}\left\{x_{ni} - E[X_{ni}]\right\}(\beta_n - \delta_i) = 0 \tag{B6b}$$

$$\varphi_{\beta_n} : \sum_{i=1}^{I}\left\{-x_{ni} + E[X_{ni}]\right\}(\beta_n - \delta_i) = 0 \tag{B6c}$$

B.2. Special case (II)-**HCMP**:

$$\Psi(t) = \cosh(t) \tag{B7}$$

In this case,

$$\Delta(i,l) \equiv \frac{\partial \log \Psi((m_i+1-l)\zeta_i)}{\partial \zeta_i} = \frac{\partial \log \cosh((m_i+1-l)\zeta_i)}{\partial \zeta_i}$$
$$= \frac{1}{\cosh((m_i+1-l)\zeta_i)} \cdot \frac{\partial \cosh((m_i+1-l)\zeta_i)}{\partial \zeta_i} \quad \text{(B8)}$$
$$= (m_i+1-l)\tanh((m_i+1-l)\zeta_i);$$

$$\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} = \frac{\partial \log \cosh(\beta_n-\delta_i)}{\partial \delta_i}$$
$$= \frac{1}{\cosh(\beta_n-\delta_i)} \cdot \frac{\partial \cosh(\beta_n-\delta_i)}{\partial \delta_i} \quad \text{(B9)}$$
$$= -\tanh(\beta_n-\delta_i);$$

and

$$\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n} = \tanh(\beta_n-\delta_i). \quad \text{(B10)}$$

The solution equations are

$$\varphi_{\zeta_i} : \sum_{n=1}^{N} \left\{ \sum_{l=1}^{x_{ni}} (m_i+1-l)\tanh((m_i+1-l)\zeta_i) - \sum_{k=0}^{m_i} \Pr\{X_{ni}=k\} \sum_{l=1}^{k} (m_i+1-l)\tanh((m_i+1-l)\zeta_i) \right\} = 0,$$
(B11a)

$$\varphi_{\delta_i} : -\sum_{n=1}^{N} \tanh(\beta_n-\delta_i) \left\{ -x_{ni} + E[X_{ni}] \right\} = 0, \quad \text{(B11b)}$$

$$\varphi_{\beta_n} : \sum_{i=1}^{I} \tanh(\beta_n-\delta_i) \left\{ -x_{ni} + E[X_{ni}] \right\} = 0. \quad \text{(B11c)}$$

Appendix C. Calculation of the expected values of second derivatives.

C.1. The expected values of second derivatives for item unit ζ_i.

According to (12a),

$$\frac{\partial \varphi_{\zeta_i}}{\partial \zeta_i} = \sum_{n=1}^{N} \left\{ \sum_{l=1}^{x_{ni}} \frac{\partial}{\partial \zeta_i} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} - \sum_{k=0}^{m_i} \frac{\partial \Pr\{X_{ni} = k\}}{\partial \zeta_i} \sum_{l=1}^{k} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} - \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} \sum_{l=1}^{k} \frac{\partial}{\partial \zeta_i} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} \right\}. \tag{C1}$$

It is evident that

$$\sum_{l=1}^{x_{ni}} \frac{\partial}{\partial \zeta_i} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} = \sum_{k=0}^{m_i} \Pr\{X_{ni} = k\} \sum_{l=1}^{k} \frac{\partial}{\partial \zeta_i} \frac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i} \Bigg\}. \tag{C2}$$

Let $\Delta(i,l) \equiv \dfrac{\partial \log \Psi((m_i + 1 - l)\zeta_i)}{\partial \zeta_i}$. Then

$$E\left(\frac{\partial \varphi_{\zeta_i}}{\partial \zeta_i}\right) = \sum_{n=1}^{N} \left\{ - \sum_{k=0}^{m_i} \frac{\partial \Pr\{X_{ni} = k\}}{\partial \zeta_i} \sum_{l=1}^{k} \Delta(i,l) \right\}. \tag{C3}$$

and

$$\frac{\partial \Pr\{X_{ni} = k\}}{\partial \zeta_i} = [\Psi(\beta_n - \delta_i)]^{m_i - k} \frac{\partial}{\partial \zeta_i} \frac{\prod_{l=1}^{k} \Psi((m_i + 1 - l)\zeta_i)}{\lambda_{ni}} \tag{C4}$$

$$\frac{\partial}{\partial \zeta_i} \frac{\prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}} = \frac{\lambda_{ni} \frac{\partial}{\zeta_i} \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) - \frac{\partial \lambda_{ni}}{\partial \zeta_i} \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}^{\ 2}}. \tag{C5}$$

It is evident that

$$\frac{\partial}{\partial \zeta_i} \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) = (\prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i)) \sum_{l=1}^{K} \Delta(i,l) \tag{C6}$$

$$\begin{aligned} \frac{\partial \lambda_{ni}}{\partial \zeta_i} &= \frac{\partial}{\partial \zeta_i} \sum_{k=0}^{m_i} \left\{ [\Psi(\beta_n - \delta_i)]^{m_i-k} \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) \right\} \\ &= \sum_{k=0}^{m_i} \left\{ [\Psi(\beta_n - \delta_i)]^{m_i-k} \frac{\partial}{\partial \zeta_i} \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) \right\} \\ &= \sum_{k=0}^{m_i} \left\{ [\Psi(\beta_n - \delta_i)]^{m_i-k} \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) \sum_{l=1}^{K} \Delta(i,l) \right\} \end{aligned} \tag{C7}$$

From (C6) and (C7), equation (C5) can be rewritten a

$$\begin{aligned} &\frac{\partial}{\partial \zeta_i} \frac{\prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}} \\ &= \frac{\prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i)) \sum_{l=1}^{K} \Delta(i,l)}{\lambda_{ni}} \\ &\quad - \frac{\{\sum_{j=0}^{m_i} \left\{ [\Psi(\beta_n - \delta_i)]^{m_i-j} \prod_{l=1}^{j} \Psi((m_i+1-l)\zeta_i) \sum_{l=1}^{j} \Delta(i,l) \right\}\} \prod_{l=1}^{j} \Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}^{\ 2}} \\ &= \frac{\prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}} \sum_{l=1}^{K} \Delta(i,l) - \frac{\{\sum_{j=0}^{m_i} \left\{ \Pr\{X_{ni}=j\} \sum_{l=1}^{j} \Delta(i,l) \right\} \prod_{l=1}^{j} \Psi((m_i+1-l)\zeta_i)}{\lambda_{ni}}. \end{aligned} \tag{C8}$$

Therefore, (C4) can be simplified as

$$\frac{\partial \Pr\{X_{ni}=k\}}{\partial \zeta_i} = \Pr\{X_{ni}=k\}\left\{\sum_{l=1}^{K}\Delta(i,l) - \sum_{j=0}^{m_i}[\Pr\{X_{ni}=j\}\sum_{l=1}^{j}\Delta(i,l)]\right\}$$

(C9)

Finally

$$E(\frac{\partial \varphi_{\zeta_i}}{\partial \zeta_i}) = -\sum_{n=1}^{N}\left\{\sum_{k=0}^{m_i}\Pr\{X_{ni}=k\}[\sum_{l=1}^{K}\Delta(i,l) - \sum_{j=0}^{m_i}\{\Pr\{X_{ni}=j\}\sum_{l=1}^{j}\Delta(i,l)]\sum_{l=1}^{k}\Delta(i,l)\right\}$$

$$= -\sum_{n=1}^{N}\left\{\Big[\sum_{k=0}^{m_i}\Pr\{X_{ni}=k\}[\sum_{l=1}^{K}\Delta(i,l)][\sum_{l=1}^{k}\Delta(i,l)]\Big] - \right.$$

$$\left. -\Big[\sum_{k=0}^{m_i}\Pr\{X_{ni}=k\}[\sum_{j=0}^{m_i}\{\Pr\{X_{ni}=j\}\sum_{l=1}^{j}\Delta(i,l)]\sum_{l=1}^{k}\Delta(i,l)\Big]\right\}$$

$$= -\sum_{n=1}^{N}\left\{\Big[\sum_{k=0}^{m_i}\Pr\{X_{ni}=k\}[\sum_{l=1}^{K}\Delta(i,l)]^2\Big] - \Big[\sum_{k=0}^{m_i}\Pr\{X_{ni}=k\}[\sum_{l=1}^{k}\Delta(i,l)]\Big]^2\right\}.$$

(C10)

C.2. The expected values of second derivatives for item location δ_i.

According to (12b),

$$\frac{\partial \varphi_{\delta_i}}{\partial \delta_i}$$

$$= \frac{\partial}{\partial \delta_i}\sum_{n=1}^{N}\left\{x_{ni} - \sum_{k=0}^{m_i} k\Pr\{X_{ni}=k\}\right\}\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \delta_i}$$

$$= \sum_{n=1}^{N}\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \delta_i}\frac{\partial}{\partial \delta_i}\left\{x_{ni} - \sum_{k=0}^{m_i} k\Pr\{X_{ni}=k\}\right\} + \left\{x_{ni} - \sum_{k=0}^{m_i} k\Pr\{X_{ni}=k\}\right\}\frac{\partial}{\partial \delta_i}\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \delta_i}$$

$$= \sum_{n=1}^{N}\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \delta_i}\left\{\sum_{k=0}^{m_i} k\frac{\partial \Pr\{X_{ni}=k\}}{\partial \delta_i}\right\} + \left\{x_{ni} - \sum_{k=0}^{m_i} k\Pr\{X_{ni}=k\}\right\}\frac{\partial}{\partial \delta_i}\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \delta_i}.$$

(C11)

Because

$$\frac{\partial \Pr\{X_{ni}=k\}}{\partial \delta_i} = \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) \cdot \frac{\partial}{\partial \delta_i} \frac{[\Psi(\beta_n-\delta_i)]^{m_i-k}}{\lambda_{ni}}$$

$$= \prod_{l=1}^{k} \Psi((m_i+1-l)\zeta_i) \cdot$$

$$\frac{\lambda_{ni}(m_i-k)[\Psi(\beta_n-\delta_i)]^{m_i-k-1}\frac{\partial[\Psi(\beta_n-\delta_i)]}{\partial \delta_i} - [\Psi(\beta_n-\delta_i)]^{m_i-k}\frac{\partial[\lambda_{ni}]}{\partial \delta_i}}{\lambda_{ni}{}^2}$$

$$= \Pr\{X_{ni}=k\} \cdot \left\{ (m_i-k)\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} - \frac{\partial \log \lambda_{ni}}{\partial \delta_i} \right\}$$

(C12)

and

$$\frac{\partial \log \lambda_{ni}}{\partial \delta_i} = \frac{1}{\lambda_{ni}} \sum_{j=0}^{m_i} \Psi((m_i+1-l)\zeta_i)(m_i-j)[\Psi(\beta_n-\delta_i)]^{m_i-j-1} \frac{\partial \Psi(\beta_n-\delta_i)}{\partial \delta_i}$$

$$= \sum_{j=0}^{m_i} \Pr\{X_{ni}=j\}(m_i-j)\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i};$$

(C13)

$$\frac{\partial \Pr\{X_{ni}=k\}}{\partial \delta_i} = \Pr\{X_{ni}=k\} \cdot \left\{ (m_i-k)\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} - \sum_{j=0}^{m_i} \Pr\{X_{ni}=j\}(m_i-j)\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} \right\}$$

$$= \Pr\{X_{ni}=k\} \cdot \frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} \cdot \left\{ (m_i-k) - \sum_{j=0}^{m_i} \Pr\{X_{ni}=j\}(m_i-j) \right\}$$

$$= \Pr\{X_{ni}=k\} \cdot \frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} \cdot \left\{ (m_i-k) - m_i + \sum_{j=0}^{m_i} j \Pr\{X_{ni}=j\} \right\}$$

$$= \Pr\{X_{ni}=k\} \cdot \frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \delta_i} \cdot \left\{ -k + \sum_{j=0}^{m_i} j \Pr\{X_{ni}=j\} \right\}.$$

(C14)

And because that

$$E(x_{ni}) = \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}; \tag{C15}$$

the expected values of second derivatives for item location δ_i is

$$\begin{aligned} E(\frac{\partial \varphi_{\delta_i}}{\partial \delta_i}) &= \sum_{n=1}^{N} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i} \left\{ \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\} \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i} \cdot [-k + \sum_{j=0}^{m_i} j \Pr\{X_{ni} = j\}] \right\} \\ &= \sum_{n=1}^{N} [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}]^2 \cdot \left\{ \sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\} \cdot [-k + \sum_{j=0}^{m_i} j \Pr\{X_{ni} = j\}] \right\} \\ &= \sum_{n=1}^{N} [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \delta_i}]^2 \cdot \left\{ \left[\sum_{k=0}^{m_i} k^2 \Pr\{X_{ni} = k\}\right] - \left[\sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right]^2 \right\}. \end{aligned}$$

(C16)

C.3. The expected values of second derivatives for item location β_n.

Similar to the derivation in C.2,

$$E(\frac{\partial \varphi_{\beta_n}}{\partial \beta_n}) = \sum_{i=1}^{I} [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \cdot \left\{ \left[\sum_{k=0}^{m_i} k^2 \Pr\{X_{ni} = k\}\right] - \left[\sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right]^2 \right\}.$$

(C17)

Appendix D. Weighted likelihood estimation of person locations within the model of equation (8)

The MLE bias function for general discrete responses is [18]

$$Bias[MLE(\beta_n,\beta_n)] \cong E[\beta_n - \beta_n \mid \beta_n] \cong -\frac{1}{2[I(\beta_n)]^2}\sum_{i=1}^{I}\sum_{k=0}^{m_I}\frac{\frac{\partial P_{ni}(k)}{\partial \beta_n}\cdot\frac{\partial^2 P_{ni}(k)}{\partial {\beta_n}^2}}{P_{ni}(k)}$$

(D1)

The solution equation is

$$\frac{\partial \log L}{\partial \beta_n} - I(\beta_n) Bias(\beta_n) = 0 \qquad \text{(D2)}$$

From (19), information function for person n is

$$I(\beta_n) = \sum_{i=1}^{I} I_{ni} = \sum_{i=1}^{I}\{E(x_{ni}^2) - [E(x_{ni})]^2\}[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \qquad \text{(D3)}$$

In addition,

$$\begin{aligned}\frac{\partial^2 P_{ni}(k)}{\partial {\beta_n}^2} &= \frac{\partial}{\partial \beta_n}[\Pr\{X_{ni}=k\}\cdot\left\{k - E_{ni}(\beta_n)\right\}\cdot\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \beta_n}] \\ &= \frac{\partial \Pr\{X_{ni}=k\}}{\partial \beta_n}\cdot\left\{k - E_{ni}(\beta_n)\right\}\cdot\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \beta_n} \\ &+ \Pr\{X_{ni}=k\}\cdot\frac{\partial\left\{k - E_{ni}(\beta_n)\right\}}{\partial \beta_n}\cdot\frac{\partial \log\Psi(\beta_n-\delta_i)}{\partial \beta_n} \\ &+ \Pr\{X_{ni}=k\}\cdot\left\{k - E_{ni}(\beta_n)\right\}\cdot\frac{\partial^2 \log\Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}\end{aligned} \qquad \text{(D4)}$$

Let

$$V_{ni}^2(\beta_n) \equiv \left\{ \left[\sum_{k=0}^{m_i} k^2 \Pr\{X_{ni} = k\}\right] - \left[\sum_{k=0}^{m_i} k \Pr\{X_{ni} = k\}\right]^2 \right\}$$
$$= \sum_{k=0}^{m_I} [k - E_{ni}(\beta_n)]^2 \cdot \Pr\{X_{ni} = k\}. \quad \text{(D5)}$$

It is evident that

$$\frac{\partial \left\{k - E_{ni}(\beta_n)\right\}}{\partial \beta_n} = -\frac{\partial \left\{\sum_{v=0}^{m_i} v P_{ni}(v)\right\}}{\partial \beta_n}$$
$$= -\sum_{v=0}^{m_i} v \frac{\partial P_{ni}(v)}{\partial \beta_n} \quad \text{(D6)}$$
$$= -\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \cdot V_{ni}^2(\beta_n).$$

Therefore

$$\frac{\partial^2 P_{ni}(k)}{\partial \beta_n{}^2} = \Pr\{X_{ni} = k\} \cdot \left\{k - E_{ni}(\beta_n)\right\} \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \cdot \left\{k - E_{ni}(\beta_n)\right\} \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}$$
$$- \Pr\{X_{ni} = k\} \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \cdot V_{ni}^2(\beta_n) \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}$$
$$+ \Pr\{X_{ni} = k\} \cdot \left\{k - E_{ni}(\beta_n)\right\} \cdot \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n}$$
$$= \Pr\{X_{ni} = k\} \left\{ [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \left\{[k - E_{ni}(\beta_n)]^2 - V_{ni}^2(\beta_n)\right\} + [k - E_{ni}(\beta_n)] \cdot \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n} \right\}$$
$$\text{(D7)}$$

And then

$$\frac{\frac{\partial P_{ni}(k)}{\partial \beta_n}\frac{\partial^2 P_{ni}(k)}{\partial \beta_n{}^2}}{P_{ni}(k)}$$
$$= [k - E_{ni}(\beta_n)] \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \cdot \Pr\{X_{ni} = k\}$$
$$\bullet \left\{ [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \left\{ [k - E_{ni}(\beta_n)]^2 - V_{ni}^2(\beta_n) \right\} + [k - E_{ni}(\beta_n)] \cdot \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n} \right\}$$
$$= [k - E_{ni}(\beta_n)]^3 [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^3 \Pr\{X_{ni} = k\}$$
$$- [k - E_{ni}(\beta_n)][\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^3 V_{ni}^2(\beta_n) \cdot \Pr\{X_{ni} = k\}$$
$$+ [k - E_{ni}(\beta_n)]^2 \cdot \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n} \cdot \Pr\{X_{ni} = k\}$$

(D8)

It is evident that

$$\sum_{k=0}^{m_I} [k - E_{ni}(\beta_n)]^3 [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^3 \Pr\{X_{ni} = k\}$$
$$= [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^3 \sum_{k=0}^{m_I} [k - E_{ni}(\beta_n)]^3 \Pr\{X_{ni} = k\}; \quad \text{(D9)}$$

and because that

$$\sum_{k=0}^{m_I} P_{ni}(k)[k - E_{ni}(\beta_n)] = \sum_{k=0}^{m_I} P_{ni}(k)k - E_{ni}(\beta_n) \sum_{k=0}^{m_I} P_{ni}(k)$$
$$= E_{ni}(\beta_n) - E_{ni}(\beta_n) = 0; \quad \text{(D10)}$$

So,

$$\sum_{k=0}^{m_I}[k-E_{ni}(\beta_n)][\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}]^3 V_{ni}^2(\beta_n)\cdot \Pr\{X_{ni}=k\}$$
$$=[\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}]^3 V_{ni}^2(\beta_n)\sum_{k=0}^{m_I}[k-E_{ni}(\beta_n)]\cdot \Pr\{X_{ni}=k\} \quad \text{(D11)}$$
$$=0;$$

In addition,

$$\sum_{k=0}^{m_I}[k-E_{ni}(\beta_n)]^2\cdot \frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}\frac{\partial^2 \log \Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}\cdot \Pr\{X_{ni}=k\}=$$
$$=\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}\frac{\partial^2 \log \Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}\sum_{k=0}^{m_I}[k-E_{ni}(\beta_n)]^2\cdot \Pr\{X_{ni}=k\}$$
$$=\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}\frac{\partial^2 \log \Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}V_{ni}^2(\beta_n)$$

(D12)

Therefore

$$\sum_{k=0}^{m_I}\frac{\frac{\partial P_{ni}(k)}{\partial \beta_n}\frac{\partial^2 P_{ni}(k)}{\partial {\beta_n}^2}}{P_{ni}(k)}$$
$$=[\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}]^3\sum_{k=0}^{m_I}[k-E_{ni}(\beta_n)]^3\Pr\{X_{ni}=k\}$$
$$+\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}\frac{\partial^2 \log \Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}V_{ni}^2(\beta_n)$$
$$=\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}\left\{[\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}]^2\sum_{k=0}^{m_I}[k-E_{ni}(\beta_n)]^3\Pr\{X_{ni}=k\}\}+\frac{\partial^2 \log \Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}V_{ni}^2(\beta_n)\right\}$$
$$=\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}\left\{[\frac{\partial \log \Psi(\beta_n-\delta_i)}{\partial \beta_n}]^2 C_{ni}^3(\beta_n)+\frac{\partial^2 \log \Psi(\beta_n-\delta_i)}{\partial \beta^2{}_n}V_{ni}^2(\beta_n)\right\}$$

(D13)

where

$$C_{ni}^{3}(\beta_n) \equiv \sum_{k=0}^{m_I} [k - E_{ni}(\beta_n)]^3 \Pr\{X_{ni} = k\}\} \qquad \text{(D14)}$$

The MLE bias function then becomes

$$Bias[MLE(\beta_n, \beta_n)] \cong E[\beta_n - \beta_n \mid \beta_n]$$

$$\cong -\frac{1}{2[I(\beta_n)^2]} \sum_{i=1}^{I} \sum_{k=0}^{m_I} \frac{\frac{\partial P_{ni}(k)}{\partial \beta_n} \frac{\partial^2 P_{ni}(k)}{\partial {\beta_n}^2}}{P_{ni}(k)}$$

$$= -\frac{\sum_{i=1}^{I} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}] \left\{ [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 C_{ni}^{3}(\beta_n) + \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n} V_{ni}^{2}(\beta_n) \right\}}{2[\sum_{i=1}^{I} [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \cdot V_{ni}^{2}(\beta_n)]^2}.$$

(D15)

The solution equation is

$$\frac{\partial \log L}{\partial \beta_n} - I(\beta_n) Bias(\beta_n)$$
$$= \sum_{i=1}^{I} \left\{ -x_{ni} + E[X_{ni}] \right\} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} - I(\beta_n) Bias(\beta_n) \qquad \text{(D16)}$$
$$= 0$$

Special case (I)-SSLMP

$Bias[MLE(\beta_n, \beta_n)]$

$$\cong -\frac{\sum_{i=1}^{I} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}] \left\{ [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 C_{ni}^3(\beta_n) + \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n} V_{ni}^2(\beta_n) \right\}}{2[\sum_{i=1}^{I}[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \cdot V_{ni}^2(\beta_n)]^2}$$

$$= -\frac{\sum_{i=1}^{I} 2(\beta_n - \delta_i) \left\{ 4[(\beta_n - \delta_i)]^2 C_{ni}^3(\beta_n) + 2V_{ni}^2(\beta_n) \right\}}{2[\sum_{i=1}^{I}[2(\beta_n - \delta_i)]^2 \cdot V_{ni}^2(\beta_n)]^2}$$

$$= -\frac{8\sum_{i=1}^{I}(\beta_n - \delta_i)^3 C_{ni}^3(\beta_n) + 2\sum_{i=1}^{I}(\beta_n - \delta_i) V_{ni}^2(\beta_n)}{2[\sum_{i=1}^{I}[2(\beta_n - \delta_i)]^2 \cdot V_{ni}^2(\beta_n)]^2}$$

$$= -\frac{4\sum_{i=1}^{I}(\beta_n - \delta_i)^3 C_{ni}^3(\beta_n) + \sum_{i=1}^{I}(\beta_n - \delta_i) V_{ni}^2(\beta_n)}{[\sum_{i=1}^{I}[2(\beta_n - \delta_i)]^2 \cdot V_{ni}^2(\beta_n)]^2}.$$

(D17)

Special case (II)-**HCMP**:

$Bias[MLE(\beta_n, \beta_n)]$

$$\cong -\frac{\sum_{i=1}^{I} \frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n} \left\{ [\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 C_{ni}^3(\beta_n) + \frac{\partial^2 \log \Psi(\beta_n - \delta_i)}{\partial \beta^2{}_n} V_{ni}^2(\beta_n) \right\}}{2[\sum_{i=1}^{I}[\frac{\partial \log \Psi(\beta_n - \delta_i)}{\partial \beta_n}]^2 \cdot V_{ni}^2(\beta_n)]^2}$$

$$= -\frac{\sum_{i=1}^{I} \tanh(\beta_n - \delta_i) \left\{ [\tanh(\beta_n - \delta_i)]^2 C_{ni}^3(\beta_n) + [\cosh(\beta_n - \delta_i)]^{-2} V_{ni}^2(\beta_n) \right\}}{2[\sum_{i=1}^{I}[\tanh(\beta_n - \delta_i)]^2 \cdot V_{ni}^2(\beta_n)]^2}$$

$$= -\frac{\sum_{i=1}^{I}[\tanh(\beta_n - \delta_i)]^3 C_{ni}^3(\beta_n) + \sum_{i=1}^{I} \frac{\tanh(\beta_n - \delta_i)}{[\cosh(\beta_n - \delta_i)]^2} V_{ni}^2(\beta_n)}{2[\sum_{i=1}^{I}[\tanh(\beta_n - \delta_i)]^2 \cdot V_{ni}^2(\beta_n)]^2}$$

(D18)

Appendix E. Simulation studies.

To show the efficiency and robustness of the two JML estimation procedures described above, results of two simulation studies are reported in this section - one study focuses on the comparison between procedures A and B, and the other examines the effect of the number of items on the recovery of item and person parameters. In both studies, the operational function used in estimation is the same as used in generating the data. The convergence criterion for all iterations is set as 0.001.

Study 1. Comparison of procedures A and B.

In each data set of this study, responses to ten items were generated. The location values of the items were evenly spaced on the interval [-2.0,2.0]. The item units were randomly generated in the interval [0.6, 1.0]. Five hundred person locations normally distributed with mean of 0.0 and variance of 2.0. Each simulation had 10 replications. All the items involved were of four categories, e.g., *0 –strongly disagree; 1 – disagree; 2 – agree and 3 – strongly agree*. In the replications, data were simulated with the simple square logistic and hyperbolic cosine operational functions respectively.

Recovery of item locations. Table 1 shows the average of the estimated item locations over the 10 replications. In all estimations, the initial signs of the item locations are obtained using the *Sign Analysis* procedure (Luo, 1999). The estimates obtained with Procedure A are compared with those obtained with Procedure B. The root mean squared error (RMSE), a commonly used measure of estimation accuracy, is calculated according to

$$MRSE_{\delta} = \sqrt{\sum_{i=1}^{I}(\delta_i - \delta_i)^2 / I} \qquad \text{(E1)}$$

or

$$MRSE_{\zeta} = \sqrt{\sum_{i=1}^{I}(\zeta_i - \zeta_i)^2 / I} \qquad \text{(E2)}$$

Another measure of the accuracy of the estimation is the correlation coefficient between the generated person locations and their corresponding estimates. Table E1 also reports the average of these correlation coefficients over the 10 replications. Figure E1 plots the recovery of item locations in Table 1.

Table E1. Recovery of item locations

	Ψ	HCM				SSLM			
Item	Generating	A	StdErr	B	StdErr	A	StdErr	B	StdErr
1	-2.000	-2.052	0.034	-2.149	0.032	-1.854	0.028	-1.989	0.028
2	-1.556	-1.632	0.034	-1.589	0.033	-1.457	0.028	-1.569	0.028
3	-1.111	-1.173	0.035	-1.134	0.034	-1.084	0.030	-1.127	0.029
4	-0.667	-0.766	0.034	-0.695	0.034	-0.675	0.030	-0.705	0.030
5	-0.222	-0.244	0.035	-0.275	0.035	-0.233	0.031	-0.241	0.031
6	0.222	0.222	0.035	0.249	0.036	0.201	0.031	0.216	0.031
7	0.667	0.690	0.035	0.692	0.036	0.636	0.031	0.676	0.031
8	1.111	1.198	0.034	1.153	0.034	1.090	0.030	1.151	0.030
9	1.556	1.720	0.034	1.645	0.033	1.526	0.031	1.609	0.031
10	2.000	2.039	0.034	2.103	0.032	1.849	0.028	1.979	0.028
Person correlation		0.939		0.940		0.989		0.989	
RMSE		0.134		0.116		0.089		0.057	

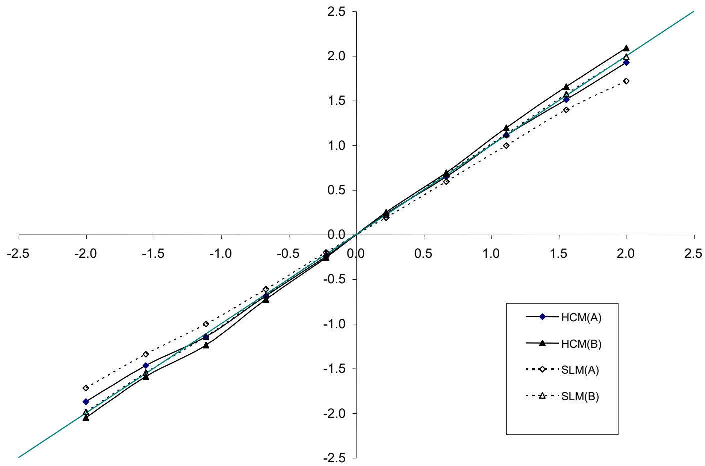


Fig. E1. Recovery of item locations in Study 1

It is seen in the results above that Procedure A produces slightly under-estimated item locations and units with the SSLM while Procedure B produces slightly over-estimated item locations and units with the HCM. The RMSE values with Procedure B are smaller than those with Procedure A for both operational functions. However, the difference in item estimates with different procedures is not significant in comparison to the corresponding standard errors, except for item 1 (the most negative item) and item 10 (the most positive item). Therefore, both sets of the estimates are acceptable for practical purposes.

Study 2. Effects of the number of items and sample size on item parameter estimates

Effect of the number of items. Table E3 lists the RMSE for the estimation of item parameters with Procedures A and B when the numbers of items involved are 5, 10, 15, and 20 respectively, while the other specifications remain the same as those in Study 1. It suggests that the item parameters can be well recovered with the number of items greater than 10.

Table E3. Values of the RMSE

	Locations				Units			
Ψ	HCM		SSLM		HCM		SSLM	
I	A	B	A	B	A	B	A	B
5	0.238	0.205	0.060	0.213	0.069	0.064	0.035	0.065
10	0.134	0.116	0.089	0.057	0.037	0.034	0.030	0.024
15	0.098	0.122	0.087	0.052	0.026	0.028	0.024	0.019
20	0.095	0.093	0.088	0.063	0.022	0.022	0.021	0.018

Effect of the sample size. In the simulation study reported above, the sample size was fixed at 500. To observe the effect of sample size on the recovery of item locations, more simulation studies were conducted, in which the sample sizes were specified in turn in increasing increments of 100 to 1000 while the other specifications of the simulations were as in Study 1. Then the estimation Procedures A and B were used respectively in recovering the item parameters. Figures E3 and E4 plot the RMSE of item locations for different sample sizes when the numbers of items involved were 5, 10, 15 and 20, and for the HCM and the SSLM respectively.

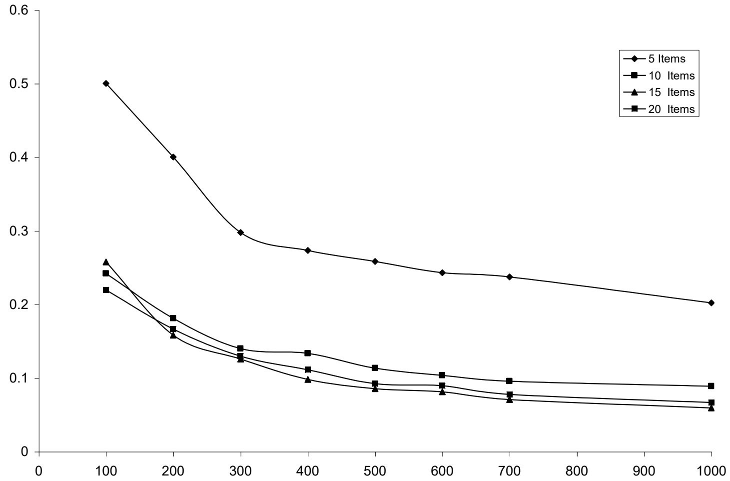


Fig. E3. MRSE of item location estimates for different sample sizes (operational function: HCM)

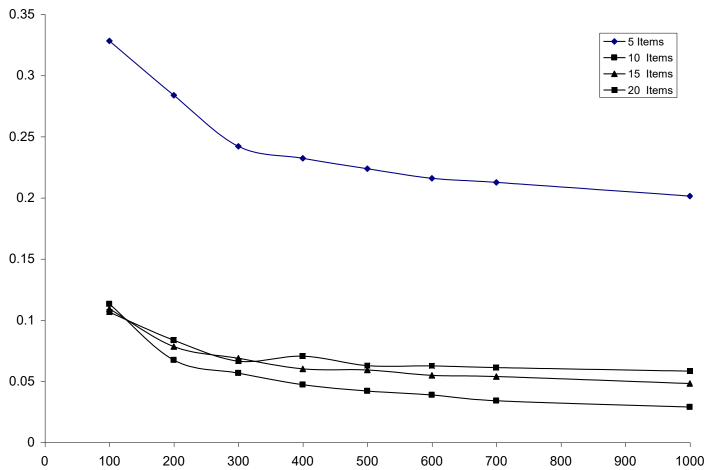


Fig. E4. MRSE of item location estimates for different sample sizes

(operational function: SLM

The Figures E3 and E4 also show that when the number of four-category items is greater than 10 and the sample size is greater than 500, the increase of sample size or the number of items does not affect the RMSE noticeably. Roberts, et. al. (2002) reported a similar result with the GGUM under the design of 15 six-category items and 750 persons.

Appendix F. A real example.

The example in this section is for illustrative purpose and was previously analyzed in [29]. It measures the attitude towards capital punishment which was thoroughly studied in literature [30] [31]. The version of the questionnaire used in [29] includes 10 statements as shown in the order from against to for capital punishment.

Item No.	Statement
9	*Capital punishment is just and necessary.*
7	*Capital punishment is justified because it stops serious crime.*
2	*Until we find a way to prevent serious crime, we need capital punishment.*
4	*Capital punishment gives criminals exactly what they deserve.*
8	*Capital punishment is necessary but I wish it were not.*
3	*I do not believe in capital punishment but it may be justified.*
5	*Capital punishment does not stop serious crime.*
1	*The state cannot teach that human life is sacred by destroying it.*
6	*Capital punishment is one of the most hideous practices in our society.*
10	*Capital punishment is never really justified.*

Three hundred and eighty students, who were from universities in Australia, Japan and Singapore, responded to the questionnaire. [29] Andrich & Luo (2003) analyzed the data using the HCM and the estimation procedure A. The analyses below compare the application of the HCM and the SSLM and estimation Procedures A and B. The convergence criterion for all iterations was again set as 0.001. Tables F1 and F2 list the estimated item parameters in these four analyses. It is seen that while the results of Procedure A is similar to those of Procedure B with the same operational function, the estimates of item unit estimated with the HCM are systematically greater than those estimated with the SSLM. Consequently, the estimated item locations with the HCM cover a larger range of the continuum than that covered by the estimated item locations with the SSLM. This arises from the relative properties of the models – the operational function of the SSLM has a much sharper shape than that of the HCM has.

Table F1. Estimates of item locations from real data

Ψ	HCM				SSLM			
Item	A	StdErr	B	StdErr	A	StdErr	B	StdErr
9	-2.876	0.037	-2.728	0.036	-1.621	0.027	-1.727	0.026
7	-3.104	0.037	-2.904	0.037	-1.610	0.027	-1.735	0.026
2	-2.794	0.037	-2.555	0.037	-1.398	0.028	-1.509	0.028
4	-2.636	0.036	-2.337	0.036	-1.251	0.029	-1.361	0.029
8	-1.580	0.035	-2.476	0.036	0.150	0.052	0.154	0.051
3	-1.873	0.035	-3.183	0.038	0.418	0.044	0.430	0.043
5	3.574	0.039	3.855	0.037	1.111	0.030	1.203	0.030
1	3.405	0.039	3.671	0.038	1.196	0.030	1.285	0.030
6	3.726	0.038	4.202	0.037	1.266	0.028	1.393	0.027
10	4.159	0.039	4.456	0.038	1.739	0.026	1.866	0.025

Table F2. Estimates of item units from real data

Ψ	HCM				SSLM			
Item	A	StdErr	B	StdErr	A	StdErr	B	StdErr
9	1.591	0.018	1.709	0.018	0.870	0.011	0.920	0.011
7	1.886	0.020	1.776	0.018	0.583	0.008	0.635	0.008
2	1.166	0.017	1.929	0.019	0.371	0.006	0.387	0.006
4	1.931	0.021	1.799	0.019	0.650	0.009	0.701	0.009
8	1.010	0.017	1.560	0.018	0.320	0.006	0.333	0.006
3	1.520	0.018	1.565	0.018	0.727	0.010	0.777	0.010
5	1.646	0.019	1.719	0.019	0.811	0.010	0.863	0.011
1	1.764	0.018	1.760	0.017	0.588	0.007	0.653	0.007
6	1.641	0.018	1.737	0.018	0.843	0.010	0.900	0.011
10	1.932	0.018	1.850	0.017	0.759	0.008	0.823	0.009

With the attempt to answer as to which model fits the data better, the following statistical indicators are presented.

1. *The log-likelihood value.* This is calculated using Equation (11) with the estimated parameter values

2. *Sum of square error.* For person n's response x_{ni} on item i, its expected value $E(x_{ni})$ can be calculated with the estimates for the person and item parameters. The corresponding square error is $[E(x_{ni}) - x_{ni}]^2$. Therefore, the sum of square error can be calculated according to

$$sse = \sum_{n=1}^{N} \sum_{i=1}^{I} [E(x_{ni}) - x_{ni}]^2 \tag{F1}$$

3. *Goodness-of-fit Statistic.* The conventional goodness-of-fit statistic involving class intervals is also used. Persons are divided into G class intervals. With each interval g, $g = 1,...,G$; the mean of person locations within the class interval is used to calculate the expected value E_{ig} for each item i. Then for each item i, the difference between this expected value and the mean score of the persons s_{ig} in the class interval are standardized by the standard deviation of this expected value $\sqrt{V(E_{ig})}$, that is

$$z_{ig} = \frac{s_{ig} - E_{ig}}{\sqrt{V(E_{ig})}} \tag{F2}$$

The overall goodness-of-fit statistic is calculated by

$$C^2 = \sum_{i=1}^{I} \sum_{g=1}^{G} z_{ig} = \sum_{i=1}^{I} \sum_{g=1}^{G} \frac{(s_{ig} - E_{ig})^2}{V(E_{ig})} \tag{F3}$$

It is noted that though the goodness-of-fit statistic of (F3) has a similar structure as that proposed by [32], whether the distribution of the statistic within the model of (1) is approximately χ^2 is to be explored. Therefore, it is considered only as a descriptive statistic for the goodness-of-fit. The values of this statistic for the real data of this section with different operational functions and estimation procedures are listed in (F3) where in the calculation of the goodness-of-fit statistic, persons were divided into 10 class intervals.

It is noted that for this particular data set, the results with the SSLM are quite robust with the two estimation procedures. It is interesting to note that though the three sets of estimates of item parameters with the SSLM in Tables F1, F2 and F3 are very close to each other, the goodness-of-fit statistics for them are different. It seems that the goodness-of-fit statistics are more sensitive than the other indicators. Although the estimates with procedure B for the SSLM have a slightly higher log-likelihood value, the estimates with procedure A for the SSLM have a much smaller overall goodness-of-fit statistic. In this sense, the results with SSLM using Procedure A seem marginally better. However, the choice of models is not conclusive.

Table F3. Statistics for the estimations from real data

Operational function	HCM		SSLM	
Procedure	A	B	A	B
Log Likelihood Value	-3784.50	-3753.95	-3672.68	-3660.54
Sum of Square Error	1684.65	1683.11	1612.87	1617.95
Goodness-of-fit statistic	659.262	486.148	279.00	322.17

The following summarizes the other results with the SSLM using Procedure A. The results with other combinations of the operational functions and procedures are similar. Figure F1 shows the distribution of the person locations with item locations estimated with the SSLM and Procedure A. The scale information function is also plotted. It is seen that the interval on which the scale information has high values captures the majority of the population. That is, the questionnaire is targeting this particular population excellently. Figure F2 plots the expected values of items 8, 9, 10 with the observed means of the person responses on these items (persons are divided into 10 class intervals according to their estimated locations). It can be seen that while the observed means of person responses on items 9 (the most negative item) and 10 (the most positive item) are monotonically decreasing and increasing respectively, those on item 8, which is in the middle of the continuum, form a single-peak. They are all as predicted by the corresponding expected values.

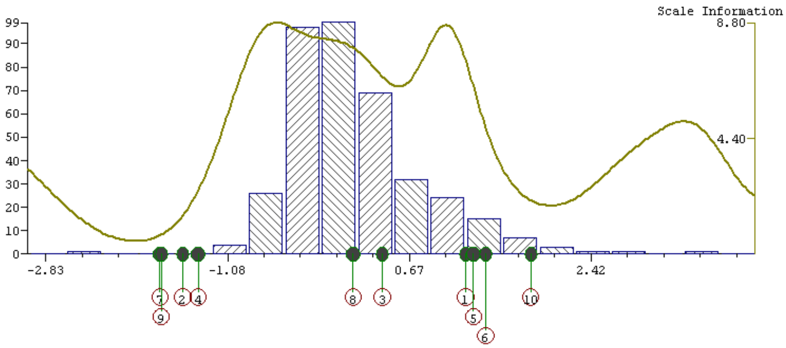
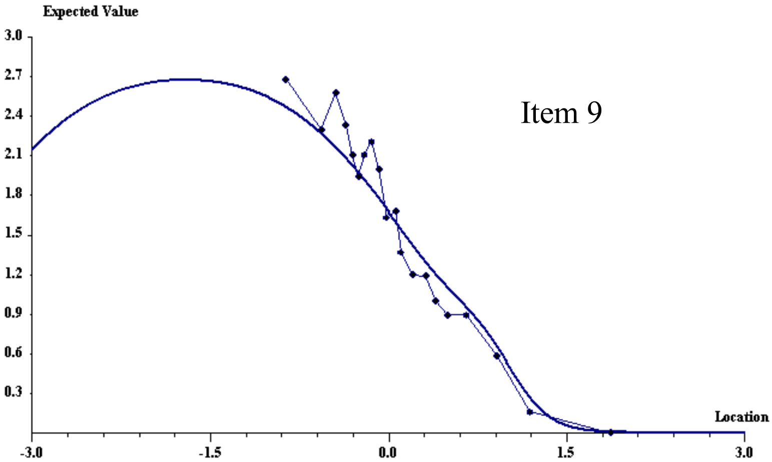
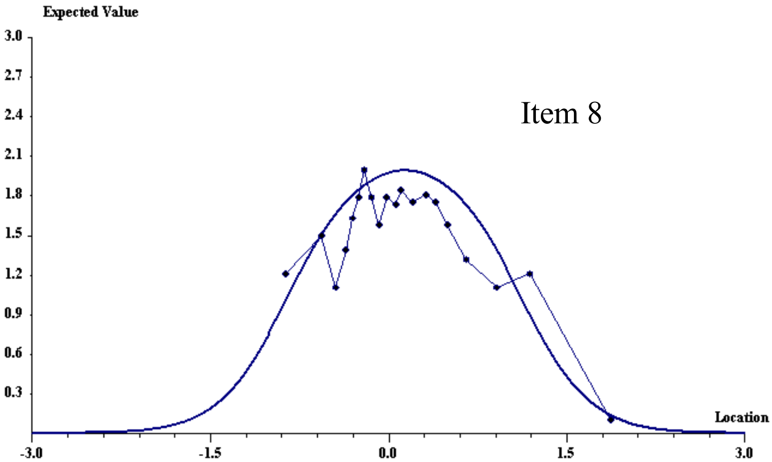


Fig. F1. The frequency of estimated person locations with item locations (labelled with item number in a circle) and the scale information function.


Expected Value
3.0
2.7
2.4
2.1
1.8
1.5
1.2
0.9
0.6
0.3
Item 9
Location
-3.0
-1.5
0.0
1.5
3.0


Expected Value
3.0
2.7
2.4
2.1
1.8
1.5
1.2
0.9
0.6
0.3
Item 8
Location
-3.0
-1.5
0.0
1.5
3.0

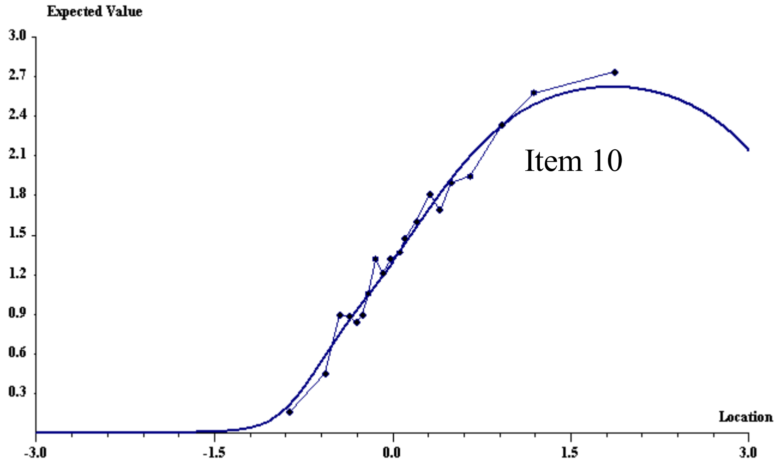


Fig. F1. The expected values of items 8, 9, 10 with the observed means of the person responses.

References

1. Thurstone, L. L. Attitude can be measured. *The American Journal of Sociology,* ***26,*** 249-269(1928).
2. Likert, R. A technique for the measurement of attitude. *Archives of Psychology,* ***140***, 5-53(1932).
3. Andrich, D. Hyperbolic Cosine Latent Trait Models for Unfolding Direct-Responses and Pairwise Preferences. *Applied Psychological Measurement,* ***20,*** 269-290(1996).
4. Roberts, J. S., Laughlin, J. E. and Wedell, D. H. Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement,* ***59,*** 211-233 (1999).
5. Van Shuur, W. H. *Structure of political beliefs.* Amsterdam: CT press(1984).
6. Hoijtink, H. The measurement of latent trait by proximity items. *Applied Psychological Measurement,* ***15,*** 153-159(1991).
7. Roberts, J. S. and Laughlin, J. E. A unidimensional item response model for unfolding responses from graded disagree-agree response scale. *Applied Psychological Measurement,* ***20,*** 231-255(1996).
8. Rost. J & Luo, G. An application of a Rasch based model to a questionnaire on adolescent centrism. Chapter 26. In J. Rost and R. Langeheine (Eds.), *Applications of Latent Trait and Latent Class Models in the Social Sciences.* Waxmann Verlag GMBH: Münster and New York, pp. 278-286(1997).

9. Roberts, J. S., Donoghue, J. R. and Laughlin, J. E. Ageneral item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement, **24,*** 3-32 (2000).
10. Roberts, J. S., Donoghue, J. R. and Laughlin, J. E. Characteristics of MML/EAP parameter estimates in the generalized graded unfolding model. *Applied Psychological Measurement, **26,*** 192-207 (2002).
11. Luo, G. A class of probabilistic unfolding models for polytomous responses. *Journal of Mathematical Psychology*, ***45,*** 224-248 (2001).
12. Andrich, D. A rating formulation for ordered response categories. *Psychometrika, **43,*** 561-573 (1978).
13. Luo, G. A general formulation for unidimensional unfolding and pairwise preference models: Making explicit the latitude of acceptance. *Journal of Mathematical Psychology*, ***42,*** 400-417 (1998).
14. Andrich, D. A extention of the Rasch model for rating providing both location and dispersion parameters. *Psychometrika, **47,*** 105-113(1982).
15. Luo, G., Andrich, D. and Styles, I. The JML estimation of the generalised unfolding model incorporating the latitude of acceptance parameter. *Australian Journal of Psychology, **50***, 187-198.(1998).
16. Luo, G. A joint maximum likelihood estimation procedure for the Hyperbolic Cosine Model for single stimulus responses. *Applied Psychological Measurement, **24,*** 33-49(1999).
17. Warm, T. A. Weighted likelihood estimation of ability in item response theory. *Psychometrika, **54,*** 427-450(1989).
18. Samejima, F. An approximation for the bias function of the maximum likelihood estimates of a latent variable for the general case where the item responses are discrete. *Psychometrika, **58***, 119-138 (1993).
19. Wang, S. & Wang, T. Precision of Warm's weighted likelihood estimates for a polytomous model in computerized adaptive testing. *Applied Psychological Measurement, **12,*** 307-314(2001).
20. Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical Theories of Mental Test Scores.* MA: Addison-Wesley (1968).
21. Wright, B. D. & Masters, G. N. Rating Scale Analysis. Chicago: MESA press(1982).
22. Masters, G. N. & Wright, B. D The Partial Credit Model. In van der Linden, W. J. & Hambleton, R. K. (edits). Handbook of modern Item Response Theory. New York: Springer(1996).
23. Neyman, J. & Scott, E. L. Consistent estimates based on partially consistent observation. *Econometrika, **16,*** 1-22(1948).
24. Dempster, A. P., Laird N. M. & Rubin D. B. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, B,* **39**,1-38(1977).
25. Bock, R. D. & Aitkin, M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, **46,*** 443-459(1981).
26. Linacre, J. M. Rasch model estimation: Further topics. *Journal of Applied Measurement, 5,* 95-110(2004).
27. Embretson, S. E. & Reise, S. P. Item Response Theory for Psychologists. New Jersey: Lawrence Erlbaum (2000).
28. Coombs, C. H. *A theory of data.* New York: Wiley (1964).

29. Andrich, D. & Luo, G. Measuring attitudes by unfolding a Likert – style questionnaire. In Keeves, J. (ed.), *The handbook on educational research in the Asia-Pacific region.* Kluwer Academic Publishers (2003).
30. Andrich, D. The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement, **12**,* 33-51(1988).
31. Andrich, D. & Luo, G. A hyperbolic cosine latent trait model for unfolding dichotomous single-stimulus responses. *Applied Psychological Measurement, **17**,* 253-276 (1993).
32. Wright, B. D. & Panchapakesan, N. (1969). A procedure for sample free analysis. *Educational and Psychological Measurement, **29**,* 23-48.