JOURNAL OF APPLIED MEASUREMENT, *6*(2), 128–146

Estimating Parameters in the Rasch Model in the Presence of Null Categories

Guanzhong Luo
Hong Kong Examinations and Assessment Authority

David Andrich
Murdoch University, Australia

A category with a frequency of zero is called a *null* category. When null categories are present in polytomous responses, then in the Rasch model for such responses, the thresholds that define the categories are inestimable with the commonly used joint maximum likelihood, marginal maximum likelihood, or standard conditional maximum likelihood estimation algorithms. The reason for this situation is that in principle, these estimation algorithms involve frequencies of each category. Andrich and Luo (2003) describe an alogorithm in which the thresholds are reparameterized into their principal components and in which the estimate of any threshold is based on a function of the frequencies of all categories of the item rather than the frequency of a particular category. This algorithm works in the presence of null categories. However, in situations where the null categories are at the extremes of a set of categories, the estimates themselves can become too extreme. This paper describes a procedure in which the solution algorithm described by Andrich and Luo is further adapted in the presence of null categories by using their expected frequencies. The procedure is demonstrated with simulated and real data.

Requests for reprints should be sent to Guanzhong Luo, Research Division, Hong Kong Examinations and Assessment Authority, 14/F Southorn Centre, 130 Hennessy Road, Wan Chai, Hong Kong, e-mail: gluo@hkeaa.edu.hk.

Introduction

This paper is concerned with estimating all thresholds of an item in the unidimensional Rasch model (RM) for ordered polytomous responses in the presence of categories which have zero frequencies. Such categories have been referred to as *null* (Wilson and Masters, 1993). In large scale testing situations where a wide range of the continuum is used, it is likely that some categories of polytomous items will be null with some populations which are not well targeted to the locations of the categories. Such categories are usually extreme and this paper is concerned with such cases. In principle, it is possible for middle categories to be null. Andrich (2004) shows such an example, but this paper is concerned only with null categories which are extreme.

In the Rasch model, any legitimate response category of an item has a positive probability, even though it may be small and even close to zero. A small theoretical probability may result in an observed null category depending on the size of the sample. The existence of null categories in real data reflects the possibilities that the data do not fit the Rasch model or that even though the data essentially fit the Rasch model, the probability of a category is very small because the distribution of person locations is not well targeted to the location of the categories. This paper focuses on the latter situation, in which the null categories are legitimate categories but with zero observed frequency essentially because persons are not located near the extreme categories.

When a category is null, the estimation of the thresholds defining this category in the RM is problematic with standard algorithms such as the conditional maximum likelihood (CML), (Andersen, 1971), the joint maximum likelihood (JML) (Wright and Masters, 1982), the marginal maximum likelihood (Dempster, Laird and Rubin, 1977), and the more recently described algorithm involving an eigenvector decomposition of a constructed pair comparison matrix (Garner and Engelhard, 2002). It is problematic with these algorithms because they directly involve frequencies of each category.

Various strategies have been suggested to overcome this problem. For example, Wright, Masters and Ludlow (1982) and Wright, Condon and Schultz (1988) collapsed the categories by summing the frequencies of adjacent categories. However, there are theoretical reasons against collapsing categories in this way in the Rasch model. First, collapsing categories of an item that fits the Rasch model in such a way corresponds to adding probabilities of adjacent categories, and this new model is no longer a Rasch model (Jansen and Roskam, 1986; Wilson and Masters; 1993; Andrich, 1995). Second, it alters the planned relationship between the substantive framework and the scoring scheme for the item.

Wilson and Masters (1993) re-formalized the RM to avoid collapsing categories. In maintaining the integrity of the RM, their re-formalization avoided the estimation of the two thresholds that define a null category while admitting their existence. Although these thresholds remain inestimable, the point where the characteristic curves for the two adjacent categories of the null category crosses is estimated. This point is considered to be the mid-point or mean of the two un-estimable thresholds. The advantage of this procedure is that even in the presence of some null categories, the thresholds not related to these null categories are recovered without altering the model itself. However, the thresholds related to null categories remain inestimable. In particular, when an extreme (lowest or highest) category of an item is null, which is often the case in real situations, the procedure of Wilson and Masters (1993) provides no information on the corresponding extreme threshold.

Andrich and Luo (2003) describe an estimation algorithm for the RM that uses a reparameterized form of the thresholds in terms of their principal components (Andrich, 1985; Pedler, 1987), where the term "principal components" is used in the sense of Guttman (1950) rather than that of the manipulation of a correlation matrix. With this algorithm, the estimate of any threshold is based on a function of the frequencies of all categories of the item rather than the frequency of particular categories. The estimation algorithm therefore reflects and takes advantage of the structural property of the model that the response in any category is a function of all thresholds, and not only the adjacent thresholds that define a category. Because it involves functions of frequencies of all categories in the estimation of all thresholds, the algorithm raises the possibility of providing estimates for all thresholds in the presence of some null categories. Even though this algorithm reaches a solution in the presence of null categories without further modification, the estimate of extreme thresholds can themselves become somewhat extreme. This paper summarises the algorithm and then describes the modifications suggested for the case of extreme null categories. The modification of the solution equations is similar to the *single imputation* technique used for the treatment of missing data (Little and Rubin, 1989). The paper shows the analysis of a real data set and corresponding simulated studies to show the degree to which the modification is successful in regressing the estimates of the extreme thresholds.

The rest of this paper is structured as follows. Section 2 summarises the reparameterisation of the Rasch model and the solution equations established in Andrich and Luo (2003); Section 3 introduces the adjustment to the equations in the presence of null categories with the imputed values; Section 4 presents an analysis of a real data set; and Section 5 presents a simulation study based on the results of this real data set. Section 6 presents simulation studies with a smaller number of categories than the real example and Section 7 provides a summary and discussion.

The Reparameterized Rasch Model and the Conditional Pairwise Estimation Algorithm

The estimation algorithm described in Andrich and Luo (2003), and modified in this paper for the presence of null categories, is a generalisation of the conditional pairwise algorithm for dichotomous responses (Choppin, 1985).

Although there are different equivalent expressions for the RM, following Andrich (1978), the one used in this paper takes the form

$$\Pr\{X_{ni} = x\} = \exp[x(\beta_n - \delta_i) + \kappa_{xi}] / \gamma_{ni}, \tag{1}$$

where

$$X_{ni} = x \in \{0,\ 1,\ 2, \ldots,\ m_i\}$$

is the random variable in which the successive integers are scoring functions of the successive categories; β_n and δ_i are the locations of the person and item respectively; $\{\tau_{xi}, x = 1,\ldots,m_i\}$ are *centralized* thresholds that separate the successive categories $x - 1$ and x for item i,

$$\sum_{x=1}^{m_i} \tau_{xi} = 0, \tag{2}$$

$$\kappa_{xi} = -\sum_{k=1}^{x} \tau_{ki},\quad x = 1, \ldots, m_i;\ \kappa_{oi} = \kappa_{mi} \equiv 0 \tag{3}$$

are category coefficients; and

$$\gamma_{ni} = 1 + \sum_{x'=1}^{m_i} \exp[x' (\beta_n - \delta_i) + \kappa_{xi}^{'}] \tag{4}$$

is a normalising factor.

From Eq. (3), it is evident that

$$\tau_{(x+1)i} = \kappa_{xi} - \kappa_{(x+1)i}\,. \tag{5}$$

Eq. (1) can be readily written as

$$\Pr\{X_{ni} = x\} = \exp[x\beta_n - x\delta_i + \kappa_{xi}] / \gamma_{ni} \tag{6}$$

from which Andrich and Luo (2003) reparameterize it into the form

$$P_{xi} = \Pr\{X_{ni} = x\} = \frac{1}{\gamma_{ni}} \exp[x\beta_n + \sum_{l=1}^{4} f_{li}(x)\omega_{li}], \tag{7}$$

where the coefficients $\{f_{li}, l = 1,\ldots,4\}$ of the fixed four principal components $\{\omega_i, i = 1,2,3,4\}$ are defined by

$$\begin{aligned} f_{1i}(x) &= -x, \\ f_{2i}(x) &= x(m_i - x), \\ f_{3i}(x) &= x(m_i - x)(2x - m_i), \\ f_{4i}(x) &= x(m_i - x)(5x^2 - 5xm_i + m_i^2 + 1). \end{aligned} \tag{8}$$

It is evident from equating terms in the numerator of Eqs. (5) and (6), that

$$\kappa_{xi} = \sum_{l=2}^{4} f_{li}(\omega_{li}) . \tag{9}$$

Before proceeding summarising the estimation algorithm based on this reparameterization, Eq. (1) can also be written

$$\Pr\{X_{ni} = x\} = \exp[x\beta_n - \sum_{k=0}^{x} \delta_{ki}] / \gamma_{ni} , \tag{10}$$

where

$$\delta_{ki} = \delta_i + \tau_{ki} \tag{11}$$

are referred to as *uncentralized* thresholds. Similar to the τ_{ki}; $k = 1,2...,m_i$, that are referenced to each item location δ_i and cannot be compared directly, the δ_{ki}, $k = 1,2...,m_i$ are referenced to a single origin and can be compared directly on a single continuum (Wright and Masters, 1982).

Proceeding with a summary of the estimation algorithm, for any item pair i, j, let their respective maximum scores be m_i, m_j. The total score r_{nij} of a person n on the item pair is defined by

$$r_{nij} = x_{ni} + x_{nj}. \tag{12}$$

Because items are always considered in pairs, for simplicity, the subscript of r_{nij} is dropped and the total score for a pair of items is denoted as r. With given r, the possible values for x_{ni} are between a minimum value $L_{ij}(r)$ and a maximum value $U_{ij}(r)$ given by

$$\begin{aligned} L_{ij}(r) &= \max(0, r - m_j), \\ U_{ij}(r) &= \min(m_i, r). \end{aligned} \tag{13}$$

Again for simplicity, these limits are denoted as L and U respectively. With given r, the conditional probability of $x_{ni} = x$, $x \in [L,U]$, is given by

$$p_{ij}(x \mid r) = P\{x_{ni} = x \mid x_{ni} + x_{nj} = r\} = \frac{P\{x_{ni} = x\} P\{x_{nj} = r - x\}}{\sum_{k=L}^{U} P\{x_{ni} = k\} P\{x_{nj} = r - k\}}$$

$$= \frac{\exp[\sum_{l=1}^{4} f_{li}(x)\omega_{li}] + \sum_{l=1}^{4} f_{lj}(r-x)\omega_{lj}]}{\sum_{k=L}^{U} \exp[\sum_{l=1}^{4} f_{li}(k)\omega_{li}] + \sum_{l=1}^{4} f_{lj}(r-k)\omega_{lj}]} . \tag{14}$$

Note that from the symmetry of (14),

$$p_{ji}(r - x \mid r) = p_{ij}(x \mid r) . \tag{15}$$

Equation (14) does not involve the person parameter β_n, and is therefore invariant across persons, given the same total score.

Consider the responses of all persons to a pair of items i, j. Let $n_{ij}(x \mid r)$ be the number of persons whose response $x_{ni} = x$ and $x_{ni} + x_{nj} = r$, and let

$$N_{ij}(r) = \sum_{x=L}^{U} n_{ij}(x \mid r) \tag{16}$$

be the number of persons who have a total score of r on items i and j. In anticipation of introducing the imputed value for a null category, note that when $N_{ij}(r)$ is given, the expectation of $n_{ij}(x \mid r)$ can be written in terms of $N_{ij}(r)$ and $p_{ij}(x \mid r)$:

$$E[n_{ij}(x \mid r)] = N_{ij}(r) p_{ij}(x \mid r) . \tag{17}$$

Returning to the original algorithm, the solution equations for the first four principal components $\{\omega_{li}, l = 1,2,3,4\}$ defined in Andrich and Luo (2003) are:

$$T_{li} - \sum_{j \neq i} \sum_{r=1}^{m_i + m_j - 1} N_{ij}(r) \sum_{k=L}^{U} f_{li}(k) p_{ij}(k \mid r) = 0; \tag{18}$$

$l = 1,...,4$; $i = 1,...,I$,

where

$$T_{li} = \sum_{j \neq i} \sum_{r=1}^{m_i + m_j - 1} \sum_{x=L}^{U} n_{ij}(x \mid r) f_{li}(x) \tag{19}$$

is the sufficient statistic of ω_{li}.

By solving Eq. (19), principal components ω_{li}, $i = 1,2,3,4$ in Eq. (7) are estimated. Then in turn, the centralized thresholds of Eq. (1) can be recovered from Eqs. (6) and (9), or more directly from

$$\tau_{xi} = \sum_{l=1}^{4} f_{li}(x-1)\omega_{li} - \sum_{l=1}^{4} f_{li}(x)\omega_{li} . \tag{20}$$

Eq. (11) is then used to recover the uncentralized thresholds $\{\delta_{ki}\}$ from the centralized thresholds of $\{\tau_{xi}\}$.

The procedure summarized above has four desirable features: (i) the method eliminates the person parameters which are the source of inconsistency in the standard joint maximum likelihood estimation; (ii) the relevant statistics for the estimate of each threshold is a function of the frequencies of all categories and not just one category, thus potentially stabilising estimates in the presence of low frequency or null categories; (iii) the procedure accommodates readily missing data in the sense that some persons do not respond to all items; and (iv) it accommodates different numbers of categories in different items. All of these features are relevant in large scale testing where the range of the continuum can be large and where, therefore, there might be categories with low or even null frequencies at extremes.

The Adjustment for Null Categories with Expected Frequencies

Only one null category for an item

When a category x of item i is not observed in the data, $n_{ij}(x \mid r) = 0$ for any item $j \neq i$ and score r. The solution equation of Eq. (18) can be written in terms of $N_{ij}(r)$ as

$$\begin{aligned}
&\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\sum_{k=L}^{U}[n_{ij}(k\mid r)-N_{ij}(r)p_{ij}(k\mid r)]f_{li}(k) \\
&=\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\sum_{\substack{k=L\\k\neq x}}^{U}[n_{ij}(k\mid r)-N_{ij}(r)p_{ij}(k\mid r)]f_{li}(k) \qquad (21)\\
&\quad+\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}[n_{ij}(x\mid r)-N_{ij}(r)p_{ij}(x\mid r)]f_{li}(x)=0,
\end{aligned}$$

where $l = 1,...,4$; $i = 1,...,I$.

If the category x is logically observable, then the cause of x being a null category is due to sampling. In this case, the best approximation for $n_{ij}(x \mid r)$ is its expectation $E[n_{ij}(x\mid r)] \neq 0$.

It is noted that the situation here with null categories is different from that which is generally referred to as missing observed values on a variable, or *missing data* (Little and Rubin, 1989). That kind of missing data occurs in the RM if the persons were not given all questions or they refused to answer some of the questions. This situation is also accommodated by the original estimation algorithm (Andrich and Luo, 2003), but without computing any values for missing responses. In the situation with null categories, however, the frequency of the category is zero even if all persons responded to all items so that there is no missing data. The conventional imputation procedure for the missing data is to insert the mean of the variable into the missing cell. In contrast, the zero frequency of the null category, which is 0, is replaced with its expected value, which in the algorithm is obtained from the parameter estimates from frequencies of all other categories. Therefore, as can be seen in the following sections, the principle of the single imputation procedure is reflected in the procedure described to adjust for a null category.

Specifically, the zero frequency $n_{ij}(x \mid r)$ of the null category is replaced by its *imputed value* $E[n_{ij}(x \mid r)]$. Consequently, when $E[n_{ij}(x \mid r)]$ is used in place of $n_{ij}(x \mid r)$ in Eq. (21), $N_{ij}(r)$, which is the total number of persons with a total score of r, also needs to be adjusted to

$$\tilde{N}_{ij}(r) = N_{ij}(r) + E[n_{ij}(x\mid r)]. \qquad (22)$$

Substituting $N_{ij}(r)$ with $\tilde{N}_{ij}(r)$ in Eq. (17) gives the adjusted expectation of $n_{ij}(x \mid r)$ as

$$E[n_{ij}(x\mid r)] = \tilde{N}_{ij}(r)p_{ij}(x\mid r). \qquad (23)$$

Substituting Eq. (23) into Eq. (22) gives

$$\tilde{N}_{ij}(r) = N_{ij}(r) + \tilde{N}_{ij}(r)p_{ij}(x\mid r). \qquad (24)$$

Therefore, the adjusted number of persons with a total score of r is

$$\tilde{N}_{ij}(r) = \frac{N_{ij}(r)}{1-p_{ij}(x\mid r)}. \qquad (25)$$

It is noted that in Eq. (25),

when $N_{ij}(r) = 0$, $\tilde{N}_{ij}(r) = 0$.

In Eq. (21), for the null category x and any r, replace $N_{ij}(r)$ by $\tilde{N}_{ij}(r)$ and replace $n_{ij}(x \mid r)$ by

$E[n_{ij}(x \mid r)]$. The solution equations after these replacements are called the *adjusted* solution equations, which have the form

$$\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\sum_{\substack{k=L\\k\neq x}}^{U}[n_{ij}(k\,|\,r)-\tilde{N}_{ij}(r)p_{ij}(k\,|\,r)]f_{li}(k)$$
$$+\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\{E[n_{ij}(x\,|\,r)]-\tilde{N}_{ij}(r)p_{ij}(x\,|\,r)\}f_{li}(x)=0, \qquad (26)$$

where l = 1,2,3,4.

According to Eq. (23), the second term of the left hand side of Eq. (26) is zero:

$$\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\{E[n_{ij}(x\,|\,r)]-\tilde{N}_{ij}(r)p_{ij}(x\,|\,r)\}f_{li}(x)=0. \qquad (27)$$

The adjusted solution equations are then simplified to

$$\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\sum_{\substack{k=L\\k\neq x}}^{U}[n_{ij}(k\,|\,r)-\tilde{N}_{ij}(r)p_{ij}(k\,|\,r)]f_{li}(k)=0. \qquad (28)$$

Because $n_{ij}(x \mid r) = 0$ for the null category x of item i,

$$\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\sum_{\substack{k=L\\k\neq x}}^{U}n_{ij}(k\,|\,r)f_{li}(k)=\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\sum_{k=L}^{U}n_{ij}(k\,|\,r)f_{li}(k)=T_{li}. \qquad (29)$$

Therefore, the adjusted solution equations (l = 1,2,3,4) can be written as

$$T_{li}-\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\tilde{N}_{ij}(r)\sum_{\substack{k=L\\k\neq x}}^{U}f_{li}(k)p_{ij}(k\,|\,r)=0\,, \qquad (30)$$

where $\tilde{N}_{ij}(r)$ is calculated with Eq. (22). In particular, when $N_{ij}(r)$, then $\tilde{N}_{ij}(r)\equiv 0$, and the term with $\tilde{N}_{ij}(r)$ drops out of Eq.(30). Therefore, Eq. (30) is valid regardless of the value of $N_{ij}(r)$.

Two or more null categories for an item

Let ϕ be the collection of all null categories for item i, and let

$$\tilde{N}_{ij}(r)=N_{ij}(r)+\sum_{x\in\phi}E[n_{ij}(x\,|\,r)]=N_{ij}(r)+[\tilde{N}_{ij}(r)\sum_{x\in\phi}p_{ij}(x\,|\,r)]. \qquad (31)$$

Therefore,

$$\tilde{N}_{ij}(r)=\frac{N_{ij}(r)}{1-\sum_{x\in\phi}p_{ij}(x\,|\,r)}. \qquad (32)$$

The adjusted solution equations (l = 1,2,3,4) then become

$$T_{li}-\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\tilde{N}_{ij}(r)\sum_{\substack{k=L\\k\notin\phi}}^{U}f_{li}(k)p_{ij}(k\,|\,r)=0\,, \qquad (33)$$

where $\tilde{N}_{ij}(r)$ is calculated using Eq. (33).

Two or more items with null categories

Let ϕ_i be the collection of all null categories for item i and ϕ_j be the collection of all null categories for item j. Then

$$\tilde{N}_{ij}(r)=N_{ij}(r)+\sum_{x\in\phi_i}E[n_{ij}(x\,|\,r)]+\sum_{\substack{y\in\phi_j\\x\notin\phi_i}}E[n_{ij}(r-y\,|\,r)]$$
$$=N_{ij}(r)+\tilde{N}_{ij}(r)[\sum_{x\in\phi_i}p_{ij}(x\,|\,r)+\sum_{\substack{y\in\phi_j\\x\notin\phi_i}}p_{ij}(r-y\,|\,r)]. \qquad (34)$$

Therefore,

$$\tilde{N}_{ij}(r)=\frac{N_{ij}(r)}{1-[\sum_{x\in\phi_i}p_{ij}(x\,|\,r)+\sum_{\substack{y\in\phi_j\\x\notin\phi_i}}p_{ij}(r-y\,|\,r)]}. \qquad (35)$$

The adjusted solution equations (l = 1,2,3,4) are

$$T_{li}-\sum_{j\neq i}\sum_{r=1}^{m_i+m_j-1}\tilde{N}_{ij}(r)\sum_{\substack{k=L\\k\notin\phi_i\\r-k\notin\phi_j}}^{U}f_{li}(k)p_{ij}(k\,|\,r)=0, \qquad (36)$$

where $\tilde{N}_{ij}(r)$ is calculated using Eq. (35).

A Real Example

This Section presents a real example from a large-scale assessment exercise in which a wide range of the continuum is used and null categories appear naturally at extrenes as a result of poor targeting. In Western Australia, a high stakes test, termed the Western Australia Literacy and

Numeracy Assessment (WALNA), has been administered to all students in years 3 and 5 since 1999, and also to year 7 students since 2001. Student performance is compared across year levels using equating techniques based on the RM. Specifically, the test was designed to cover the instructional objectives of Year 3 through Year 10 so that the results of students from various years can be placed on the same scale. The writing test of the WALNA, which is the main interest of this Section, gives the same prompt (instruction) to students of different year levels, and their essays are marked using the same marking guide. The judges are instructed to mark the essays based on the following aspects:

(0) On balance judgement (OBJ)

(1) Spelling (SP)

(2) Vocabulary (V)

(3) Sentence control (SC)

(4) Punctuation (P)

(5) Form of Writing (F)

(6) Subject matter (SM)

(7) Text organization (TO)

(8) Purpose and audience (PA)

The performance of each student on each aspect and irrespective of year level is graded using the same system of ordered categories. A higher category means a better performance. The numbers of categories are 6 for SP, 7 for P and 8 for all other aspects. The marking results on a subset of students who took the WALNA writing test of 2000 are analyzed in this Section. This data set was used for equating the student performance over the period of 1999 to 2002. Each aspect of the marking guide was considered as an item. However, it is evident that Item 0 (OBJ), which effectually summarizes the performance on all the other aspects, potentially causes a violation of local independence. For this reason, the following analysis is carried out with this item removed from the data. The response frequencies of the remaining 8 items are shown in Table 1.

Table 1 shows null categories at the high end for all items except Items 1 and 4. This is a symptom of poor targeting of the test because students of Year 10 level are targeted in designing the marking key but are not included in the sample.

When RUMM2020 (Andrich, Sheridan and Luo; 2003) is used to analyse the data, which implements the algorithm summarised above, the principal components parameters of Eq. (2) are estimated with the model of four principal component parameters and is therefore of reduced rank for all items with more than five categories. The convergence criterion was set to 0.001. The data were analyzed twice in this study. First, the data were analyzed without the adjustment for null categories. This is possible because, as indicated above, even without the adjustment for null categories, the algorithm provides estimates for all thresholds. The estimation iterations converged after 354 loops. Tables 2 and 3 show the estimates of the principal components without adjustment for null categories and the corresponding thresholds respectively. The person distribution with a plot of thresholds is shown in

Table 1

Frequencies of categories of the real example.

		Category Frequency							
Item	Label	0	1	2	3	4	5	6	7
1	Sp	4	122	243	337	172	18		
2	V	84	91	183	322	108	101	7	0
3	SC	90	110	290	217	92	94	3	0
4	P	141	120	197	266	101	70	1	
5	F	87	126	222	258	83	118	2	0
6	SM	141	80	270	226	68	109	2	0
7	TO	150	52	194	305	105	86	4	0
8	PA	123	98	227	264	61	120	3	0

Figure 1. It is noted that the last thresholds of the items with null categories are quite large. As an extreme case, the last threshold of item 5 is 17.16. Figure 1 shows the distribution of the estimated person locations with the thresholds estimated without the adjustment for null categories.

The data were analyzed again using the adjustment algorithm for null categories with the same convergence criterion of 0.001. The iterations converged after 305 loops compared to 354 loops without the adjustment. The estimated values of the principal components and the corresponding thresholds are shown in Table 4 and Table 5 respectively. Figure 2 shows the distribution of estimated person locations with thresholds estimated with the adjustment for null categories.

A comparison of Tables 3 and 5 shows that although most threshold estimates in the two sets of are quite similar, for the last threshold of the items with null categories, the estimates with the adjustment are noticeably smaller than the corresponding estimates without the adjustment. To show the difference between these two sets of threshold estimates, Figure 3 shows a plot with the *x*-values being the thresholds estimated with the adjustment, and the *y*- values being the differences between the corresponding thresholds estimated without the adjustment and that with the adjustment. It is evident that the greatest differences appear at the positive end of the continuum where the thresholds related to null categories are located. Specially, the thresholds estimated with the adjustment are regressed towards the origin. Because of the constraint that the mean of all item locations is zero in both analyses, the values of other thresholds estimated with the adjustment are also slightly altered to compensate for the regression of the threshold estimates at the positive end of the continuum. The simulation study described in the next Section shows that this regression of the extreme thresholds, when the related extreme categories are null, is in the correct direction.

Table 2

Estimated principal components of the real example (without adjustment for null categories).

Item	Location	*SE*	Unit	*SE*	Skewness	*SE*	Kurtosis	*SE*
1	-4.721	0.082	3.163	0.042	0.051	0.008	0.009	0.003
2	0.065	0.071	2.240	0.022	0.008	0.004	-0.008	0.001
3	0.966	0.072	2.257	0.022	-0.005	0.004	-0.003	0.001
4	-0.174	0.071	2.368	0.025	0.078	0.005	-0.001	0.001
5	1.156	0.071	2.424	0.022	0.044	0.004	0.001	0.001
6	1.103	0.070	2.075	0.021	0.002	0.004	-0.005	0.001
7	0.675	0.071	2.057	0.021	0.004	0.004	-0.013	0.001
8	0.931	0.070	2.171	0.021	0.020	0.004	-0.003	0.001
Mean	0.000	0.072	2.344	0.025	0.025	0.005	-0.003	0.001
Std Dev	1.969	0.004	0.355	0.007	0.029	0.001	0.006	0.001

Table 3

Estimated uncentralized thresholds of the real example (without adjustment for null categories).

	Thresholds						
Item	1	2	3	4	5	6	7
1	-16.99	-10.90	-5.33	0.85	8.77		
2	-12.17	-9.88	-5.53	-0.12	5.39	10.01	12.76
3	-12.31	-8.48	-3.88	1.09	5.99	10.41	13.94
4	-10.39	-7.68	-3.85	0.99	6.70	13.18	
5	-12.19	-8.41	-4.35	0.10	5.08	10.72	17.16
6	-10.65	-7.83	-3.72	1.05	5.85	10.04	12.98
7	-10.04	-9.07	-5.02	0.59	6.24	10.42	11.61
8	-11.11	-8.14	-4.15	0.46	5.30	10.00	14.17

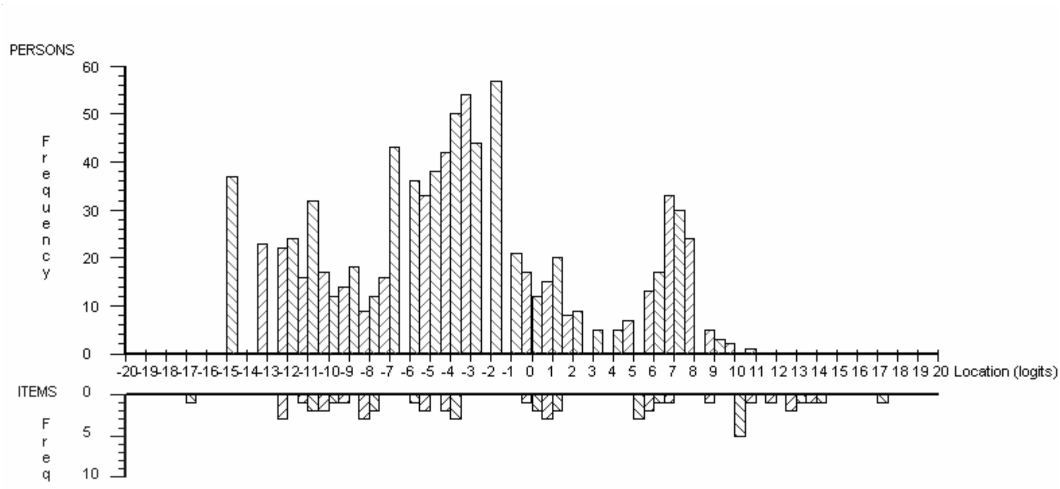

Figure 1. The distribution of estimated person locations and thresholds of the real example (without adjustment for null categories).

Table 4

Estimated principal components of the real example (with adjustment for null categories).

Item	Location	*SE*	Unit	*SE*	Skewness	*SE*	Kurtosis	*SE*
1	-4.609	0.082	3.132	0.042	0.049	0.008	0.009	0.003
2	0.036	0.071	2.187	0.021	0.002	0.004	-0.009	0.001
3	0.948	0.071	2.210	0.021	-0.010	0.004	-0.004	0.001
4	-0.127	0.070	2.334	0.025	0.074	0.005	-0.001	0.001
5	1.165	0.071	2.384	0.022	0.040	0.004	0.001	0.001
6	1.100	0.070	2.033	0.021	-0.002	0.004	-0.005	0.001
7	0.545	0.070	1.971	0.021	-0.007	0.004	-0.014	0.001
8	0.941	0.070	2.132	0.021	0.016	0.004	-0.003	0.001
Mean	0.000	0.072	2.298	0.024	0.020	0.005	-0.003	0.001
Std Dev	1.924	0.004	0.364	0.007	0.031	0.001	0.007	0.001

Table 5

Estimated uncentralized thresholds of the real example (with adjustment for null categories).

	Thresholds						
Item	1	2	3	4	5	6	7
1	-16.76	-10.76	-5.19	0.95	8.71		
2	-12.01	-9.74	-5.39	0.00	5.41	9.81	12.18
3	-12.16	-8.34	-3.74	1.19	5.99	10.23	13.46
4	-10.24	-7.54	-3.71	1.09	6.69	12.94	
5	-12.05	-8.26	-4.21	0.20	5.10	10.59	16.79
6	-10.51	-7.69	-3.58	1.15	5.86	9.89	12.58
7	-9.88	-8.96	-4.89	0.72	6.24	10.05	10.53
8	-10.97	-7.99	-4.02	0.56	5.32	9.87	13.81

Simulation Study 1: Reduced Rank Reflecting the Real Example

Two simulation studies are described in this paper to demonstrate the efficiency of the parameter estimation algorithm in the presence of null categories. This Section presents Study 1, which mimics the structure of the real example when the model with four principal components of Eqs. (5) and (6) and is of reduced rank.

In this study, the category numbers of the 8 items involved were the same as in the real example of the previous Section. The estimated uncentralized thresholds in Table 5 were used to simulate the responses. In order that the person distribution is similar to that in Figure 2, three groups of normally distributed persons were simulated, using the statistics specified in Table 6.

The over all distribution of the generated person locations is shown in Figure 4.

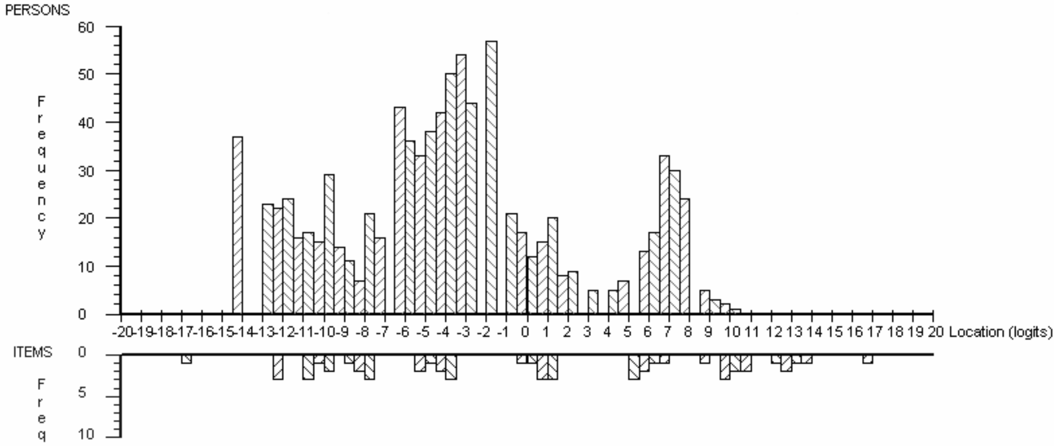

Figure 2. The distribution of estimated person locations and thresholds of the real example (with adjustment for null categories).

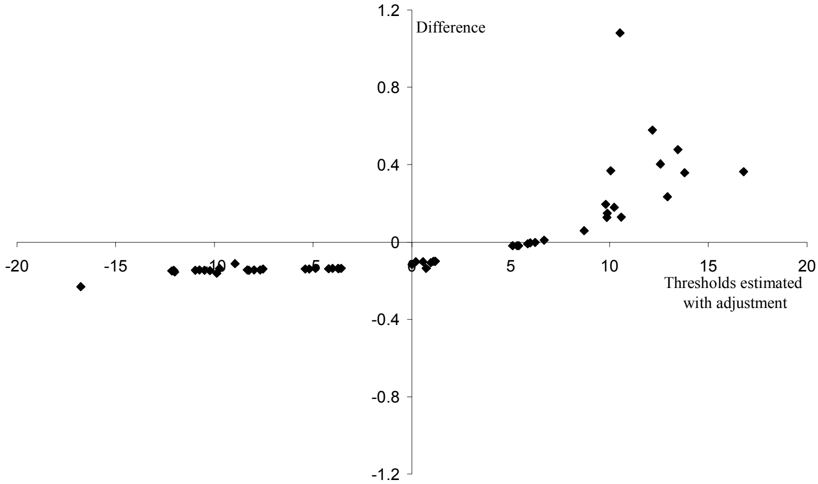

Figure 3. The thresholds estimated with the adjustment against the difference between the corresponding thresholds estimated without the adjustment and these values.

The RM of Eq.(1) was used to generate the responses. The frequency table of the generated data is shown in Table 7.

It can be seen that the data are similar to the real data in that the highest categories of all items except item 1 are null categories. To see the effect of the algorithm of the null category adjustment, the data were again analyzed with and without the adjustment for null categories. In both analyses, the convergence criterion was set at 0.001. The estimation iterations without the adjustment for null categories converged after 421 loops. The estimation iterations with the adjustment converged after 365 loops. The estimated values and standard errors of the principal components are

Table 6

Descriptive statistics of three groups of generated person locations in Study 1.

Group	Persons	Mean	St Dev	Min	Max
1	250	-11.50	3.00	-15.00	-9.00
2	500	-3.00	4.00	-9.00	3.00
3	144	6.00	1.50	2.00	9.50
Total	894	-3.85	6.23		

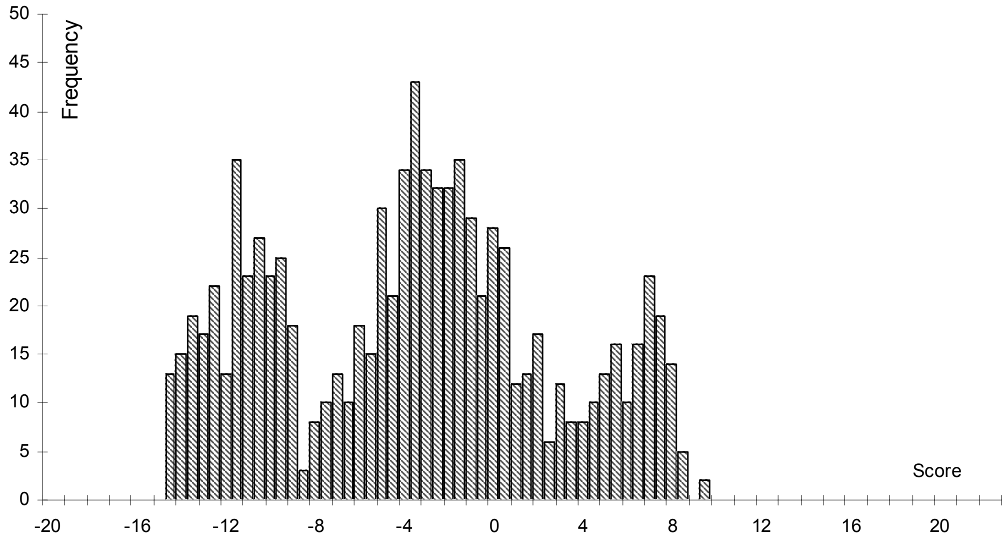

Figure 4. The distribution of generated person locations in Study 1.

Table 7

Frequencies of categories of simulated data in Study 1.

	Category							
Item	0	1	2	3	4	5	6	7
1	2	173	177	348	181	15		
2	101	109	144	300	155	81	6	0
3	110	141	187	269	99	87	3	0
4	180	93	180	242	147	54	0	
5	103	155	172	227	136	99	4	0
6	173	96	180	249	123	73	2	0
7	196	59	133	289	154	59	6	0
8	140	124	156	256	134	81	5	0

shown in Table 8. The recovered values of the uncentralized thresholds are shown in Table 9.

As a measure of estimation accuracy, the root mean square error (RMSE) was calculated for both sets of recovered thresholds against their generating values:

$$MRSE = \sqrt{\sum_{i=1}^{I}\sum_{k=1}^{m_i}(\hat{\tau}^*_{ik} - \tau^*_{ik})^2 / \sum_{i=1}^{I} m_i} \;, \qquad (38)$$

where $\hat{\tau}^*_{ik}$ is an estimate of τ^*_{ik}. Based on the values of { τ^*_{ik} } in Table 9 and their corresponding estimates in Table 12, the RMSEs for the two sets of thresholds estimates are 0.433 and 0.363, without and with the adjustment for null categories respectively.

Although the value of RMSE is only slightly smaller for the algorithm with the adjustment, the improvement is substantial on the last thresholds of the items that have the highest category as a null category. Figure 5 shows the distributions of estimated person and threshold locations of the two analyses. Figure 6 shows the difference between the corresponding estimated thresholds and the generating thresholds. It shows that the estimates with the adjustment generally have smaller difference from their generating values. Figure 7 shows the category characteristic curves of Item 8 from these two analyses. It shows that although the curves for the other categories are very similar in the two analyses, the curve for category 6 is noticeably narrower with the adjustment, making the estimate of the last threshold closer to the generating value. To highlight this point, a vertical line is drawn at the generated value of 13.808 for the last threshold. It is relevant to note that despite null and low frequency categories in both real and simulated data, the threshold estimates are in their natural order.

Table 8

Estimated principal component parameters of simulated data in Study 1.

(a) without adjustment

Item	Location	*SE*	Unit	*SE*	Skewness	*SE*	Kurtosis	*SE*
1	-4.427	0.082	3.039	0.040	0.085	0.009	0.007	0.002
2	0.004	0.070	2.235	0.020	0.011	0.004	-0.008	0.001
3	0.702	0.071	2.171	0.020	-0.004	0.004	-0.004	0.001
4	-0.236	0.072	2.350	0.025	0.082	0.005	-0.001	0.001
5	1.116	0.070	2.412	0.021	0.054	0.003	0.004	0.001
6	1.286	0.071	2.149	0.020	0.019	0.003	-0.005	0.001
7	0.542	0.072	1.977	0.021	0.003	0.004	-0.012	0.001
8	1.013	0.070	2.244	0.020	0.033	0.003	-0.002	0.001
Mean	0.000	0.072	2.322	0.023	0.035	0.004	-0.003	0.001
Std Dev	1.866	0.004	0.318	0.007	0.035	0.002	0.006	0.000

(b) with adjustment

Item	Location	*SE*	Unit	*SE*	Skewness	*SE*	Kurtosis	*SE*
1	-4.310	0.082	3.006	0.039	0.080	0.009	0.006	0.002
2	-0.014	0.070	2.177	0.020	0.004	0.004	-0.008	0.001
3	0.715	0.071	2.125	0.020	-0.009	0.003	-0.004	0.001
4	-0.231	0.072	2.289	0.025	0.072	0.005	-0.002	0.001
5	1.117	0.070	2.361	0.021	0.049	0.003	0.003	0.001
6	1.275	0.071	2.096	0.020	0.013	0.003	-0.005	0.001
7	0.434	0.071	1.890	0.020	-0.008	0.004	-0.013	0.001
8	1.015	0.070	2.194	0.020	0.027	0.003	-0.002	0.001
Mean	0.000	0.072	2.267	0.023	0.029	0.004	-0.003	0.001
Std Dev	1.822	0.004	0.330	0.007	0.035	0.002	0.006	0.000

Simulation Study 2: A Full Rank Model with Five Response Categories

The simulation study in Section 5 demonstrated the working of the algorithm in the case where the model was of reduced rank in the sense that less parameters than the maximum possible were estimated. This Section investigates the quality of the algorithm when the model is of full rank. In this study, responses to ten items with five response categories were simulated using the RM of Eq. (1). The location values of the items were evenly spaced on the interval [-3.0,3.0]. The thresholds of item *i* were randomly generated in the interval $[\delta_i - 3.0, \delta_i + 3.0]$ and then placed in their natural order. One thousand person locations which were normally distributed with mean of 0.5 and variance of 1.0 were generated. The generating item locations and uncentralized thresholds for the ten items are shown in Table 10. The item locations, which are the means of the uncentralized thresholds for each item, are included in the tables to study their stability. The frequencies for different categories of the simulated data are shown in Table 11. It is evident that null categories appear in Items 1 and 2, which had the lowest thresholds and which were not targeted well to the population of persons. In addition, low frequencies appear in the highest category for Items 9 and 10, which had the greatest thresholds and were also not well targeted to the population.

Two analyses were again conducted on the simulated data, one without, and one with the adjustment for null categories. The convergence criterion of 0.001 was used in the estimation. The estimated thresholds from these two analyses are shown in Table 12. The RMSEs for the two sets of thresholds estimates are 0.513 and 0.315, without and with the adjustment for null categories respectively.

Table 12 shows that the two sets of estimated thresholds are quite similar except for the items

Table 9

Estimated uncentralized thresholds of simulated data in Study 1.

(a) without adjustment

	Thresholds						
Item	1	2	3	4	5	6	7
1	-15.75	-10.66	-5.44	0.79	8.92		
2	-12.15	-9.86	-5.59	-0.26	5.20	9.87	12.81
3	-12.02	-8.42	-4.00	0.81	5.56	9.83	13.16
4	-10.32	-7.67	-3.92	0.84	6.55	13.11	
5	-12.15	-8.11	-4.26	-0.19	4.54	10.34	17.64
6	-10.47	-7.89	-3.93	0.84	5.83	10.46	14.16
7	-9.73	-8.86	-4.96	0.46	5.92	9.94	11.02
8	-11.26	-8.18	-4.28	0.23	5.13	10.21	15.24

(b) with adjustment

	Thresholds						
Item	1	2	3	4	5	6	7
1	-15.53	-10.49	-5.27	0.91	8.82		
2	-11.97	-9.69	-5.42	-0.12	5.23	9.66	12.21
3	-11.85	-8.24	-3.83	0.93	5.59	9.67	12.74
4	-10.14	-7.53	-3.75	1.00	6.50	12.54	
5	-11.97	-7.94	-4.10	-0.05	4.57	10.17	17.14
6	-10.30	-7.72	-3.76	0.97	5.84	10.27	13.62
7	-9.55	-8.72	-4.80	0.62	5.94	9.58	9.96
8	-11.09	-8.01	-4.11	0.36	5.16	10.04	14.75

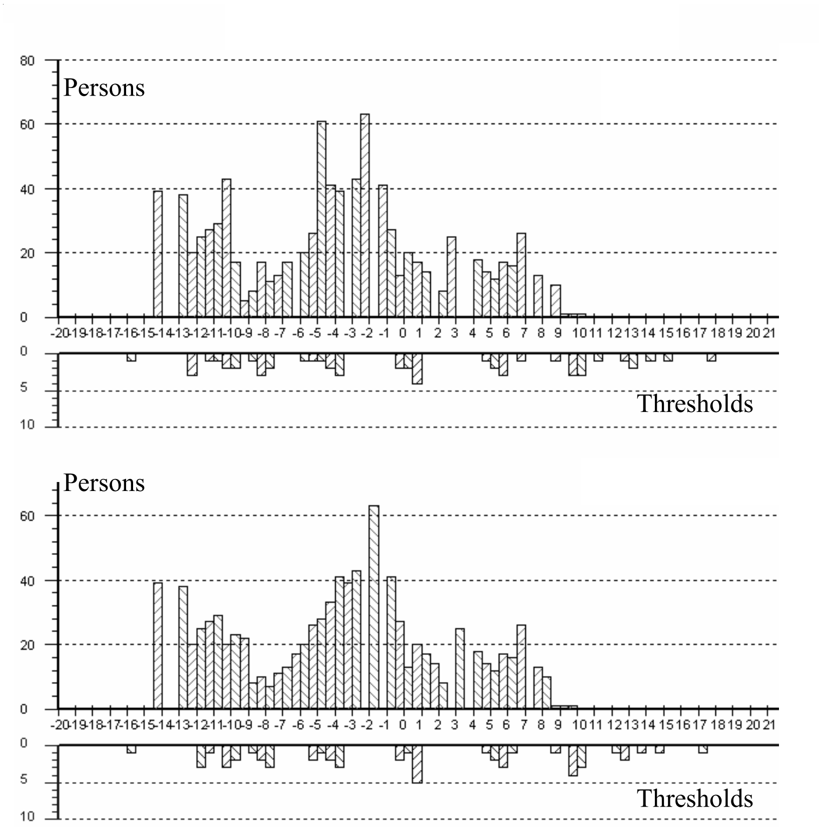

Figure 5. The distribution of estimated person locations and thresholds (top: without adjustment; bottom: with adjustment) in Study 1.

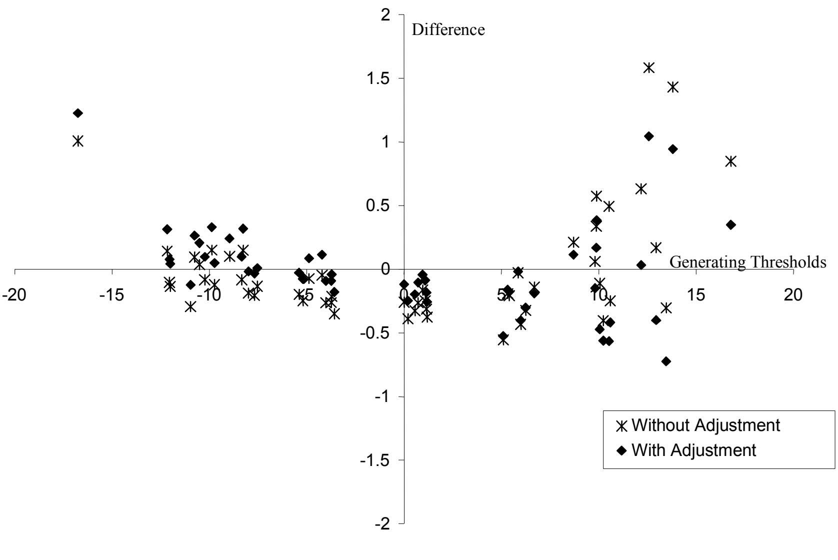

Figure 6. The generating thresholds verses the difference between the corresponding estimated thresholds and the generating thresholds in Study 1.

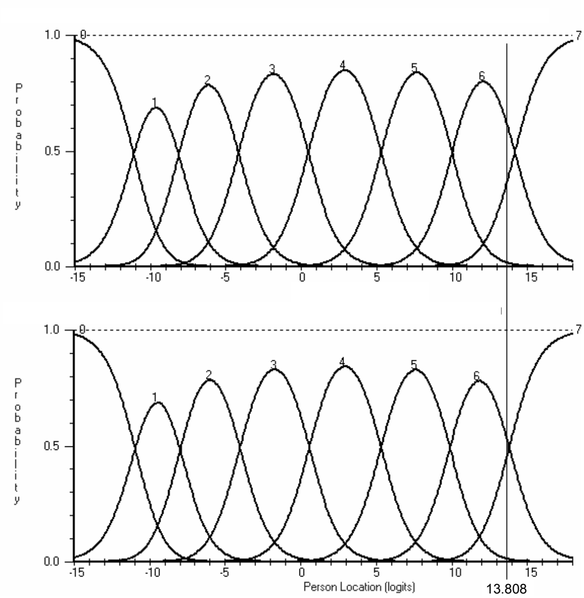

Figure 7. Item characteristic curves of Item (top: without adjustment; bottom: with adjustment) 8 in Study 1.

Table 10

The generating values of item locations and thresholds in Study 2

Item	Generating Location	Generated thresholds			
		1	2	3	4
1	-3.000	-4.08	-3.06	-2.86	-2.01
2	-2.333	-3.70	-2.63	-2.10	-0.90
3	-1.667	-3.52	-2.05	-2.01	0.90
4	-1.000	-2.99	-1.54	0.11	0.43
5	-0.333	-1.15	-0.76	-0.29	0.87
6	0.333	-0.69	0.02	0.32	1.69
7	1.000	-1.03	-0.69	1.60	4.12
8	1.667	1.04	1.28	1.73	2.62
9	2.333	-0.02	2.46	2.62	4.27
10	3.000	1.74	2.53	3.10	4.64

Table 11

Category frequencies of the simulated data in Study 2.

Item	Categories				
	0	1	2	3	4
1	0	0	8	105	887
2	0	7	34	211	748
3	2	17	63	501	417
4	15	78	222	268	417
5	62	104	180	342	312
6	136	188	207	287	182
7	89	178	486	230	17
8	464	225	151	104	56
9	367	486	95	46	6
10	694	228	58	15	5

with null categories. For both Items 1 and 2 which have null categories, in particular, the estimated thresholds with the adjustment are much closer to their corresponding generated values than those estimated without the adjustment. Figure 8 shows this comparison. It can be seen that although the estimates of the thresholds in the middle of the continuum are quite robust, the greatest effect of the adjustment is on the thresholds at the two ends of the continuum. Overall, the estimated thresholds with the adjustment are closer to the diagonal line than those estimated without the adjustment. In addition, it is seen by comparing Table 12 with Table 10 that the estimates of item locations, which are the means of the uncentralized thresholds of each item, are closer to the generating values when the adjustment is applied. In addition, it is evident that the estimate of one threshold at the other end of the continuum is noticeably different from its generating value regardless of the algorithm used. This is the last threshold of Item 10, and it is no coincidence that the last two categories of this item have low frequencies of 5 and 15 respectively, as shown in Table 11.

Similar overall results were observed when the simulation study was repeated with the same design but different random seeds for data simulation. However, the presence of null categories may be different with different random seeds. Therefore, it is not meaningful to present the average of the estimated thresholds over the repetitions of the simulation studies.

Summary and Discussions

The conditional pairwise maximum likelihood algorithm for estimating parameters in the reparameterized Rasch model (Andrich and Luo, 2003) for ordered polytomous responses provides estimates of the thresholds in the presence of null categories. This paper describes an adjustment that can be made to the algorithm in the presence of null categories with a view to improving the stability of the estimates. Two simulation studies and a real example show the degree of efficiency of the adjustment is made when either the full rank or the reduced rank reparameterized RM is applied.

The null category adjustment of this paper is applicable only to items in which all categories are logically observable. Often the cause for this phenomenon is poor targeting of the thresholds of the item to the person distribution. However, it is immaterial in the algorithm where a null category appears in the item, although in real situations null categories are often extreme, that is, the lowest category 0 or the highest category *m* and this is the case described in the real data and in the two simulation studies. A case where null categories might be said to occur by design

Table 12

Recovery of thresholds in Study 2.

	Without adjustment					With adjustment				
		Thresholds					Thresholds			
Item	Location	1	2	3	4	Location	1	2	3	4
1	-3.57	-3.64	-5.64	-3.17	-1.82	-3.08	-3.60	-3.53	-3.27	-1.93
2	-2.62	-4.64	-2.93	-2.15	-0.77	-2.22	-2.72	-3.03	-2.26	-0.89
3	-1.61	-3.39	-2.00	-2.07	1.01	-1.66	-3.24	-2.09	-2.19	0.89
4	-0.79	-2.56	-1.28	0.15	0.54	-0.89	-2.64	-1.40	0.03	0.42
5	-0.18	-1.16	-0.62	-0.13	1.21	-0.29	-1.26	-0.73	-0.25	1.09
6	0.49	-0.52	0.19	0.47	1.84	0.38	-0.64	0.07	0.35	1.72
7	1.13	-0.95	-0.63	1.84	4.26	1.01	-1.06	-0.75	1.72	4.13
8	1.72	1.12	1.38	1.84	2.53	1.60	1.00	1.26	1.72	2.41
9	2.47	0.17	2.82	2.51	4.38	2.33	0.05	2.70	2.39	4.19
10	2.95	1.94	2.82	3.55	3.49	2.83	1.82	2.70	3.42	3.39

is where a category system is designed for a wide part of the continuum (e.g. Years 3 to 10) in assessment of some performance, but only students in Year 3 are being assessed. The categories given to the judges might not include the extreme categories at the end reflecting greater performance on the assumption that students will not reach them. In principle, this is a design problem and this paper does not adress such an example, though in principle the algorithm can be applied in such situations.

When no adjustment is made, the estimates of the extreme thresholds corresponding to the null categories are stretched away from the arbitrary constraint of a mean value of 0 for the threshold estimates. The adjustment regresses the estimates of these extreme thresholds to more realistic values in the case of real data to values closer to the generating values in simulated data. Even though the estimation procedure of this paper can be used to produce the estimates of the thresholds in the presence of null categories, the precision of the estimates of the thresholds remains affected by the lack of responses in the null categories. In general, a poorly targeted test cannot be fully rescued by the procedure described in this paper. Therefore, the adjustment cannot obviate the need for professional judgment based on experience and results from analysing relevant data that is better targeted before concluding on the validity of the estimates for use in any particular circumstances. Model fit, as usual, plays a role in the quality of the estimates.

A more extreme case of mistargeting with multimodal person distributions than those shown in Figures 1 and 4, with modes further apart and with narrower distributions, could result in null categories which are in the middle rather at the extreme of the categories. Both the original and adjusted algorithms work in this case too. It remains to be investigated under which conditions the adjustment for null categories provides better estimates than the algorithm without adjustment. It is conceivable that with a middle null category, any stretching in estimates from one direction, evident in the case of an extreme null category, is compensated for by stretching from the other direction.

One other case where null categories can appear and which needs further study is where relatively few persons respond to items which have many categories. For this case, the targeting may be sound and therefore null categories may appear throughout the continuum rather than only at extremes. From preliminary studies, a

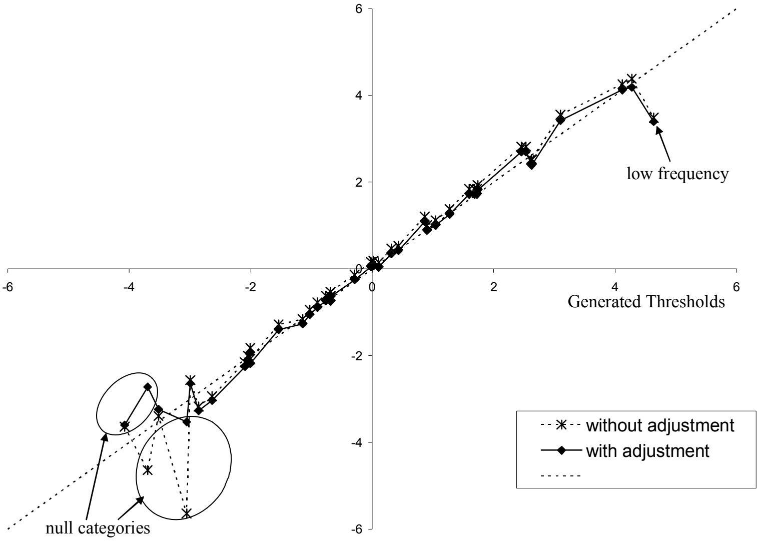

Figure 8. The plot of estimated thresholds against their generated values in Study II.

reduced rank model with the null category adjustment in this case appears to provide a better recovery of the thresholds compared to the case where null categories are present because of poor targeting. However, this is yet to be confirmed with a larger number of studies.

Poor targeting which produces null categories is also likely to produce categories with low frequencies. The simulation studies in this paper revealed, not surprisingly, that not only null categories, but also categories with low frequencies, affect the quality of the recovery of the thresholds. The adjustment for null categories presented in this paper does not adjust for categories with low frequencies. Some preliminary studies suggest that the quality of the threshold estimates for categories with low frequencies may be improved with an adjustment similar to that for null categories. However, there is a problem in deciding when a frequency is sufficiently small that the adjustment should be applied. This, too, requires further study.

Finally, it is pointed out that the analyses of the real and simulated data in this paper, in which the threshold estimates can retain their natural order in the presence of null or low frequency categories, demonstrates that null or low frequency categories do not necessarily lead to thresholds estimates in which there are order reversals. This is also shown in Andrich (2004) in a case of two items with 11 categories each and a bimodal person distribution in which one of the central categories of one of the items has a frequency of 0 and one of the central categories of the other item has a frequency of 2.

Acknowledgments

This Research was funded in part by grants from the Australian Research Council and Murdoch University. The permission to use the real data in this paper was granted by the Department of Education and Training, Western Australia.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-574.

Andrich, D. (l985). An elaboration of Guttman scaling with Rasch models for measurement. In N. Brandon-Tuma (Ed.), *Sociological Methodology* (pp. 33-80). San Francisco: Jossey-Bass.

Andrich, D. (1995). Distinctive and incompatible properties of two common classes of IRT models for graded responses. *Applied Psychological Measurement, 19*, 101- 119.

Andrich, D. (2004) The Rasch model explained. In S. Alagumalai, D. D. Curtis, and N. Hungi (Eds.) *Applied Rasch measurement: A book of exemplars* (pp. 27-59). Springer-Kluwer.

Andrich, D., Sheridan, B. S., and Luo, G. (2003). *RUMM2020: Rasch unidimensional models for measurement.* Perth, Australia: RUMM Laboratory.

Andrich, D., and Luo, G. (2003). Conditional pairwise estimation in the Rasch model for ordered response categories using principal components. *Journal of Applied Measurement, 4*, 205-221.

Andersen, E. B. (1971). The numerical solution of a set of conditional estimation equations. *Journal of the Royal Statistical Society. Serial B*, *34*, 42-50.

Dempster, A. P., Laird N. M. and Rubin D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Serial B, 39*, 1-38.

Choppin, B. (1985). A fully conditional estimation procedure for Rasch model parameters. *Evaluation in Education, 9*, 29-42.

Garner, M., and Engelhard, G. (2002). An eigenvector method for estimating item parameters of the dichotomous and polytomous Rasch models. *Journal of Applied Measurement, 3*, 107-112.

Guttman, L. (1950). The principal components of scale analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, and J. A. Clausen (Eds.), *Measurement and Prediction* (pp.312-361). New York: Wiley.

Jansen, P. G. W., and Roskam, E. E. (1986). Latent trait models and dichotomization of graded responses. *Psychometrika*, *51*, 69-91.

Little, R. J. A., and Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research, 18,* 292-326.

Pedler, P. (1987). *Accounting for psychometric dependence with a class of latent trait models*. Unpublished Ph.D. thesis, Department of Education, The University of Western Australia.

Wilson, M., and Masters, G. N. (1993). The partial credit model and null categories. *Psychometrika*, *58*, 87-99.

Wright, B. D., and Masters, G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.

Wright, B. D.,Condon, R., and Schultz, M. (1988). MSTEPS[computer program]. Chicago: University of Chicago, MESA Psychometric Laboratory.

Wright, B. D., Masters, G. N., and Ludlow, L. H. (1982). CREDIT. [computer program]. Chicago: MESA Press.