DOCUMENT RESUME

ED 383 721 TM 023 142

AUTHOR	Chang, Hua-Hua; Mazzeo, John
TITLE	The Unique Correspondence of Item Response Functions and Item Category Response Functions in Polytomously Scored Item Response Models.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-93-53
PUB DATE	Nov 93
NOTE	25p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Classification; *Item Response Theory; *Scores; Scoring; *Test Items
IDENTIFIERS	Graded Response Model; Item Category Response Functions; Item Response Functions; Partial Credit Model; *Polytomous Scoring; *Rasch Model; Test Characteristic Curve

ABSTRACT

The item response function (IRF) for a polytomously scored item is defined as a weighted sum of the item category response functions (ICRF, the probability of getting a particular score for a randomly sampled examinee of ability theta). This paper establishes the correspondence between an IRF and a unique set of ICRFs for two of the most commonly used polytomous item response theory (IRT) models (the partial credit models and the graded response model). Specifically, a proof of the following assertion is provided for these models: If two items have the same IRF, then they must have the same number of categories; moreover, they must consist of the same ICRFs. As a corollary, for the Rasch dichotomous model, if two tests have the same test characteristic function, then they must have the same number of items. Moreover, for each item in one of the tests, an item in the other test with an identical IRF must exist. Theoretical as well as practical implications of these results are discussed. (Contains 11 references.) (Author)

RESEARCH REPORT

THE UNIQUE CORRESPONDENCE OF ITEM RESPONSE FUNCTIONS AND ITEM CATEGORY RESPONSE FUNCTIONS IN POLYTOMOUSLY SCORED ITEM RESPONSE MODELS

Hua-Hua Chang
John Mazzeo

Educational Testing Service
Princeton, New Jersey
November 1993

The Unique Correspondence of Item Response Functions and Item Category Response Functions in Polytomously Scored Item Response Models [1]

Hua-Hua Chang & John Mazzeo[2]
Educational Testing Service

November 5, 1993

[1]This research was supported by Educational Testing Service Allocation Projects No.79409 and No.79413.

[2]The authors wish to thank John Donoghue, Ming-Mei Wang, Rebecca Zwick, and Zhiliang Ying for their useful comments and discussions. The authors' mailing address: Hua-Hua Chang and John Mazzeo, Educational Testing Service, Rosedale Road, Princeton, NJ 08541.

ABSTRACT

The item response function (IRF) for a polytomously scored item is defined as a weighted sum of the item category response functions (ICRF, the probability of getting a particular score for a randomly sampled examinee of ability θ). This paper establishes the correspondence between an IRF and a unique set of ICRFs for two of the most commonly used polytomous IRT models (the partial credit models and the graded response model). Specifically, a proof of the following assertion is provided for these models: **If two items have the same IRF, then they must have the same number of categories; moreover, they must consist of the same ICRFs.** As a corollary, for the Rasch dichotomous model, if two tests have the same test characteristic function (TCF), then they must have the same number of items. Morover, for each item in one of the tests, an item in the other test with an identical IRF must be exist. Theoretical as well as practical implications of these results are discussed.

Key words: item response theory, polytomous item, partial credit model, generalized partial credit model, graded response model, invariance, ordered categories.

1 Introduction

With the increased use of Item Response Theory (IRT) models for ordinally scored polytomous items in educational achievement tests (e.g., essays, performance tasks, and testlets), an important question presents itself about the correspondence between an item response function (IRF, defined as the regression of item score on ability) and sets of item category response functions (ICRF, defined as the probability curves for each of an item's possible response categories). For commonly used dichotomous parametric IRT models (like the 1-, 2-, and 3-parameter logistic model), two items with identical IRFs have identical item parameters. However, for polytomous models, like the partial credit (Masters, 1982), generalized partial credit (Muraki, 1992), and graded response (Samejima, 1969, 1972) models, the item structure of an (m+1)-category item is determined by m ICRFs, each of which is determined by a set of parameters. It is not clear whether an IRF for polytomously scored items corresponds to a unique set of ICRFs, and hence, a unique parametric structure. Whereas in the dichotomous case, items with identical IRFs must have identical conditional item score distributions at all levels of ability, two polytomous items with identical IRFs could exhibit conditional item score distributions that differ substantially beyond the first moment at one or more levels of ability.

This article investigates the relationship between IRFs and ICRFs for the most commonly used ordinal item response models the partial credit models and the graded response model. Specifically, we provide proofs of the following assertion for these models: **If two items have the same IRF, then they must have the same number of score categories. Moreover, they must consist of the same ICRFs.** An example is provided to show that this uniqueness assertion does not hold in general. Lastly, some theoretical and practical implications of these results are discussed.

2 The Basic Notation and Problem

Consider a test consisting of items intended to measure some unidimensional proficiency of interest. Assume that examinee responses to the items can be categorized into one of a set of $(m+1)$ ordered categories. Let X_j be the score for a randomly selected examinee on the jth item; $X_j = k, 0 \leq k \leq m$, and let Θ denote the ability for a randomly chosen examinee. Let $P_{j,k}(\theta)$ denote the probability of getting score k on item j for a randomly sampled examinee with ability θ, that is,

$$P_{j,k}(\theta) \equiv Prob\{X_j = k | \Theta = \theta\}. \tag{1}$$

In other words

$$X_j = \begin{cases} 0, & \textit{with probability } P_{j,0}(\theta), \\ 1, & \textit{with probability } P_{j,1}(\theta), \\ 2, & \textit{with probability } P_{j,2}(\theta), \\ \vdots & \vdots \\ m-1, & \textit{with probability } P_{j,m-1}(\theta), \\ m, & \textit{with probability } P_{j,m}(\theta). \end{cases} \tag{1a}$$

$P_{j,k}(\theta)$ is referred to as item category response function (ICRF) (Muraki, 1992). Assume the domain of θ to be $(-\infty, \infty)$ or some subinterval of this range. Further, assume

$$\sum_{k=0}^{m} P_{j,k}(\theta) = 1. \tag{2}$$

If $m = 1$, then X_j is dichotomous and the model defined in $(1a)$ can be equivalently written as

$$X_j = \begin{cases} 0, & \textit{with probability } 1 - P_j(\theta), \\ 1, & \textit{with probability } P_j(\theta). \end{cases}$$

where $P_j(\theta) = P_{j,1}(\theta)$, the IRF for item j. Note that for dichotomous X_j, the $P_{j,1}(\theta)$ can also be viewed as the regression of item score on proficiency, i.e.

$$P_{j,1}(\theta) = E\{X_j|\theta\}.$$

Analogously, one can define IRFs for items with $m > 1$ categories as the regression of item score on proficiency. For an item with $m+1$ score categories, the IRF is defined as

$$\begin{aligned} E\{X_j|\theta\} &= P_{j,1}(\theta) + 2P_{j,2}(\theta)+, ..., +(m-1)P_{j,m-1}(\theta) + mP_{j,m}(\theta), \\ &= \sum_{k=1}^{m} kP_{j,k}(\theta). \end{aligned} \tag{3}$$

Thus, the item response function for a polytomously scored item is a weighted sum of ICRFs.

In the dichotomous context, there is no distinction between an item's IRF and its parametric structure. In the polytomous context, however, an item's parametric structure corresponds to a set of ICRFs and the IRF is a weighted sum of these ICRFs. The IRF may not be uniquely determined. There may exist n ICRFs $Q_{i,1}(\theta), Q_{i,2}(\theta), ..., Q_{i,n}(\theta)$, which are different from those in (3), but produce an equivalent IRF. Specifically, it could be the case that

$$E\{X_j|\theta\} = \sum_{k=1}^{m} kP_{j,k}(\theta) = \sum_{l=1}^{n} lQ_{i,l}(\theta), \qquad for\ all\ \theta, \tag{4}$$

where n is not necessarily equal to m. Hence, investigations of the relationship between the IRF and the ICRFs are important.

3 The Basic Results

In the following section we will investigate the correspondence between IRFs and ICRFs for the most commonly used polytomous models: (a) the partial credit models of Masters (1982) and Muraki (1992), and (b) the graded response model of Samejima (the homogeneous case in Samejima, 1969 and 1972). Specifically, we will provide proofs that, for each model, if (4) holds under the definition of (3), then

$$n = m \quad and \quad P_{j,k}(\theta) = Q_{i,k}(\theta) \;\; for\ all\ \theta, \qquad k = 0, ..., m. \tag{5}$$

Thus, the item structure (i.e., all ICRFs) is uniquely identified by the IRF.

Before going through the major proofs, we will include some explanations and background information about these models. For convenience, the item subscripts are suppressed.

3.1 Partial Credit Models

A partial credit model was proposed by Masters (Masters, 1982) for items with scores that have $m+1$ ordered levels, $0, 1, ..., m$ (m may vary from item to item). Let X be the item score. The general expression of the ICRF, the probability of a randomly sampled examinee with ability θ obtaining score k, is given by

$$P_k(\theta) = \frac{\exp \sum_{i=0}^{k}(\theta - b_i)}{\sum_{l=0}^{m} \exp \sum_{i=0}^{l}(\theta - b_i)}, \qquad k = 0, 1, ..., m \tag{6}$$

where $\sum_{k=0}^{0}(\theta - b_i) \equiv 0$ for notational convenience. The parameter b is referred to as step difficulty (Masters, 1982) and it governs how likely it is that a person with ability θ will reach level k rather than level $k-1$, for $k = 1, ..., m$.

A generalized partial credit model was obtained from the partial credit model by Muraki (Muraki, 1992) by incorporating a slope parameter for each item. The ICRF for an item can be expressed:

$$P_k(\theta) = \frac{\exp \sum_{i=0}^{k} \alpha(\theta - b_i)}{\sum_{l=0}^{m} \exp \sum_{i=0}^{l} \alpha(\theta - b_i)}, \qquad k = 0, 1, ..., m \tag{7}$$

where α is a slope parameter, which may vary from item to item, and the b-parameters have the same meaning as those in the partial credit model (6). It should be evident that Masters' partial credit model is a special case of the generalized partial credit model with $\alpha = 1$. Note that when $m = 1$, the generalized partial credit model is equivalent to the 2-parameter logistic (2PL) model. If, in addition, $\alpha = 1$, the generalized partial credit model is equivalent to the Rasch model for dichotomously scored items. It should also be noted that the partial credit models are actually special versions of the nominal model proposed by Bock. (See Bock, 1972, and also see Thissen & Steinberg, 1986.)

Since the IRF is the expected value of X conditioning on θ, according to (3) and (7), the IRF for the generalized partial credit model can be expressed as

$$E\{X|\theta\} = \frac{\sum_{k=1}^{m} k \exp\{\sum_{i=1}^{k} \alpha(\theta - b_i)\}}{1 + \sum_{k=1}^{m} \exp\{\sum_{i=1}^{k} \alpha(\theta - b_i)\}}. \tag{8}$$

Suppose that $E\{X|\theta\}$ can be expressed by another set of ICRFs, i.e., there exist a_1,...,a_n, and β such that

$$E\{X|\theta\} = \frac{\sum_{k=1}^{n} k \exp\{\sum_{i=1}^{k} \beta(\theta - a_i)\}}{1 + \sum_{k=1}^{n} \exp\{\sum_{i=1}^{k} \beta(\theta - a_i)\}}. \tag{9}$$

As shown in Theorem 1, if the expressions on the right-hand side of (8) and (9) are equal, then the number of categories must be the same and the two sets of parameters must be identical.

First, we consider the restricted case where $\alpha = \beta = 1$ (i.e., Masters' partial credit model).

Lemma 1 *Let $m, n \geq 1$ and $a_1, \ldots, a_n$, $b_1, \ldots, b_m$, be constants. If*

$$\frac{\sum_{k=1}^{m} k \exp\{\sum_{i=1}^{k}(\theta - b_i)\}}{1 + \sum_{k=1}^{m} \exp\{\sum_{i=1}^{k}(\theta - b_i)\}} = \frac{\sum_{k=1}^{n} k \exp\{\sum_{i=1}^{k}(\theta - a_i)\}}{1 + \sum_{k=1}^{n} \exp\{\sum_{i=1}^{k}(\theta - a_i)\}} \tag{10}$$

for all θ, then $n = m$ and $a_i = b_i$, $i = 1, \ldots, m$.

Proof. Note that

$$\sum_{i=1}^{k}(\theta - b_i) = k\theta - \sum_{i=1}^{k} b_i$$

and

$$\frac{d}{d\theta} e^{\sum_{i=1}^{k}(\theta - b_i)} = k e^{\sum_{i=1}^{k}(\theta - b_i)}.$$

Therefore, (10) can be rewritten as

$$\frac{d}{d\theta}\{\log[1 + \sum_{k=1}^{m} \exp\{\sum_{i=1}^{k}(\theta - b_i)\}]\} = \frac{d}{d\theta}\{\log[1 + \sum_{k=1}^{n} \exp\{\sum_{i=1}^{k}(\theta - a_i)\}]\}. \tag{11}$$

Equation (11) implies

$$\log[1 + \sum_{k=1}^{m} \exp\{k\theta - \sum_{i=1}^{k} b_i\}] = \log[1 + \sum_{k=1}^{n} \exp\{k\theta - \sum_{i=1}^{k} a_i\}] + C \tag{12}$$

for some constant C that does not depend on θ. Letting $\theta \to -\infty$, we see that

$$\log(1) = \log(1) + C, \quad or \quad C = 0.$$

Exponentiating both sides of (12) and subtracting 1 from each side gives

$$\sum_{k=1}^{m} \exp\{-\sum_{i=1}^{k} b_i\}[\exp(\theta)]^k = \sum_{k=1}^{n} \exp(-\sum_{i=1}^{k} a_i)[\exp(\theta)]^k. \tag{13}$$

Setting $u_k = \exp(-\sum_{i=1}^{k} b_i)$, $v_k = \exp(-\sum_{i=1}^{k} a_i)$ and $x = \exp(\theta)$, (13) becomes

$$\sum_{k=1}^{m} u_k x^k = \sum_{k=1}^{n} v_k x^k, \quad for\ x > 0.$$

But if two polynomials are equal on a common interval, then $n = m$ and $u_k = v_k$, $k = 1, ..., m$. Thus

$$\sum_{i=1}^{k} b_i = \sum_{i=1}^{k} a_i, \quad k = 1, ..., m$$

or equivalently $a_i = b_i$, $i = 1, ..., m$. ■

We now consider the general case where α and β are not constrained (i.e., the generalized partial credit model).

Theorem 1 *Let $m, n \geq 1$ and α, β, $a_1, ..., a_n$, $b_1, ..., b_m$ be constants. If*

$$\frac{\sum_{k=1}^{m} k \exp\{\sum_{i=1}^{k} \alpha(\theta - b_i)\}}{1 + \sum_{k=1}^{m} \exp\{\sum_{i=1}^{k} \alpha(\theta - b_i)\}} = \frac{\sum_{k=1}^{n} k \exp\{\sum_{i=1}^{k} \beta(\theta - a_i)\}}{1 + \sum_{k=1}^{n} \exp\{\sum_{i=1}^{k} \beta(\theta - a_i)\}} \qquad (14)$$

for all θ, then $n = m$, $\alpha = \beta$, and $a_i = b_i$, $i = 1, ..., m$.

Proof: If $\alpha = \beta$, then (14) can be rewritten as (10) by replacing a_i with $a_i^* = a_i\alpha$ and b_i with $b_i^* = b_i\alpha$, and θ with $\theta^* = \theta/\alpha$. By Lemma 1, we must have $m = n$, and $a_i^* = b_i^*$, which implies $a_i = b_i$, $i = 1, ..., m$. So to prove Theorem 1, it is sufficient to prove that if (14) holds, we must have $\alpha = \beta$. Suppose $\alpha \neq \beta$, say, $\alpha < \beta$. (14) can be equivalently written as

$$\{\sum_{k=1}^{m} k \exp\{\sum_{i=1}^{k} \alpha(\theta - b_i)\}\{1 + \sum_{k=1}^{n} \exp\{\sum_{i=1}^{k} \beta(\theta - a_i)\}\} =$$
$$\{\sum_{k=1}^{n} k \exp\{\sum_{i=1}^{k} \beta(\theta - a_i)\}\{1 + \sum_{k=1}^{m} \exp\{\sum_{i=1}^{k} \alpha(\theta - b_i)\}\}. \qquad (15)$$

Letting $x = e^{\theta}$, (15) becomes

$$(A_1 x^{\alpha} + A_2 x^{2\alpha} +, ..., + A_m x^{m\alpha})(1 + B_1 x^{\beta} +, ..., + B_n x^{n\beta}) =$$
$$(C_1 x^{\beta} + C_2 x^{2\beta} +, ..., + C_n x^{n\beta})(1 + D_1 x^{\alpha} +, ..., + D_m x^{m\alpha}), \ for\ x > 0.$$

where

$$A_j = je^{-\alpha \sum_{i=1}^{j} b_i}, \quad D_j = e^{-\alpha \sum_{i=1}^{j} b_i}, \quad j = 1, ..., m$$

and

$$B_k = ke^{-\beta \sum_{i=1}^{k} a_i}, \quad C_k = e^{-\beta \sum_{i=1}^{k} a_i}, \quad k = 1, ..., n.$$

After multiplying the two "polynomials" on both sides of the equation, we have

$$A_1 x^{\alpha} +, ..., + A_m x^{m\alpha} + A_1 B_1 x^{\alpha+\beta} + A_2 B_1 x^{2\alpha+\beta} +, ..., + A_m B_n x^{m\alpha+n\beta} =$$
$$C_1 x^{\beta} +, ..., + C_n x^{n\beta} + C_1 D_1 x^{\alpha+\beta} + C_2 D_1 x^{\alpha+2\beta} +, ..., + C_n D_m x^{m\alpha+n\beta}. \quad (16)$$

Notice that the coefficients are all > 0, and the power function x^{α} only appears on the left hand side ($l.h.s.$) of (16). Subtracting the right from the left and collecting terms where appropriate, we have

$$c_{1,0} x^{\alpha} +, ..., + c_{m,0} x^{m\alpha} + c_{0,1} x^{\beta} +, ..., + c_{0,n} x^{n\beta} + c_{1,1} x^{\alpha+\beta} +, ..., + c_{m,n} x^{m\alpha+n\beta} = 0 \quad (17)$$

for all $x > 0$. The $l.h.s.$ of (17) is a linear combination of $(m+1)(n+1) - 1$ power functions of x. Since $\alpha < \beta$, the smallest exponent among all possible exponents in (16) is α. Since x^{α} only appears on the $l.h.s$ of (16), after we subtract the $r.h.s.$, the coefficient of x^{α} must remain unchanged. It is not difficult to verify that

$$c_{1,0} = A_1 = e^{-\alpha b_1} > 0. \quad (18)$$

If the $(m+1)(n+1) - 1$ exponents in (17) are all different,

$$x^{\alpha}, ..., x^{m\alpha}, x^{\beta}, ..., x^{n\beta}, x^{\alpha+\beta}, ..., x^{m\alpha+n\beta}$$

are linearly independent. [1] Thus, (17) holds if and only if for all the $c_{i,j}$'s are 0. This contradicts (18), and therefore, α must equal β.

If, on the other hand, some of the terms in (17) share a common exponent, by appropriate algebraic operations (additions or subtractions to those terms with common exponents of x), for some $k_1 \leq m$ and $k_2 \leq n$, (17) can be reduced to

$$C_1 x^{\alpha} +, ..., + C_{k_1,k2} x^{k_1\alpha + k2\beta} = 0, \tag{19}$$

where all the exponents are different. Since α is still the smallest exponent in (17), we still have $C_1 = c_{1,0} = e^{-\alpha b_1} > 0$, and (19) is still impossible by virtue of the same argument concerning the linear independence of power functions. Therefore α must equal β. ■

3.2 The Graded Response Model

The graded response model was proposed by Samejima (the homogeneous case in Samejima, 1969, 1972). Let X be the score of an item which has $m + 1$ category levels, $0, 1, .., m$. Suppose there are m boundaries

$$b_1 < b_2 < ... < b_m$$

on the θ scale. According to Samejima's graded response model, the probability that a randomly sampled person with ability θ produces a performance that is scored at or above the k-th level is determined by the boundary b_k and an item discrimination parameter,

[1] Let $f_1(x), f_2(x), ..., f_n(x)$ be n functions in (a, b). Assuming there exist n constants $c_1, c_2, ..., c_n$, not all zero, such that

$$c_1 f_1(x) + c_2 f_2(x) +, ..., + c_n f_n(x) = 0 \textit{ for every } x \textit{ in } (a, b), \qquad (**)$$

then $f_1(x), ..., f_n(x)$ are said to be linearly dependent on (a, b). If (**) holds if and only if all $c_1, ..., c_n$ are zeros, then $f_1(x), ..., f_n(x)$ are linearly independent on (a, b). Of course, a set of power functions with different exponents are linearly independent on any given interval.

say α. Let

$$Y_k = \begin{cases} 1, & if\ X \geq k, \\ 0, & otherwise. \end{cases}$$

Define

$$P_k^*(\theta) = Prob\{Y_k = 1|\theta\} = \frac{\exp(\alpha(\theta - b_k))}{1 + \exp(\alpha(\theta - b_k))} \qquad k = 1, ..., m \tag{20}$$

where b_k is referred to as the category boundary parameter and α is the discrimination parameter, and $P_k^*(\theta)$ is the probability of the event "score $\geq k$" for randomly sampled examinee with ability θ. Note that for $1 \leq k \leq m$, $P_k^*(\theta)$ is given by the 2PL model. By convention, $P_0^*(\theta) \equiv 1$ and $P_{m+1}^*(\theta) \equiv 0$.

The ICRF (i.e., $P\{X = k|\theta\}$) can be obtained by subtracting the 2PL item response functions:

$$P_k(\theta) = P_k^*(\theta) - P_{k+1}^*(\theta), \quad k = 0, 1, ..., m. \tag{21}$$

Equations (3), (21) and (20) can be used to get

$$\begin{aligned} E\{X|\theta\} &= \sum_{k=1}^{m} kP_k(\theta), \\ &= P_1^*(\theta) + P_2^*(\theta) +, ..., + P_m^*(\theta), \\ &= \sum_{k=1}^{m} \frac{\exp(\alpha(\theta - b_k))}{1 + \exp(\alpha(\theta - b_k))}. \end{aligned} \tag{22}$$

The above expression indicates that the item response function from the graded response model is the sum of m 2PL dichotomous item response functions, in which the difficulty parameters are all different and the discrimination parameters are the same. Now we can state our assertion about Samijima's graded response model:

If two items satisfy the graded response model, and if they have the same IRF, then they must have the same number of categories; moreover, they must consist of the same ICRFs.

The mathematical description of the above assertion is given by the following theorem.

Theorem 2 *Let α, β, $a_1, ..., a_n$, $b_1, ..., b_m$, be constants, and $a_1 < a_2 < ... < a_n$ and $b_1 < b_2 < ... < b_m$. If*

$$\sum_{i=1}^{m} \frac{\exp(\alpha(\theta - b_i))}{1 + \exp(\alpha(\theta - b_i))} = \sum_{i=1}^{n} \frac{\exp(\beta(\theta - a_i))}{1 + \exp(\beta(\theta - a_i))} \qquad (23)$$

for all θ, then $n = m$, $\alpha = \beta$ and $a_i = b_i$, $i = 1, ..., m$.

Proof: First we consider a special case $\alpha = \beta = 1$. Rewrite (23) as

$$\frac{d}{d\theta}\{\sum_{i=1}^{m} \log[1 + \exp(\theta - b_i)]\} = \frac{d}{d\theta}\{\sum_{i=1}^{n} \log[1 + \exp(\theta - a_i)]\}.$$

Thus

$$\log\{\prod_{i=1}^{m}[1 + \exp(\theta - b_i)]\} = \log\{\prod_{i=1}^{n}[1 + \exp(\theta - a_i)]\} + C \qquad (24)$$

for some constant C that does not depend on θ. Letting $\theta \to -\infty$, we see that

$$\log(1) = \log(1) + C, \quad or\ C = 0.$$

Exponentiating both sides of (24) gives

$$\prod_{i=1}^{m}[1 + \exp(\theta - b_i)] = \prod_{i=1}^{n}[1 + \exp(\theta - a_i)]. \qquad (25)$$

Set $x = \exp(\theta)$, $c_i = \exp(-b_i)$, and $d_i = \exp(-a_i)$. Then (25) becomes

$$\prod_{i=1}^{m}[1 + c_i x] = \prod_{i=1}^{n}[1 + d_i x]. \qquad (26)$$

Each side of (26) is a non-reducible factorization of the same polynomial. Note that

$$c_1 > c_2 > ... > c_m \quad and \quad d_1 > d_2 > ... > d_n.$$

According to the polynomial factorization law[2], $n = m$ and $c_i = d_i$, $i = 1, ..., m$, or equivalently

$$a_i = b_i, \quad i = 1, ..., m.$$

Now we consider $\alpha \neq \beta$, say, $\alpha < \beta$. Rewrite (23) as

$$\frac{(\sum_{i=1}^{m} c_i)x^{\alpha} + 2(\sum_{i<j} c_i c_j)x^{2\alpha} + ... + m(c_1...c_m)x^{m\alpha}}{1 + (\sum_{i=1}^{m} c_i)x^{\alpha} + (\sum_{i<j} c_i c_j)x^{2\alpha} + ... + (c_1...c_m)x^{m\alpha}} = \frac{(\sum_{i=1}^{n} d_i)x^{\beta} + 2(\sum_{i<j} d_i d_j)x^{2\beta} +, ... + n(d_1...d_n)x^{n\beta}}{1 + (\sum_{i=1}^{n} d_i)x^{\beta} + (\sum_{i<j} d_i d_j)x^{2\beta} +, ... + (d_1...d_n)x^{n\beta}} \quad (27)$$

where $c_i = \exp(-\alpha b_i)$ and $d_i = \exp(-\beta a_i)$. Multiplying the numerator on each side with the other side's denominator, and then subtracting the right from the left, we have

$$C_{1,0}x^{\alpha}+, ..., +C_{m,0}x^{m\alpha} + C_{0,1}x^{\beta}+, ..., +C_{0,n}x^{n\beta} + C_{1,1}x^{\alpha+\beta}+, ..., +C_{m,n}x^{m\alpha+n\beta} = 0 \quad (28)$$

for all $x > 0$. Except for coefficient differences, (28) is the same as (17). It is obvious that α is the smallest exponent in (28) and $C_{1,0} = \sum_{i=1}^{m} \exp(-\alpha b_i) > 0$. By the same discussion given for the generalized partial credit model, α must equal β. ■

4 Additional Results

The test characteristic function (TCF) for a set of items is commonly defined as the expected score on the item set conditioning on θ. For dichotomous IRT models, the TCF is defined as the sum of the IRFs for each of the items in the set. It is apparent from

[2]This is a general result in algebra: If polynomial $f(x)$ can be expressed:

$$f(x) = p_1(x)p_2(x)...p_m(x) = q_1(x)q_2(x)...q_n(x)$$

where $p_i(x)$ and $q_j(x)$ are non-reducible polynomials, then $n = m$, and after some possible order changes,

$$p_i(x) = e_i q_i(x), \quad i = 1, ..., m$$

where $e_i \neq 0$. (Note: In equation (26) $p_i(x) = 1 + c_i x$ and $q_i(x) = 1 + d_i x$, hence $e_i = 1$, $i = 1, ..., m$).

(22) that the IRF for the graded response model can be expressed as the sum of a set of dichotomous logistic IRF with a common slope. Thus, as a by-product, Theorem 2 implies the following corollary:

Corollary 1 *For the Rasch dichotomous model, if two tests have the same test characteristic function (TCF), then they must have the same number of items. Moreover, for each item in one of the tests, an item in the other test with an identical IRF must exist.*

The two theorems in Section 3 are based on the assumption that the score sequence for the item category levels is a sequence of integers

$$0, 1, ..., m. \tag{29}$$

As a matter of fact these theorems will also hold if the scores in (29) are replaced with any given ordered sequence

$$k_0 < k_1 < ... < k_m. \tag{30}$$

Note that, k_i in (30) can be viewed as an increasing function of i, i.e.,

$$k_i = G(i), \quad i = 0, 1, ..., m.$$

Therefore, the score sequence in (30) is viewed as a monotone transformation of (29).

Let X and Y be the scores for two items, where X is scored $0, 1, ..., m$ and Y is scored $0, 1, ..., n$. Let $P_k(\theta)$, $k = 0, .., m$, and $Q_k(\theta)$, $k = 0, ..., n$ be the ICRFs of X and Y respectively. For the partial credit model, the generalized partial credit model, and the graded response model, we have the following theorem:

Theorem 3 *If*

$$E\{G(X)|\theta\} = E\{G(Y)|\theta\} \quad for\ all\ \theta, \tag{31}$$

where G is any monotone transformation of the sequence $0, 1, 2, .., max\{n, m\}$, then

$$m = n, \quad and \quad P_k(\theta) = Q_k(\theta), \quad k = 0, ..., m.$$

Proof: Analogous to the proofs of Theorem 1 – 2, and omitted. ■

If one defined a monotone function G as

$$G(x) = x^l, \quad for\ l \geq 1,$$

then condition (31) in Theorem 3 becomes

$$E\{X^l|\theta\} = E\{Y^l|\theta\}, \quad for\ all\ \theta. \tag{32}$$

Since (32) clearly implies

$$E\{X|\theta\} = E\{Y|\theta\}, \quad for\ all\ \theta,$$

the immediate conclusion is that Theorem 3 includes Theorem 1 – 2 as special cases.

5 A Limitation on the Uniqueness Results

It should be noted that the uniqueness assertions proved in section 3 do not hold in general. If $Q_k(\theta)$ is not restricted to follow the parametric models considered here, one can construct $P_k(\theta)$ and $Q_k(\theta)$ such that (4) holds but (5) does not hold. The following example illustrates this fact.

Example 1 *Let $P_k(\theta)$, $k = 0, 1, 2, 3$, satisfy the partial credit model defined in (6). Define $Q_k(\theta)$ as following:*

$$Q_0(\theta) = P_0(\theta) + P_1(\theta)/4, \tag{33}$$

$$
\begin{aligned}
Q_1(\theta) &= P_1(\theta)/2, \\
Q_2(\theta) &= P_1(\theta)/4 + P_2(\theta), \\
Q_3(\theta) &= P_3(\theta).
\end{aligned}
$$

The reader can readily verify that in the above example,

$$P_k(\theta) \neq Q_k(\theta), \quad k = 0, 1, 2,$$

despite the fact that

$$\sum_{k=1}^{3} kP_k(\theta) = \sum_{k=1}^{3} kQ_k(\theta).$$

For certain choices of b_is, the shapes of $Q_k(\theta)$s defined in (33) satisfy the usual assumptions of ICRFs for polytomous IRT models (i.e., uni-modal functions of θ or strictly monotone functions of θ, etc., see Samejima, 1972). For example, if the $P_k(\theta)$s are defined as (6) with $b_1 = -0.91$, $b_2 = 0.98$, and, $b_3 = 0.19$, the resulting $Q_k(\theta)$s from (33) look much like ICRFs that arise in the partial credit model (see Figure 1). However, it should be particularly indicated that for many choices of b_1, b_2, and b_3, $Q_0(\theta)$ and $Q_2(\theta)$ will not be either strictly monotone or uni-modal, thus, the $Q_k(\theta)$ is not a member of the IRT parametric models family. The $Q_k(\theta)$s were explicitly constructed to demonstrate that one can obtain two different sets of ICRFs which have identical IRFs (see Figure 1).

Are there sufficient conditions under which the uniqueness assertion will hold for all polytomous IRT models? The following is a very restrictive case where X and Y are scored by $0, 1, ..., m$ and their 1-st to m-th moments are all equal (a very strong condition). That is, if

$$E\{X^l|\theta\} = E\{Y^l|\theta\}, \quad for\ l = 1, 2, ..., m, \quad and\ for\ all\ \theta, \tag{34}$$

then

$$P_k(\theta) = Q_k(\theta), \quad k = 0, ..., m.$$

The proof is rather straightforward. Equation (34) implies

$$\sum_{k=1}^{m} k^n [P_k(\theta) - Q_k(\theta)] = 0, \quad n = 1, .., m. \tag{35}$$

Since the coefficient matrix of the set equations defined in (35) is full rank, (35) holds if and only if $P_k(\theta) = Q_k(\theta)$ for $k = 1, ...m$.

It may seem desirable to specify general and less-restrictive nonparametric conditions under which the uniqueness assertions would hold for all polytomous IRT models. But, Example 1 suggests that it may not be possible to identify such a set of conditions. Nevertheless, the theorems established in the preceding sections do provide sufficient evidence for all well-defined ordinally scored parametric IRT models that an item's structure (i.e., all ICRFs) is uniquely identified by its IRF.

6 Conclusion

The main theorems of this paper establish the unique correspondence between an IRF and a set of ICRFs for the most commonly used ordinal polytomous IRT models, the restricted and generalized versions of the partial credit model and the graded response model. The results of this paper indicate that if both members of a pair of items follow one of the models stated above and have identical IRFs, then they also have identical ICRFs. An additional theorem established that, for any monotonically increasing function of item score, identity across items of the regression of this function on ability implies identity of ICRFs across items. It should be noted that the uniqueness assertions established in this article are applicable to other well defined parametric polytomous models, such as Andrich's (1978) rating scale model, and Bock's nominal model (some special cases), etc.

There are a number of potential uses of these results in both theoretical and practical work with the polytomous IRT models discussed here. One potential area is in test assembly. Many testing programs have sets of statistical specifications designed to ensure the construction of multiple interchangeable parallel test forms. The results reported here imply that two items following the above models can be treated as equivalent, provided that their IRFs coincide. From a psychometric point of view, this equivalence entails identical information functions as well as identical conditional (on θ) item score distributions. Thus, for tests using items that follow one of the models discussed here, it is sufficient of express statistical specifications for the assembly of test forms in terms of a set of target IRFs. Computer-based test assembly methods, in conjunction with graphical procedures to evaluate the match of estimated IRFs to target IRFs may provide an effective (and user-friendly) procedure for ensuring psychometric equivalence across alternate forms. Corollary 1 implies that for tests constructed using items that follow the Rasch dichotomous model, it is sufficient to express statistical specifications in terms of a target TCF. Test assembly procedures which involve the matching of actual TCFs to target TCFs may be particularly easy for test assemblers to work with.

The results presented here may also be helpful in providing a parsimonious conceptualization of item-parameter invariance assumptions for ordinally scored polytomous items that follow one of the models discussed here. Within the context of IRT for dichotomous items, assumptions concerning the invariance of IRFs are equivalent to assumptions of parametric invariance and are fundamental to many IRT applications. For example, assumptions about IRF invariance across different contexts are fundamental to the study of item context effects. As a second example, assumptions about IRF invariance across groups of students are fundamental to study of differential item functioning (DIF). The results presented above establish this same identity between an item's IRF and its parametric structure for ordinally scored polytomous items following commonly used IRT models. Consequently, a straightforward generalization from dichotomous IRT applications based on IRF invariance assumptions to situations that involve ordinally scored polytomous items may be possible.

The results presented here may be particularly useful with regard to the modeling and detection of DIF. A natural generalization of an IRT model-based definition of null DIF for an ordinally scored polytomous item would be to require that the regression of ordinal item score on ability be identical for two groups under study. Such a generalization would allow the extension of available dichotomous DIF methodologies, such as Shealy-Stout's (Shealy,& Stout, 1993) DIF analysis methods, to ordinally scored polytomous items in a fairly straightforward way (see Chang, Mazzeo, & Roussos, 1993).

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 516-573.

Bock, R. D. (1972). Estimating item parameters and latent ability when the responses are scored in two or more nominal categories. *Psychometrika, 37*, 29-51.

Chang, H., Mazzeo, J., & Roussos, L. (1993). *Extension of Shealy-Stout's DIF procedures to polytomously scored items.* Paper presented at the 1993 NCME meeting, April, 1993, Atlanta, GE.

Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement, 16*, 159-176.

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika, Monograph Supplement, No. 17.*

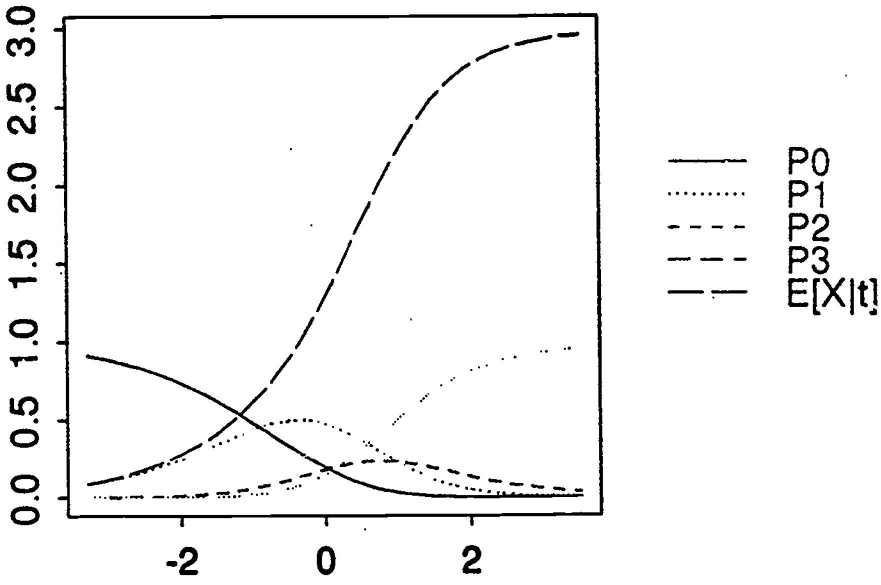
Samejima, F. (1972). A general model for free-response data. *Psychometrika Monograph, No 18.*

Shealy, R. & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*, 159-194.

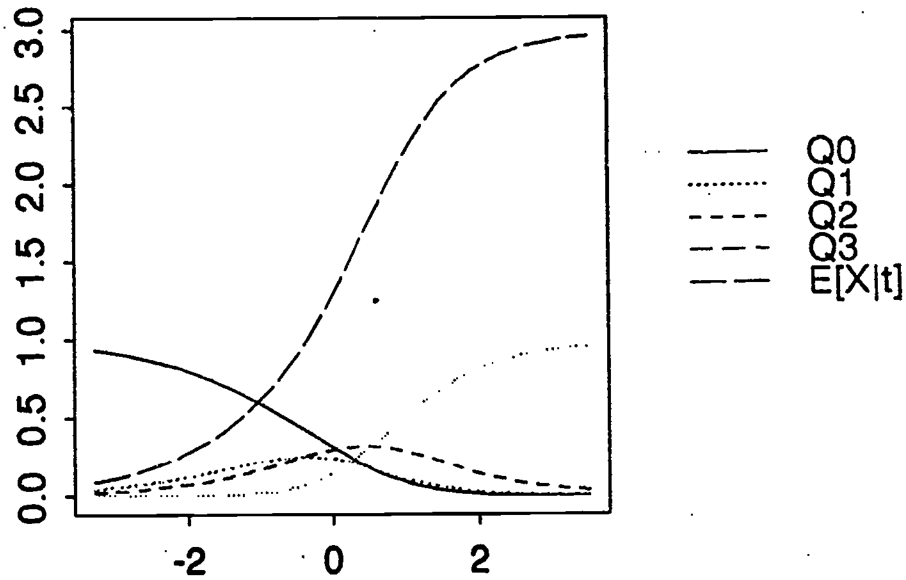
Thissen, D. & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika, 51*, 567-577.

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

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



Item 1 with ICRFs $P_0(\theta)$, $P_1(\theta)$, $P_2(\theta)$, and $P_3(\theta)$.



Item 2 with ICRFs $Q_0(\theta)$, $Q_1(\theta)$, $Q_2(\theta)$, and $Q_3(\theta)$.

Figure 1: An example of two items with different ICRFs but identical IRFs. In other words, $\sum_{k=1}^{3} kP_k(\theta) = \sum_{k=1}^{3} kQ_k(\theta)$ but $P_k(\theta) \neq Q_k(\theta)$, $k = 0, 1, 2$