PSYCHOMETRIKA—VOL. 73, NO. 4, 561–578
DECEMBER 2008
DOI: 10.1007/S11336-008-9071-2

# GRADED RESPONSE MODEL BASED ON THE LOGISTIC POSITIVE EXPONENT FAMILY OF MODELS FOR DICHOTOMOUS RESPONSES

FUMIKO SAMEJIMA

UNIVERSITY OF TENNESSEE

Samejima (Psychometrika 65:319–335, 2000) proposed the logistic positive exponent family of models (LPEF) for dichotomous responses in the unidimensional latent space. The objective of the present paper is to propose and discuss a graded response model that is expanded from the LPEF, in the context of item response theory (IRT). This specific graded response model belongs to the general framework of graded response model (Samejima, Psychometrika Monograph, No. 17, 1969 and No. 18, 1972; Handbook of modern item response theory, Springer, New York, 1997; Encyclopedia of Social Measurement, Academic Press, San Diego, 2004), and, in particular to the heterogeneous case (Samejima, Psychometrika Monograph, No. 18, 1972). Thus, the model can deal with any number of ordered polytomous responses, such as letter grades (e.g., A, B, C, D, F), etc.

For brevity, hereafter, the model will be called the LPEF graded response model, or LPEFG. This model reflects the opposing two principles contained in the LPEF for dichotomous responses, with the logistic model (Birnbaum, Statistical theories of mental test scores, Addison Wesley, Reading, 1968) as their transition, which provide a reasonable rationale for partial credits in LPEFG, among others.

Key words: item response theory, graded response model, logistic positive exponent family of models.

## 1. Logistic Positive Exponent Family of Models

Let $\theta$ be the unidimensional latent trait, or *ability*, which represents the target hypothetical construct, and is assumed to take on any real number. Let $g$ denote an item, which is the smallest unit of manifest entity in observable data (e.g., "$[3/7]-[11/13]=$" in an arithmetic test). The item characteristic function (ICF), $P_g(\theta)$, of a dichotomous item $g$ is defined as

$$P_g(\theta) \equiv \text{prob.}[U_g = 1|\theta], \tag{1}$$

where $U_g$ is the binary item score.

Birnbaum (1968) proposed the logistic model for dichotomous responses as a substitute for the normal ogive model (Lord and Novick, 1968) in which a simple sufficient statistic $t(v)$ for $V=v$ exists, such that

$$t(v) = \sum_{u_g \in v} a_g u_g, \tag{2}$$

where $u_g$ is the realization of $U_g$, $V$ is the response pattern or a sequence of (binary) item scores of $n$ dichotomous items with $v$ as its realization, and $a_g(>0)$ is the item discrimination parameter in the logistic (and normal ogive) model, whose ICF, defined by (1), is specified by

$$P_g(\theta) = \frac{1}{1+\exp[-Da_g(\theta - b_g)]} \equiv \Psi_g(\theta), \tag{3}$$

where $b_g$ is the item difficulty parameter and $D$ is a scaling factor usually set equal to 1.702 to approximate the ICF in the normal ogive model with the same values of item parameters. Hence,

Requests for reprints should be sent to Fumiko Samejima, Department of Psychology, 108 Conference Center Bldg., University of Tennessee, Knoxville, TN 37996-4100, USA. E-mail: fsamejim@utk.edu

the maximum likelihood estimate (MLE) of the examinee's ability level, $\hat{\theta}_v$, is obtainable as the solution of

$$\sum_{g=1}^{n} a_g P_g(\theta) = t(v).$$

Samejima (2000) observed that models with point-symmetric ICFs, such as those in the normal ogive and logistic models, have an intrinsic contradiction in the principle of ordering the examinees' MLEs of $\theta$, though the contradiction is degenerated in the logistic model in which the sufficient statistic $t(v)$ (2) does not include the difficulty parameter $b_g$.

We hear a casual remark such as: "So-and-so *succeeded* in solving that *difficult* question: what a *bright* man (woman)!" If we replace *succeeded* by *failed*, *difficult* by *easy*, and *bright* by *dumb* in the above comment, it will become a remark of somewhat different implications. Although both comments come out naturally, it should be noted that they are based on two opposing principles or philosophies. When the first philosophy is used, a greater credit is given for solving a difficult item than an easy item, and when the second principle is used, failure in solving an easy item is more strongly penalized than that in solving a difficult item. If our mathematical model provides a symmetric ICF, that is exemplified by the normal ogive model, these two opposing rules are mixed and selection of one of the two principles is determined by an incidental factor.

Samejima (2000, p. 324) illustrated an example of five dichotomous items following the normal ogive model, with the same discrimination parameter ($a_g = 1.0$) and equally distanced difficulty parameters ($b_g = -3.0, -1.5, 0.0, 1.5, 3.0$, respectively). We assume that local independence (Lord and Novick, 1968, Chap. 16) practically holds. Suppose two response patterns, which are sequences of binary item scores, are identical, except for items $g$ and $h$, where $u_g = 0$ and $u_h = 1$ in the first response pattern (0–1 response pattern), and $u_g = 1$ and $u_h = 0$ (1–0 response pattern) in the second response pattern. Choice of items $g$ and $h$ are arbitrary, but for convenience here, item 1 and item 5 are used. Note that if the principle of *giving greater credit for solving a more difficult question* is intrinsic in our mathematical model, a 0–1 response pattern should indicate greater ability than the corresponding 1–0 response pattern in each pair of response patterns, indicating the examinee's success and honor in solving the more difficult problem, and this order is reversed if our model is based on the principle of *penalizing failure in solving an easier question*, for with the 1–0 response pattern the examinee is cleared from the shame of failing in solving the easier problem.

There are eight possible *subresponse patterns* for items 2, 3, and 4, and they are (000), (001), (010), (100), (011), (101), (110), (111), and for each subresponse pattern there are 1–0 and 0–1 full response patterns. If those sixteen response patterns are picked up from the total set of thirty two response patterns in Table 1 of Samejima (2000) and the corresponding values of the maximum likelihood estimates of $\theta$ are compared within each pair, they are as follows:

| Subresp. pat. | 1–0 response pat. | 0–1 response pat. |
|---|---|---|
| **(000)** | MLE(10000) = −2.28385 | −**0.86577** = MLE(00001) |
| **(100)** | MLE(11000) = −0.75034 | −**0.27309** = MLE(01001) |
| **(010)** | MLE(10100) = −0.75012 | −**0.36062** = MLE(00101) |
| **(001)** | MLE(10010) = −0.34310 | −**0.19116** = MLE(00011) |
| **(110)** | MLE(11100) = **0.75034** | 0.27309 = MLE(01101) |
| **(101)** | MLE(11010) = **0.75021** | 0.36062 = MLE(01011) |
| **(011)** | MLE(10110) = **0.34310** | 0.19116 = MLE(00111) |
| **(111)** | MLE(11110) = **2.28385** | 0.86577 = MLE(01111) |

It can be seen that the value of MLE is greater for the 0–1 response pattern in each of the four pairs out of eight, and greater for the 1–0 response pattern in the other four pairs, meaning that the order of two MLEs within a pair is determined by the subresponse pattern, that is, an incidental factor.

The LPEF was proposed as a model that does not have this intrinsic contradiction, whose ICF is defined by

$$P_g(\theta) = \left[\Psi_g(\theta)\right]^{\xi_g}, \tag{4}$$

where $\Psi_g(\theta)$ is the logistic ICF given by (3), and the third item parameter $\xi_g(>0)$, called the *acceleration parameter*, has a critical role in this family of models. In ordering the examinees' MLEs of $\theta$, the ICF follows the principle of *penalizing failure in solving an easier item* when $0 < \xi_g < 1$, whereas it follows the opposing principle, that is, *greater credit is given for solving a more difficult item* when $\xi_g > 1$.

In proving the relationships between the parameter $\xi_g$ and the two opposing principles, Samejima (2000) defined the *model/item feature function* $S_g(\theta)$, such that

$$S_g(\theta) = \frac{P_g'(\theta)}{P_g(\theta)Q_g(\theta)},$$

where $P_g(\theta)$ is the ICF defined by (1),

$$Q_g(\theta) = 1 - P_g(\theta),$$

and $P_g'(\theta)$ is the first derivative of $P_g(\theta)$ with respect to $\theta$. Showing that this model/item feature function is strictly decreasing in $\theta$ for $\xi_g > 1$, equals a constant for $\xi_g = 1$, and strictly increasing in $\theta$ for $0 < \xi_g < 1$, the relationship described in the preceding paragraph is proved. For details, the reader is directed to Figs. 4a, b, and c (pp. 331–332) in Samejima (2000), where examples of $\xi_g = 2.0$, 1.0, and 0.3 are given.

Note that when $\xi_g = 1$ (4) becomes (3), namely, the logistic ICF. Thus, the ICF in the logistic model does not follow either of the two opposing principles, and can be interpreted as a *transition* from one principle to the other.

Figure 1 illustrates ICFs in the LPEF for seven dichotomous items with the common item parameters $a_g = 1.0$ and $b_g = 0.0$, respectively, and of seven different $\xi_g$'s, 0.3, 0.5, 0.8, 1.0, 1.5, 2.0, 3.0.

## 2. General Graded Response Model

### 2.1. Rationale

The general graded response model was proposed by Samejima (1969, 1972, 1997, 2004), which represents a family of mathematical models that deal with *ordered polytomous responses* in general. This expansion of IRT resulted in substantial enhancement of applicability of IRT for various areas of social and natural sciences.

Let $X_g$ (=0, 1, ..., $m_g$) denote a graded item score to item $g$ and $x_g$ be its realization. Note that unlike dichotomous responses, the values of $m_g$'s can be different for separate items; the feature that makes the graded response model more widely applicable to different types of data. The dichotomous response model is included in the graded response model as its special case where $m_g = 1$ for all items.

The *operating characteristic*, $P_{x_g}(\theta)$, of the graded item score $x_g$ is defined by

$$P_{x_g}(\theta) \equiv \text{prob.}[X_g = x_g \mid \theta], \tag{5}$$

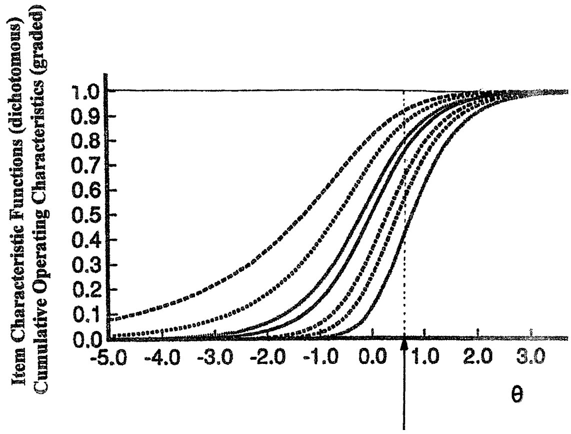


FIGURE 1.
Illustration that the item characteristic functions of seven dichotomous items following the logistic positive exponent family of models with $a_g = 1$, $b_g = 0$, and $\xi_g = 0.3$, 0.5, 0.8, 1.0, 1.5, 2.0, 3.0, respectively, can be the cumulative operating characteristics of a single graded response item with $m_g = 7$.

and the general graded response model is represented by

$$P_{x_g}(\theta) = \left[\prod_{u \le x_g} M_u(\theta)\right]\left[1 - M_{(x_g+1)}(\theta)\right], \tag{6}$$

where $M_{x_g}(\theta)$ is called the *processing function*, which is strictly increasing in $\theta$ except

$$M_{x_g}(\theta) \begin{cases} =1 & \text{for } x_g = 0, \\ =0 & \text{for } x_g = m_g + 1. \end{cases} \tag{7}$$

Let $P^*_{x_g}(\theta)$ be the *cumulative operating characteristic* of the graded item score, $X_g = x_g$, defined by

$$P^*_{x_g}(\theta) \equiv \text{prob.}[X_g \ge x_g \mid \theta] = \prod_{u \le x_g} M_u(\theta). \tag{8}$$

It is obvious from (7) and (8) that

$$P^*_{x_g}(\theta) \begin{cases} =1 & \text{for } x_g = 0, \\ =0 & \text{for } x_g = m_g + 1. \end{cases} \tag{9}$$

From (5) through (9), the operating characteristic $P_{x_g}(\theta)$ can be written as

$$P_{x_g}(\theta) = P^*_{x_g}(\theta) - P^*_{(x_g+1)}(\theta) \tag{10}$$

for $x_g = 0, 1, 2, \ldots, m_g$.

Note that each of the $P^*_{x_g}(\theta)$'s for $x_g = 1, 2, \ldots, m_g$ can be considered as the ICF when graded scores are changed to dichotomous scores using one of the $m_g$ borderlines of adjacent two graded categories as the cutting point. This is exemplified by a usual practice of changing letter grades A, B, C, D, and F into P (pass) and F (fail) setting the borderline between C and D. Thus, seven ICFs in Fig. 1 can also be considered as the cumulative operating characteristics $P^*_{x_g}(\theta)$'s for $x_g = 1, 2, 3, 4, 5, 6, 7$ of a single graded response item, and from (10), it can be seen that $P_{x_g}(\theta)$ is represented by a dotted line segmented by two adjacent $P^*_{x_g}(\theta)$'s, given $\theta$, in the figure.

The operating characteristic $P_{x_g}(\theta)$ is the sole basis of the graded response model, and added by the assumption of *local independence* (Lord and Novick, 1968, Chap. 16), all other important functions such as various types of information functions, the basic function for each of the $m_g + 1$ grades, and the bias function of the maximum likelihood estimate of $\theta$ (MLE bias function) (Samejima, 1993) of a test are derived from $P_{x_g}(\theta)$ (cf. Samejima, 1993). In other words, (5) has an equivalent importance for the general graded response model as the item characteristic function defined by (1) does for the general dichotomous response model. It can be seen that (1) is a special case of (5), which is obtainable by replacing the graded item score $X_g$ by the binary item score $U_g$, and its realization $x_g$ by a specific value of binary item scores, 1.

It is also noted from (6), (7), (8), (9), and (10) that any specific mathematical model that belongs to the general graded response model can be represented, alternately, either by $M_{x_g}(\theta)$ or $P^*_{x_g}(\theta)$.

### 2.2. *Homogeneous and Heterogeneous Cases*

It is noted that when $m_g > 1$, the $m_g + 1$ graded response categories can be redichotomized by choosing one of the $m_g$ borderlines between any pair of adjacent graded response categories into two binary categories, as is done when the letter grades A, B, and C are categorized into "Pass" and D and F into "Fail." When the borderline was set between the categories $(x_g - 1)$ and $x_g$, the cumulative operating characteristic $P^*_{x_g}(\theta)$, defined by the right-hand side of (8), equals the ICF defined by (1). If these $m_g$ ICFs are identical except for the positions alongside the $\theta$ dimension, the model is said to belong to the *homogeneous case*, and otherwise to the *heterogeneous case*.

In the homogeneous case, a model can be represented by

$$P^*_{x_g}(\theta) = \int_{-\infty}^{a_g(\theta - b_{x_g})} \psi(t)\,dt, \tag{11}$$

where $\psi(\cdot)$ denotes a four times differentiable density function with $\lim_{t\to-\infty}\psi(t) = 0$, and the *item response parameter* $b_{x_g}$ satisfies

$$-\infty = b_0 < b_1 < b_2 < \cdots < b_{m_g} < b_{m_g+1} = \infty. \tag{12}$$

Models in the homogeneous case imply that each and every graded response boundary has the same discrimination power (Lord and Novick, 1968), the principle that may fit categorical judgment data. If, for instance, in (11), $\psi(t)$ is replaced by the standard normal density function, it will provide the cumulative operating characteristic in the normal ogive model for graded responses (Samejima, 1969, 1972, 1997, 2004). Another widely used model in the homogeneous case is the logistic model (Samejima, 1969, 1972) expanded from the same model for dichotomous responses proposed by Birnbaum (1968).

Several examples of models which belong to the heterogeneous case can be seen in those graded response models modified from Bock's (1972) nominal response model, and the acceleration model (Samejima, 1995b). The graded response model that is proposed in this paper, i.e., LPEFG, also belongs to the heterogeneous case.

## 3. The LPEF Graded Response Model

### 3.1. Model Specification

In LPEFG, defining $\xi_{-1} = 0$, the processing function is given by

$$M_{x_g}(\theta) = \left[\Psi_g(\theta)\right]^{\xi_{x_g} - \xi_{x_g - 1}} \tag{13}$$

for $x_g = 0, 1, 2, \ldots, m_g, m_g + 1$, where $\Psi_g(\theta)$ is the logistic function defined by (3), and $\xi_{x_g}$ denotes the item response parameter, called *acceleration parameter*, satisfying the relationships:

$$0 = \xi_0 < \xi_1 < \cdots < \xi_{m_g} < \xi_{m_g+1} = \infty. \tag{14}$$

The partial derivative of $M_{x_g}(\theta)$ with respect to $\theta$ is obtained from (13), and because of (14), it can be written as

$$\frac{\partial}{\partial\theta} M_{x_g}(\theta) = (\xi_{x_g} - \xi_{x_g-1})\left[\Psi_g(\theta)\right]^{\xi_{x_g} - \xi_{x_g-1} - 1} Da_g \Psi_g(\theta)\left[1 - \Psi_g(\theta)\right] > 0. \tag{15}$$

It is obvious that (13) and (14) satisfy (7), and from (13) and (15), it can be seen that the $M_{x_g}(\theta)$'s for $x_g = 1, 2, \ldots, m_g$ are strictly increasing functions of $\theta$ with zero and unity as their lower and upper asymptotes.

It is interesting to recall that in the logistic model of the homogeneous case for graded responses, unlike in the normal ogive model, the lower asymptotes of the processing functions are positive values, not zero, for $x_g = 1, 2, \ldots, m_g$, respectively (cf. Samejima, 1972, p. 43). This asymptote value is larger, if the difficulty parameter $b_{x_g}$ is closer to that of the preceding item score, $b_{x_g-1}$.

From (8) and (13), $P^*_{x_g}(\theta)$ can be written as

$$P^*_{x_g}(\theta) = \left[\Psi_g(\theta)\right]^{\xi_{x_g}}, \tag{16}$$

for $x_g = 0, 1, \ldots, m_g$. Thus, it is obvious that $P^*_{x_g}(\theta)$ strictly increases in $\theta$ for $x_g = 1, 2, \ldots, m_g$, and for all $\theta$ equals 1 and 0 for $x_g = 0$ and $x_g = m_g + 1$, respectively, which satisfy (9), and

$$P^*_{x_g}(\theta) > P^*_{(x_g+1)}(\theta) \quad \text{for } x_g = 0, 1, \ldots, m_g. \tag{17}$$

Figures 2 and 3 present examples of the processing functions $M_{x_g}(\theta)$ and the cumulative operating characteristics $P^*_{x_g}(\theta)$. In these examples, $m_g = 5$, $a_g = 1$ and $b_g = 0$, and $\xi_{x_g} = 0.3, 0.5, 0.8, 1.5, 3.0$ for $x_g = 1, 2, 3, 4, 5$, respectively. It is obvious from Fig. 3 that LPEFG belongs to the heterogeneous case for these $P^*_{x_g}(\theta)$'s are not identical in shape.

Note, however, that the processing functions, $M_{x_g}(\theta)$'s depend on the difference between the two adjacent acceleration parameters, and this means that there are no restrictions in ordering the positions of their curves for $x_g = 1, 2, \ldots, m_g$ in accordance with the item score, $x_g$, as is the case with the cumulative operating characteristics, $P^*_{x_g}(\theta)$'s.

Substituting (16) into (10), we obtain the operating characteristics of the graded response $x_g$

$$P_{x_g}(\theta) = \left[\Psi_g(\theta)\right]^{\xi_{x_g}} - \left[\Psi_g(\theta)\right]^{\xi_{x_g+1}} > 0 \quad \text{for } x_g = 0, 1, \ldots, m_g. \tag{18}$$

From (18), the first partial derivative of $P_{x_g}(\theta)$ with respect to $\theta$ is given by

$$\frac{\partial}{\partial\theta} P_{x_g}(\theta) = Da_g\left[1 - \Psi_g(\theta)\right]\left[\xi_{x_g}\Psi_g(\theta)^{\xi_{x_g}} - \xi_{x_g+1}\Psi_g(\theta)^{\xi_{x_g+1}}\right]$$
$$\text{for } x_g = 1, 2, \ldots, m_g - 1. \tag{19}$$

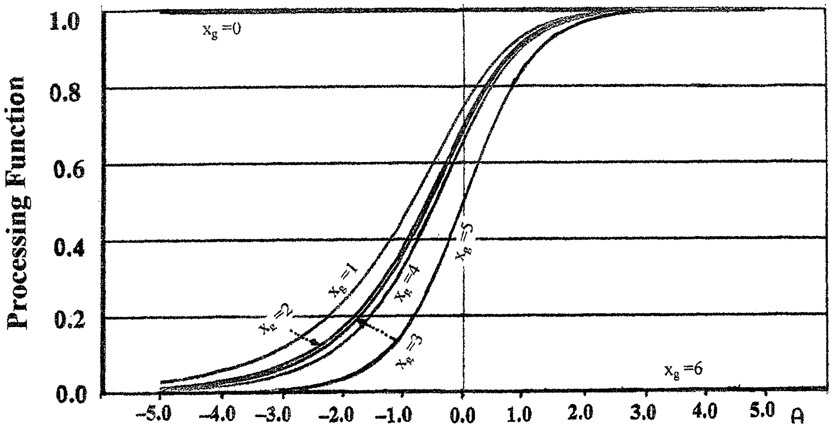


FIGURE 2.
Example of a set of seven processing functions of a graded response item following the LPEFG with $m_g = 5$, $a_g = 1$, $b_g = 0$, and $\xi_{x_g} = 0.3, 0.5, 0.8, 1.5, 3.0$, for $x_g = 0, 1, 2, 3, 4, 5, 6$, respectively.

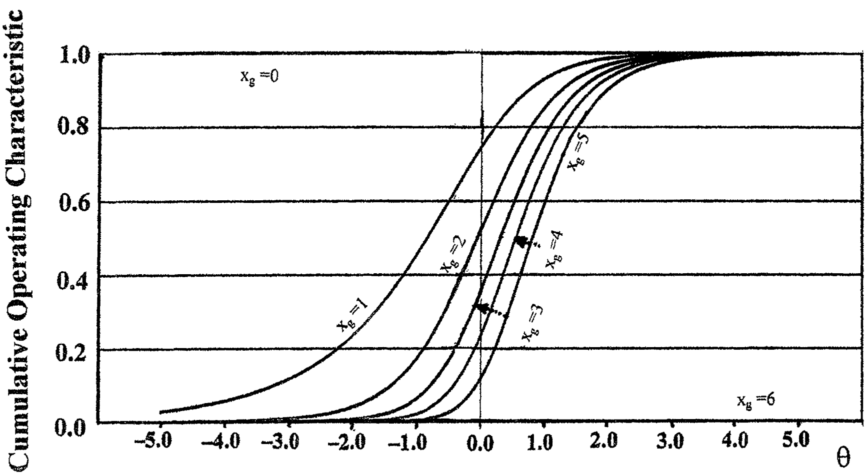


FIGURE 3.
The set of seven cumulative operating characteristics of the same graded response item for $x_g = 0, 1, 2, 3, 4, 5, 6$, respectively.

Note that because of (14) the first term of the last factor in the right-hand side of (19) disappears when $x_g = 0$, and when $x_g = m_g$ its second term disappears to provide

$$\frac{\partial}{\partial\theta} P_{x_g}(\theta) \begin{cases} = -Da_g[1-\Psi_g(\theta)]\xi_{x_g+1}[\Psi_g(\theta)]^{\xi_{x_g+1}} & \text{for } x_g = 0, \\ = Da_g[1-\Psi_g(\theta)]\xi_{x_g}[\Psi_g(\theta)]^{\xi_{x_g}} & \text{for } x_g = m_g. \end{cases} \tag{20}$$

From (18), (19), and (20) it can be seen that the $P_{x_g}(\theta)$ is strictly decreasing in $\theta$ with a terminal maximum at $\theta = -\infty$ for $x_g = 0$, strictly increasing in $\theta$ with a terminal maximum at $\theta = \infty$ for $x_g = m_g$, and it is unimodal for each of the other values of $x_g$. The local modal points of these $(m_g - 1)$ $P_{x_g}(\theta)$'s are the values of $\theta$ to set (19) equal to zero, which is equivalent to the

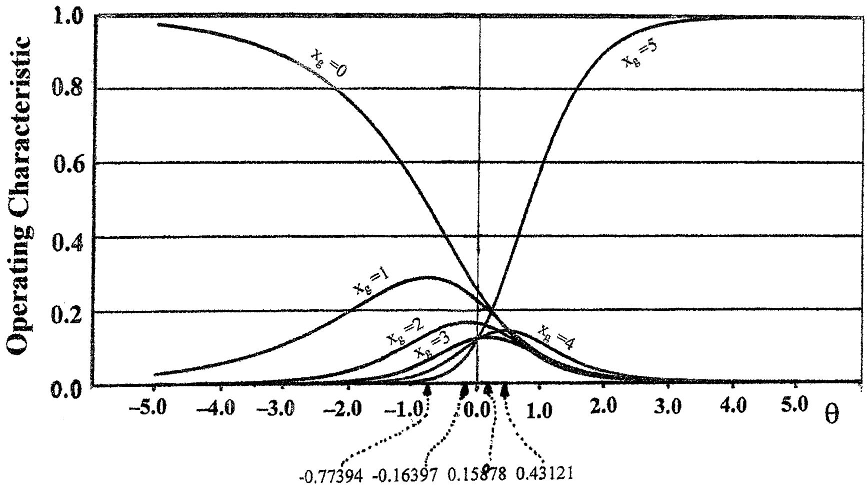


FIGURE 4.
The set of six operating characteristics of the same graded response item for $x_g = 0, 1, 2, 3, 4, 5$, respectively.

solutions of

$$\frac{\xi_{x_g+1}}{\xi_{x_g}} = \left\{1 + \exp\left[-Da_g(\theta - b_g)\right]\right\}^{\xi_{x_g+1}-\xi_{x_g}}. \tag{21}$$

This modal point, $\theta_{\text{mod}}$, can be obtained from (21) and is given by

$$\theta_{\text{mod}} = b_g - (Da_g)^{-1} \log\left[\sqrt[\xi_{x_g+1}-\xi_{x_g}]{\frac{\xi_{x_g+1}}{\xi_{x_g}}} - 1\right]. \tag{22}$$

Figure 4 presents the six operating characteristics of the same example item used in Figs. 2 and 3. These modal points computed from (22) turned out to be: −0.77394, −0.16397, 0.15878, and 0.43121 for $x_g = 1, 2, 3, 4$, respectively.

### 3.2. *Satisfaction of the Unique Maximum Condition*

This satisfaction is considered as one of the desirable features of specific graded response models, together with additivity, ordered modal points of $P_{x_g}(\theta)$'s in accordance with the graded item scores $x_g$'s, etc. (cf. Samejima, 1996).

The *basic functions*, $A_{x_g}(\theta)$, (Samejima, 1969, 1997, 2004) in the general graded response model is defined by

$$A_{x_g}(\theta) \equiv \frac{\partial}{\partial\theta} \log P_{x_g}(\theta) = \frac{\frac{\partial}{\partial\theta} P_{x_g}(\theta)}{P_{x_g}(\theta)}. \tag{23}$$

A sufficient condition that a unique maximum exists in the likelihood function of each and every possible response pattern when all the items follow a specific model is given by

1. $A_{x_g}(\theta)$ defined by (23) is strictly decreasing in $\theta$, and
2. either
   (a) its upper asymptote is positive and lower asymptote is nonpositive, or
   (b) its upper asymptote is nonnegative and lower asymptote is negative

(cf. Samejima, 1969, 1972, 1997, 2004). For brevity, this condition is called the *unique maximum condition*. Satisfaction of this condition is a desirable feature of a model.

In LPEFG, the basic function can be written from (18) and (19) as

$$A_{x_g}(\theta) = \frac{Da_g(1-\Psi_g(\theta))[\xi_{x_g}\Psi_g(\theta)^{\xi_{x_g}} - \xi_{x_g+1}\Psi_g(\theta)^{\xi_{x_g+1}}]}{[\Psi_g(\theta)]^{\xi_{x_g}} - [\Psi_g(\theta)]^{\xi_{x_g+1}}} \tag{24}$$

for $x_g = 1, 2, \ldots, m_g - 1$.

Similar simplifications as pointed out for (19) are possible in the numerator of (24) when $x_g = 0$ and $x_g = m_g$, respectively. It is also noted that in the denominator of (24), when $x_g = 0$, the first term becomes unity, and when $x_g = m_g$, the second term disappears. Thus, for these two cases, we can write

$$A_{x_g}(\theta) \begin{cases} = \frac{-Da_g\ [1-\Psi_g(\theta)]\ \xi_{x_g+1}\ [\Psi_g(\theta)]^{\xi_{x_g+1}}}{1-[\Psi_g(\theta)]^{\xi_{x_g+1}}} & \text{for } x_g = 0, \\ = Da_g\ [1-\Psi_g(\theta)]\xi_{x_g} & \text{for } x_g = m_g, \end{cases} \tag{25}$$

respectively.

Using l'Hospital's rule, it can be seen that the basic functions provided by (24) and (25) are strictly decreasing in $\theta$ for all $x_g$'s, with the two asymptotes

$$\lim_{\theta\to-\infty} A_{x_g}(\theta) \begin{cases} = 0, & x_g = 0, \\ = Da_g\xi_{x_g} > 0, & x_g = 1, 2, \ldots, m_g, \end{cases} \tag{26}$$

and

$$\lim_{\theta\to\infty} A_{x_g}(\theta) \begin{cases} = -Da_g < 0, & x_g = 0, 1, 2, \ldots, m_g - 1, \\ = 0, & x_g = m_g. \end{cases} \tag{27}$$

LPEFG satisfies the unique maximum condition, therefore, indicating that for each and every response pattern a unique maximum likelihood estimate of $\theta$ exists. These MLEs assume finite values for all possible response patterns except for the two extreme cases in which all item scores are zero and all item scores are the highest scores, $m_g$'s, respectively, where the MLEs are negative and positive infinities.

Figure 5 presents the six basic functions of the same example of graded item responses shown in Figs. 2–4, with the values of upper and lower asymptotes shown at both ends of each curve, which were calculated through (26) and (27).

It is noted that these curves never cross each other, and the values of $\theta$ at which these curves touch the $\theta$ axis are ordered in accordance with the values of $x_g$. These values are: $-\infty$, $-0.77394$, $-0.16397$, $0.15878$, $0.43121$, and $\infty$ for $x_g = 0, 1, 2, 3, 4, 5$, respectively. They are the modal points of the operating characteristics of these graded scores, shown in Fig. 4.

### 3.3. *Item Response Information Function*

For the general graded response model, the *item response information function* $I_{x_g}(\theta)$ is defined by

$$I_{x_g}(\theta) \equiv -\frac{\partial^2}{\partial\theta^2}\log P_{x_g}(\theta) = \left[\frac{\frac{\partial}{\partial\theta}P_{x_g}(\theta)}{P_{x_g}(\theta)}\right]^2 - \frac{\frac{\partial^2}{\partial\theta^2}P_{x_g}(\theta)}{P_{x_g}(\theta)}, \tag{28}$$

for $x_g = 0, 1, 2, \ldots, m_g$ (cf. Samejima, 1973b, 1997, 2004), where $\frac{\partial^2}{\partial\theta^2}$ indicates the second partial derivative of the function with respect to $\theta$. The *item information function* is given as the

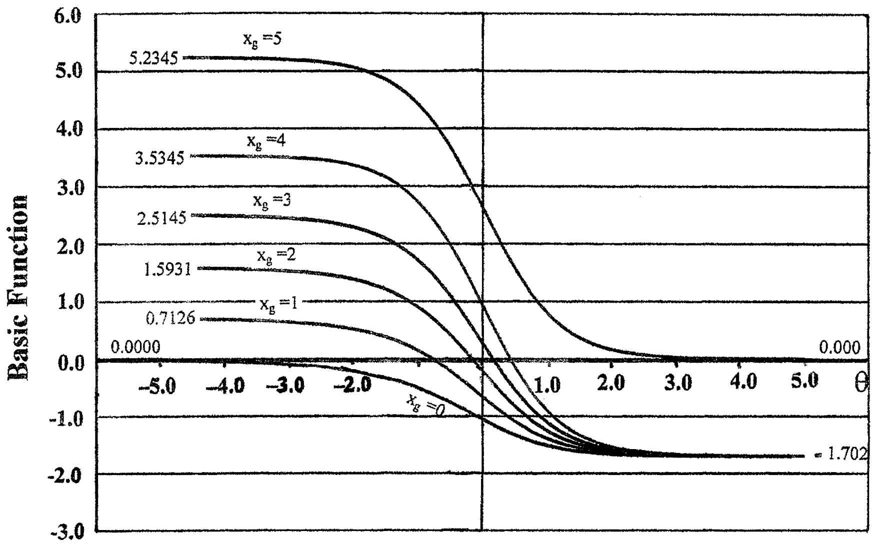


FIGURE 5.
The set of six basic functions of the same graded response item for $x_g = 0, 1, 2, 3, 4, 5$, respectively.

conditional expectation of the item response information function, given $\theta$, which can be written as

$$I_g(\theta) \equiv E\big[I_{x_g}(\theta) \mid \theta\big] = \sum_{x_g} I_{x_g}(\theta) P_{x_g}(\theta) = \sum_{x_g} \frac{[\frac{\partial}{\partial\theta} P_{x_g}(\theta)]^2}{P_{x_g}(\theta)}. \tag{29}$$

Note that (29) includes Birnbaum's (1968) item information function for the dichotomous item, which is based on somewhat different rationale as a special case.

Samejima (1973b, 1997, 2004) also proposed the *response pattern information function*, $I_v(\theta)$, for graded responses that can be written as

$$I_v(\theta) \equiv -\frac{\partial^2}{\partial\theta^2} \log P_v(\theta) = -\sum_{x_g \in v} \frac{\partial^2}{\partial\theta^2} \log P_{x_g}(\theta) = \sum_{x_g \in v} I_{x_g}(\theta), \tag{30}$$

where $P_v(\theta)$ denotes the operating characteristic of the response pattern $V = v$ provided by

$$P_v(\theta) \equiv \text{prob.}[V = v \mid \theta] = \prod_{x_g \in v} P_{x_g}(\theta). \tag{31}$$

The test information function, $I(\theta)$, is defined as the conditional expectation of $I_v(\theta)$, given $\theta$, which from (29), (30), and (31) can be written as

$$I(\theta) \equiv E\big[I_v(\theta) \mid \theta\big] = \sum_v I_v(\theta) P_v(\theta) = -\sum_v \left[\frac{\partial^2}{\partial\theta^2} \log P_v(\theta)\right] P_v(\theta) = \sum_{g=1}^{n} I_g(\theta). \tag{32}$$

The outcome of (32) also includes Birnbaum's (1968) test information function for dichotomous responses as a special case.

Figure 6 presents the item response information functions for $x_g = 0, 1, 2, 3, 4, 5$ for the same example shown in Figs. 2, 3, 4, and 5. It is noted that in this example $I_{x_g}(\theta)$'s for $x_g = 4$

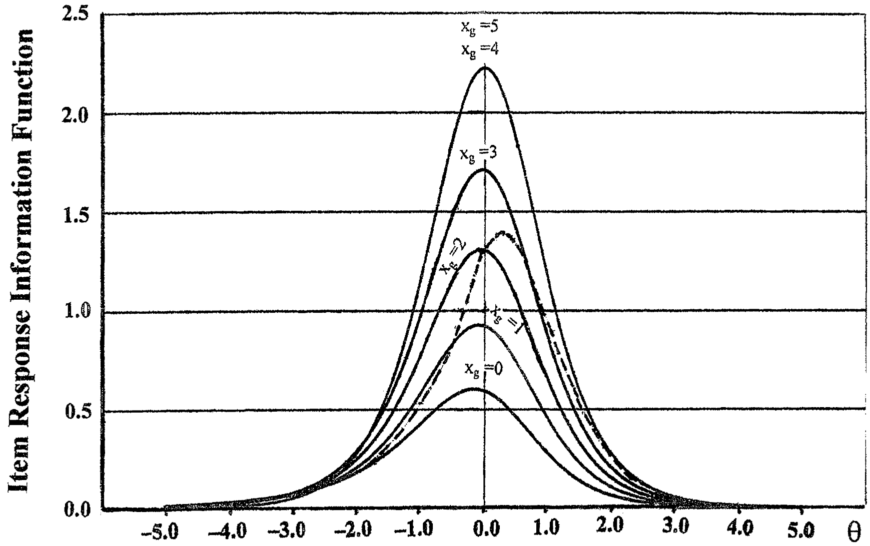


FIGURE 6.
The set of six item response information functions of the same graded response item for $x_g = 0, 1, 2, 3, 4, 5$, respectively, and item information function (*dashed line*).

and 5 are overlapping each other. This comes from the fact that in this example the acceleration parameters for the item scores 4 and 5 happen to be:

$$\xi_5 = 3.07913 = 2.07913 + 1 = \xi_4 + 1,$$

and from (24) and the second line of (25), it can be written:

$$\begin{aligned}
A_5(\theta) - A_4(\theta) &= Da_g\big[1 - \Psi_g(\theta)\big]\left[\xi_5 - \frac{Da_g[1-\Psi_g(\theta)][\xi_4\Psi_g(\theta)^{\xi_4} - \xi_5\Psi_g(\theta)^{\xi_5}]}{\Psi_g(\theta)^{\xi_4} - \Psi_g(\theta)^{\xi_5}}\right] \\
&= Da_g\big[1 - \Psi_g(\theta)\big][1+\xi_4] \\
&\quad - \frac{Da_g[1-\Psi_g(\theta)][\xi_4\Psi_g(\theta)^{\xi_4} - [1+\xi_4]\Psi_g(\theta)^{\xi_4}\Psi_g(\theta)]}{\Psi_g(\theta)^{\xi_4} - \Psi_g(\theta)^{[1+\xi_4]}} \\
&= Da_g\big[1 - \Psi_g(\theta)\big][1+\xi_4] - \frac{Da_g[1-\Psi_g(\theta)]\Psi_g(\theta)^{\xi_4}[\xi_4 - [1+\xi_4]\Psi_g(\theta)]}{\Psi_g(\theta)^{\xi_4}[1-\Psi_g(\theta)]} \\
&= Da_g\big[1 - \Psi_g(\theta)\big][1+\xi_4] - Da_g\big[\xi_4 - [1+\xi_4]\Psi_g(\theta)\big] \\
&= Da_g[(1+\xi_4) - \Psi_g(\theta)(1+\xi_4) - \xi_4 + \Psi_g(\theta)[1+\xi_4]] \\
&= Da_g. \qquad (33)
\end{aligned}$$

The vertical distance of $A_5(\theta)$ from $A_4(\theta)$ equals a constant, $Da_g$, for all $\theta$, therefore, can be visualized in Fig. 5. Since from (23) and (28), it can be written as

$$I_{x_g}(\theta) = -\frac{\partial}{\partial\theta}A_{x_g}(\theta), \qquad (34)$$

the item response information functions for $x_g = 4$ and $x_g = 5$ are also identical.

It is noted that in the example given here, the maximal values of the item response information functions are ordered in accordance with the item score. Since the lower asymptote of the basic function is zero for $x_g = m_g$ while it is uniformly $-Da_g$ otherwise (cf. (27), however, this order does not necessarily exist). If, for instance, $\xi_5$ is just slightly greater than $\xi_4$, the curve for $x_g = 5$ may become flatter, and the maximal value of the item response information response function will assume a less value than the corresponding value for $x_4$.

In the same figure, the item information function is shown by a dashed line.

### 3.4. *Additivity*

Samejima (2004) categorizes those models in the family of general graded response model into:

1. models that can be naturally expanded to those for continuous responses, and
2. those which are discrete in nature.

It can be shown that LPEFG is naturally expanded to a model for continuous responses, that is, it belongs to category 1. The continuous response model thus obtained will be discussed in a separate paper, however.

Most models in the first category also have *additivity* (Samejima, 1996), in the sense that:

1. If two or more adjacent graded response categories (e.g., A, B, and C) are combined into one new category (e.g., Pass), the operating characteristic of the resulting new category, which is the sum total of the original separate operating characteristics, also belongs to the original model, and
2. If an original graded response category (e.g., "agree") is divided into two or more graded categories (e.g., "strongly agree" and "moderately agree") the operating characteristics of these new categories can be found in the original mathematical model. (Note that in the acceleration model, however, additivity 1 does not rigorously hold, but at least it approximately does in most practical situations (cf. Samejima, 1995b).)

LPEFG has such a desirable characteristic described above, that is, additivity.

### 3.5. *The Relationship with the Acceleration Model*

Let $\xi^*_{x_g}$ be defined as

$$\xi^*_{x_g} = \xi_{x_g} - \xi_{x_g-1}. \tag{35}$$

From inequality (14), it is obvious that $\xi^*_{x_g} > 0$ for $x_g = 1, 2, \ldots, m_g$, and $\xi^*_{x_g} = 0$ for $x_g = 0$ and $\xi^*_{x_g} = \infty$ for $x_g = m_g + 1$. Let $\Psi_{x_g}(\theta)$ denote

$$\Psi_{x_g}(\theta) = \frac{1}{1 + \exp[-D\alpha_{x_g}(\theta - \beta_{x_g})]}, \tag{36}$$

where $\alpha_{x_g}$ (>0) and $\beta_{x_g}$ are item response discrimination and location parameters, respectively.

It is noted that if in (13) $\Psi_g(\theta)$ is replaced by $\Psi_{x_g}(\theta)$ defined by (36) with the item response parameter $\alpha_{x_g}$ and rewriting $\xi_{x_g} - \xi_{x_g-1}$ as $\xi^*_{x_g}$, it will become the processing function of a special case of the acceleration model, with $\xi^*_{x_g}$ as *the step acceleration parameter* (cf. Samejima, 1995b).

Thus, it is legitimated to consider that LPEFG belongs to the family of acceleration models. Aside from the differences in the objectives of these models, however, LPEFG has an advantage over more general acceleration models in the sense that as was pointed out earlier, it has perfect additivity and orderliness of the modal points of the operating characteristics, while for more general acceleration models only their *robustnesses* are assured (cf. Samejima, 1996).

## 4. Limitation of Curve Fittings in Model Validation

The reader might expect the author to use a set of simulated data for demonstrating the fitness of LPEFG to a certain type of data. This can be done by creating a hypothetical set of data that follows LPEFG using the Monte Carlo method, estimating the cumulative operating characteristic, $P^*_{x_g}(\theta)$, for each graded response $x_g$ by one of the nonparametric estimation methods such as Levine's (1984), Ramsay's (1991), and Samejima's (1998, 2001). The processing functions $M_{x_g}(\theta)$'s can be computed from the $P^*_{x_g}(\theta)$'s, and the operating characteristic $P_{x_g}(\theta)$'s, and then find out how the outcomes fit LPEFG with a certain set of parameter values.

A problem of graded response models is that such a curve fitting may not validate the model, because two or more mathematical models that are substantially different in principle may provide a very similar set of curves. Samejima (1996, 1997) illustrated this fact using the acceleration model (1995b) and Bock's nominal response model (1972).

Figure 7 is the combination of Figs. 2 (upper graph) and 1 (lower graph) in Samejima (1996) arranged in a reversed order because the parameter values of the latter were calculated from those

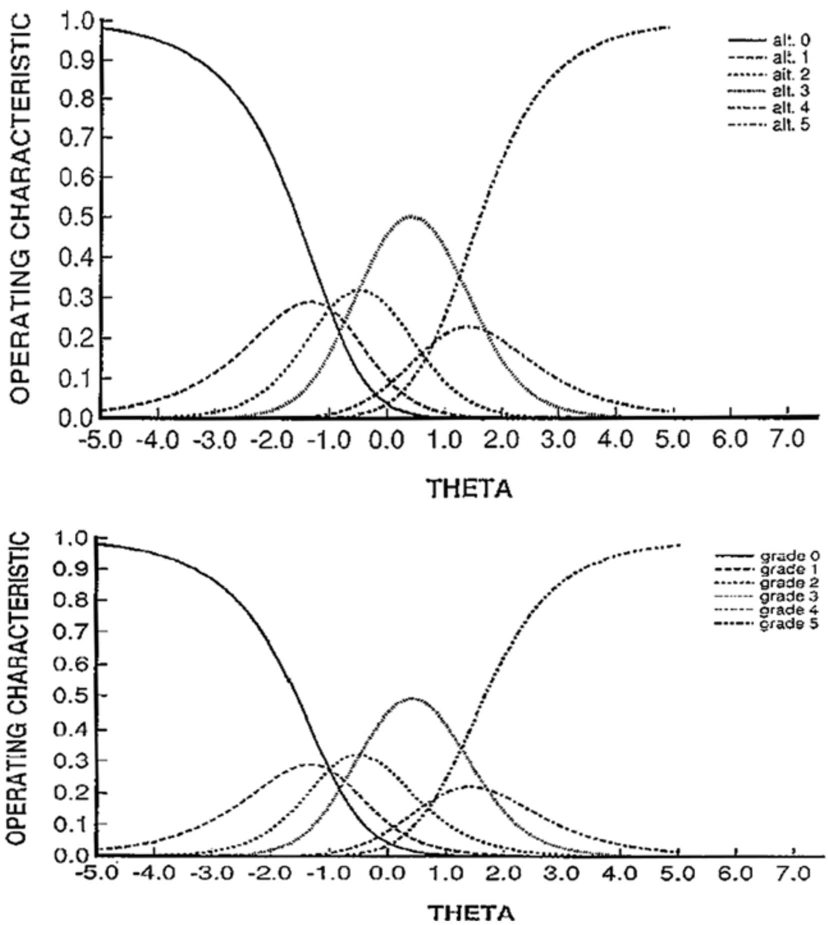


FIGURE 7.
A set of operating characteristics of six discrete responses following Bock's nominal model, with $\alpha_{x_g} = 1, 2, 3, 4, 5, 6$ and $\beta_{x_g} = 1.0, 2.0, 3.0, 3.5, 1.8, 1.0$, for $x_g = 0, 1, 2, 3, 4, 5, 6$, respectively (*upper graph*), and those of six steps in the acceleration model (*lower graph*), approximating the processing functions in Bock's model, whose parameter values turned up to be: $\alpha_{x_g} = 1.35719, 1.03009, 0.87689, 1.07899, 0.58824$, $\beta_{x_g} = -0.95250, -0.77484, 0.04352, 1.32708, 0.80000$, and $\xi_{x_g} = 0.42671, 0.52069, 0.53958, 0.62196, 1.00000$, for $x_g$'s, 1 through 5.

of the former to make each of the resulting operating characteristics very close to those in the upper graph.

A close observation of the two graphs in Fig. 7 and straightforward logic are convincing enough to say that any set of data that fit the acceleration model (lower graph) also fit the nominal response model (upper graph), which means failure in validating either of the two models when curve fittings alone are used.

Actually, noting similarities in the shapes of the processing functions in Bock's nominal response model and those in the acceleration model, the parameter values of the acceleration model were calculated from those in the nominal response model by fitting each curve at $M_{x_g}(\theta) = 0.1, 0.5, 0.9$. Changing these three values of $M_{x_g}(\theta)$ in several different ways, say, to $M_{x_g}(\theta) = 0.3, 0.6, 0.7$, etc., it was confirmed that the outcomes were almost identical (cf. Samejima, 1995a). The parameter values in the acceleration model illustrated here were obtained by using the latter set of three values, i.e., 0.3, 0.6, 0.7, and have as many as five digits after the decimal point. Details of this method will be presented and discussed in a separate paper.

## 5. Substantive Modeling: Examples

From the standpoint of substantive mathematical modeling, one type of psychological phenomena that LPEFG may fit well is *problem solving*, that includes proving a mathematical theorem. In proving a theorem, various levels of problem complexity are conceivable.

It is said, Pythagoras' theorem has as many as 367 different proofs. In most problem solving, it is usual that there are multiple e.g. cognitive process sequences that lead to the correct answer. In real research, therefore, it may be advisable to conduct good pilot studies, and then from their outcomes pick up several typical cognitive sequences that the majority of examinees in the pilot studies followed, and then analyze them separately and then eventually combined (cf. Samejima, 1983).

In this section, a couple of relatively simple examples to which LPEFG is likely to fit will be illustrated.

### 5.1. *Example 1*

The first item taken as an example here, item $D_{2.2}$, is a reasoning test item, taken from the LIS Measurement Scale of Nonverbal Reasoning Ability (Indow and Samejima, 1962). The test is a paper and pencil test, and each item was originally scored dichotomously, i.e., either correct ($u_g = 1$) or incorrect ($u_g = 0$). The test was administered to 883 junior high school students (584 eighth graders and 299 ninth graders) in Tokyo, Japan. Based on this group of examinees, the scale was constructed using the normal ogive model for dichotomous responses. Out of 883 eighth graders, 202 examinees, or 22.88% answered the item correctly and the other 681 did not.

Item $D_{2.2}$, is an *addition*, such that

$$\begin{array}{r} \text{Y A M A} \\ +\text{U M I} \\ \hline \text{K A W A} \end{array}$$

that consists of seven alphabetical letters A, I, K, M, U, W, and Y, each of which represents one of the positive integers, 0, 3, 4, 6, 7, 8, and 9, and the examinee is to find out which alphabetical letter represents which number.

It is noted that the right most column of the above addition is A + I = A, and because the column indicates these numbers are of the lowest digit, it can be reasoned that I = 0. Thus, this first process provides:

$$\begin{array}{r} \text{Y A M A} \\ +\ \text{U M 0} \\ \hline \text{K A W A} \end{array}$$

The graded item score given to those who answered I = 0 correctly, but could not go further, will be $x_g = 1$.

After this first process has been successfully followed, the six unused letters are A, K, M, U, W, and Y, and six unused numbers are 3, 4, 6, 7, 8, and 9, respectively. Then it is noted that in the third column from the right there is A + U = A. Because 0 has already been matched to I, U cannot be zero. This is also confirmed by the fact that in the fourth column Y and K must represent two different numbers. Thus, the only logical answer is U = 9. This second reasoning process provides:

$$\begin{array}{r} \text{Y A M A} \\ +\ \text{9 M 0} \\ \hline \text{K A W A} \end{array}$$

The graded score given to those who answered U = 9 correctly in addition to I = 0, but could not go further will be $x_g = 2$.

When the author checked the examinees' writings years ago, it was noted that out of the 681 students who could not go through all the processes, as many as 187 had obviously reasoned up to I = 0 and U = 9 correctly, and then could not go further. This fact is a good reason to use the item as a graded response item, instead of a dichotomous response item.

Today, unlike early 1960s, it is easy to trace an examinee's performance in the computerized testing environment instead of paper and pencil testing, and identify each examinee's level of performance.

It is obvious from (7) and (8) that when $x_g = 1$, the processing function $M_{x_g}(\theta)$ equals the cumulative operating characteristic $P^*_{x_g}(\theta)$, and the latter will also be the ICF when the item score is redichotomized by setting the borderline of "pass" and "fail" between $x_g = 1$ and $x_g = 2$. Thus, using one of the nonparametric estimation methods such as Levine's (1984), Ramsay's (1991), and Samejima's (1998, 2001), the processing function for $x_g = 1$ can be estimated. Then likewise the cumulative operating characteristic, $P^*_{x_g}(\theta)$, for $x_g = 2$ can be estimated, setting the borderline of redichotomization between $x_g = 2$ and $x_g = 3$, and from (8) the conditional ratio of this estimated $P^*_{x_g}(\theta)$ for $x_g = 2$, given $\theta$, to the one for $x_g = 1$ will provide the estimate of $M_{x_g}(\theta)$ for $x_g = 2$.

If most examinees who have found out I = 0 can also find out U = 9, then the processing function $M_{x_g}(\theta)$ for $x_g = 2$ will be strictly increasing in $\theta$ but very close to the constant function $M_{x_g}(\theta) = 1$. Note that unlike $P^*_{x_g}(\theta)$'s, there is no ordering relationship for $M_{x_g}(\theta)$ in accordance with the graded item score $x_g$, and yet $P^*_{x_g}(\theta)$'s are always ordered, as is obvious from (8), because $0 < M_{x_g}(\theta) < 1$ for $x_g = 1, 2, \ldots, m_g$, for all $\theta$.

The remaining five unused letters are A, K, M, W, and Y, and the five unused numbers are 3, 4, 6, 7, and 8, respectively. It is noted that K = Y + 1, because the addition is of two numbers only. From the list of unused integers, it can be seen that either Y = 3 and K = 4, Y = 6 and K = 7, or Y = 7 and K = 8. Also, M must be greater than 5 and W = 2M − 10, so either M = 7 and W = 4 or M = 8 and W = 6, because M = 6 is impossible because 2 is not included in last subset of unused numbers.

If Y = 6 or 7, however, neither of the above two possible pairs of M and W can be realized. Thus, Y must be 3 and K must be 4, that is,

$$\begin{array}{r} 3\ \text{A M A} \\ +\,9\ \text{M}\ 0 \\ \hline 4\ \text{A W A} \end{array}$$

This outcome receives $x_g = 3$.

Then the three unused letters are A, M, and W, and the three unused numbers are 6, 7, and 8, respectively. Thus, it is obvious that M = 8, W = 6, and A = 7, to make the formula

$$\begin{array}{r} 3\ 787 \\ +\,980 \\ \hline 4\ 767 \end{array}$$

which is the correct answer that deserves $x_g = 4$ (=$m_g$).

We could separate Y = 7 from the above process, and make the highest item score category $x_g = 5$. However, after the process of finding M = 8 and W = 6, no other values but 7 remain that can be assigned to A. Thus, it is very possible that when analyzing the data, we will find this last step a degenerated category, or $M_{x_g}(\theta) = 1.0$ for $x_g = 5$ to provide an identical $P^*_{x_g}(\theta)$ with that of $x_g = 4$. Thus, the highest item score category $x_g = 5$ will be meaningless.

Table 1 presents possible graded scores of item $D_{2.2}$ when it is given as a paper and pencil test item and a computerized test item. It is noted that in the environment of computerized testing not only when matching of a number and a letter is completed, but a reasoning process itself can be scored, as exemplified by $x_g = 3$ and $x_g = 4$. It is also noted that these two processes can be reversed, that is, some examinees might work on (Y, K) first and then (M, W) second, and others might work in a reversed order. It is recommended that the examinee group should be divided into two subgroups in accordance with the two orders, and the resulting estimated processing functions are compared.

Note that even in paper and pencil testing, if detailed instructions are given, and if scorers read the examinees' handwritings intensively as they do in scoring projective tests, then $m_g$ can be 5 just as in computerized testing. Such a scoring will be far more time consuming, however.

TABLE 1.
Item $D_{22}$ and its graded item scores that may be optimal in two different testing environments.

| Paper–pencil testing | Computerized testing | Subprocess |
|---|---|---|
| 0 | 0 | (None) |
| 0 | 1 | J = 0, Left: (3 4 6 7 6 9) |
| 1 | 2 | U = 9, Left: (3 4 6 7 8) |
| 1 | 3 | K = Y · 1,<br>(Y, K) = (3, 4) or (6, 7) or (7, 8) |
| 1 | 4 | W = 2M − 10(M, W) = (7, 4) or (8, 6) |
| 2 | 5 | M = 8<br>W = 6<br>Y = 3<br>K = 4, Left: (7) |
| 2 | 5 | A = 7 |

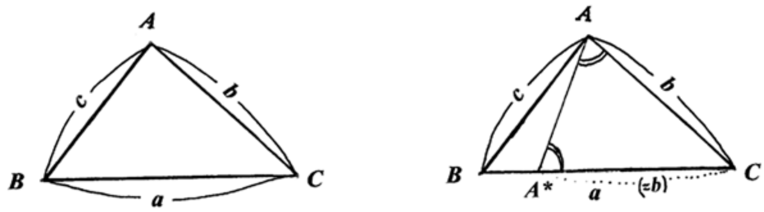


FIGURE 8.
Given triangle (*left*) and the same triangle after used for reasoning processes (*right*).

### 5.2. *Example 2*

The second example is the proof of an elementary geometry theorem. Suppose in the left-hand side triangle in Fig. 8 $a > b > c$. Prove that

$$\text{angle } BAC > \text{angle } ABC > \text{angle } ACB.$$

*Proof:* Here, the proof will be shown for angle $BAC >$ angle $ABC$ only because the other inequality, angle $ABC >$ angle $ACB$ can be worked out in the same way.

(a) The proof starts from taking the same length as $b$ on $a$, starting from $C$, and call the other end of the new line $A^*$. Since $a > b$, $A^*$ must be located on the line $a$, as shown in the right-hand side triangle in Fig. 8.
(b) Combine $A$ and $A^*$ by a straight line.
(c) Thus, angle $BAC =$ angle $A^*AC +$ angle $BAA^*$, and
(d) angle $ABC =$ angle $ABA^*$.
(e) angle $ABA^* =$ angle $AA^*C -$ angle $BAA^*$, since for any triangle the sum of the three angles equals $\pi$.
(f) Because triangle $AA^*C$ is an isosceles triangle, angle $AA^*C =$ angle $A^*AC$.
(g) Thus, angle $ABC =$ angle $A^*AC -$ angle $BAA^*$.
(h) angle $BAC -$ angle $ABC = 2$ angle $BAA^* > 0$.
(i) Therefore, angle $BAC >$ angle $ABC$.

We could score 1 through 9 for these processes. Note that in this proof process (a) is essential, so the processing function for (a) is expected to have a large acceleration parameter. □

## 6. Discussion and Practical Implications

LPEFG, a new graded response model that belongs to the heterogeneous case, has been proposed and its characteristics have been observed. The author has been in search of a graded response model that belongs to the heterogeneous case and can be naturally expanded to a continuous response model, that is comparable to the normal ogive and logistic models for graded responses in the homogeneous case (Samejima, 1973a). The model proposed in the present paper is an answer to the author's search.

It is fairly common that when one tries to solve a mathematical problem, for example, early processes are relatively easy in the sense even people of lower ability levels have high probabilities of finding and completing each process, but on later stages such probabilities will decrease. After following one or more right processes in solving the problem, one might have to bang his/her head on a wall trying to go further without success. LPEFG may be appropriate to use in giving partial credit to such an outcome.

References

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability, Contributed chapters. In F.M. Lord & M.R. Novick (Eds.), *Statistical theories of mental test scores*. Reading, MA: Addison Wesley. Chaps. 17–20.

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

Indow, T., & Samejima, F. (1962). *LIS measurement scale of non-verbal reasoning ability*. Tokyo, Japan: Nippon Bunka Kagakusha (in Japanese).

Levine, M.V. (1984). *An introduction to multilinear formula score theory* (Office of Naval Research Report 84-4, N00014-83-K-0397). Champaign, IL: Model-Based Measurement Laboratory, University of Illinois.

Lord, F.M., & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison Wesley.

Ramsay, J.O. (1991). Kernel smoothing approaches to nonparametric item characteristic estimation. *Psychometrika*, *56*, 611–630.

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

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

Samejima, F. (1973a). Homogeneous case of the continuous response model. *Psychometrika*, *38*, 203–219.

Samejima, F. (1973b). A comment on Birnbaum's three-parameter logistic model in the latent trait theory. *Psychometrika*, *38*, 221–233.

Samejima, F. (1983). *A latent trait model for differential strategies in cognitive processes* (Office of Naval Research Report 83-1, N00014-81-C-0569). Knoxville, TN: University of Tennessee.

Samejima, F. (1993). An approximation for the bias function of the maximum likelihood estimate of a latent variable for the general case where the item, responses are discrete. *Psychometrika*, *58*, 119–138.

Samejima, F. (1995a). *An easy method to switch to the acceleration model from other models for graded responses*. Paper presented at the 1995 annual meeting of the National Council on Measurement in Education on April 21, 1995.

Samejima, F. (1995b). Acceleration model in the heterogeneous case of the general graded response model. *Psychometrika*, *60*, 549–572.

Samejima, F. (1996). Evaluation of mathematical models for ordered polychotomous responses. *Behaviormetrika*, *23*, 17–35.

Samejima, F. (1997). Graded response model. In W.J. Van der Linden & R.K. Hambleton (Eds.), *Handbook of modern item response theory*. New York: Springer. Chap. 5.

Samejima, F. (1998). Efficient nonparametric approaches for estimating the operating characteristics of discrete item responses. *Psychometrika*, *63*, 111–130.

Samejima, F. (2000). Logistic positive exponent family of models: Virtue of asymmetric item characteristic curves. *Psychometrika*, *65*, 319–335.

Samejima, F. (2001). *Nonparametric on-line item calibration* (Law School Admission Council Final Report). Knoxville, TN: Mathematical Psychology Laboratory, University of Tennessee.

Samejima, F. (2004). Graded response model. In *Encyclopedia of social measurement* (pp. 140–153). San Diego: Academic Press.

*Manuscript Received: 3 NOV 2004*
*Final Version Received: 11 MAR 2008*
*Published Online Date: 19 AUG 2008*