DOCUMENT RESUME

ED 413 336 TM 027 571

AUTHOR	Samejima, Fumiko
TITLE	Ability Estimates That Order Individuals with Consistent Philosophies.
PUB DATE	1997-03-28
NOTE	37p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Ability; *Estimation (Mathematics); *Item Response Theory; Philosophy; *Scores; Test Interpretation; Test Items; Validity
IDENTIFIERS	*Order Relations; Ordering Operations

ABSTRACT

Latent trait models introduced the concept of the latent trait, or ability, as distinct from the test score. There is a recent tendency to treat the test score as through it were a substitute for ability, largely because the test score is a convenient way to place individuals in order. F. Samejima (1969) has shown that, in general, the amount of test information decreases if the test score is the estimate of ability. This paper introduces a new family of models that has a high level of substantive validity and inner consistency in ordering individuals. This family is called the logistic positive exponent family (LPEF). The development of these models is traced, and how to define the item characteristic curve in the LPEF family is demonstrated. One of the most important characteristics of the LPEF family is that a point-symmetric (logistic) model is treated as one of the infinitely many models of the family. LPEF provides more appropriate models for human behavior those based on error distributions. (Contains 2 tables, 5 figures, and 12 references.) (SLD)

ABILITY ESTIMATES THAT ORDER INDIVIDUALS WITH CONSISTENT PHILOSOPHIES[1]

Fumiko Samejima

The University of Tennessee

The 1997 Annual AERA Meeting

March 28, Chicago, IL

1 Requests for reprints should be sent to Fumiko Samejima, Department of Psychology, 405 Austin Peay Bldg., University of Tennessee, Knoxville, Tennessee 37996-0900. E-mail: samejima@psych1.psych.utk.edu

I. Objective

Latent trait models introduced the concept of latent trait or *ability*, which is distinct from the test score. There is a recent tendency, however, to treat the test score as if it were an authorized substitute of ability as exemplified by studies on the *monotone likelihood ratio*. One of the reasons for this tendency is that the test score is a convenient measure to order individuals.

Samejima (1969) has shown that, in general, the amount of test information decreases if we use the test score as the estimate of ability. This implies that the test score includes a substantial amount of error as an estimate of ability, with exceptional situations where the test score is a sufficient statistic, as is the case with the Rasch model. The results of data analysis indicate, however, that in most cases Rasch model with a single item parameter does not fit. Accepting this fact, we must say that it is illegitimate to use the test score as the substitute for ability in evaluating models.

The present paper introduces a new family of models that has a high level of substantive validity and inner consistency in ordering individuals. This family is called the logistic positive exponent family (LPEF) (Samejima, 1972, 1997b).

II. Contradiction in Symmetric Item Characteristic Curve

Let θ be the latent trait, or *ability*, which assumes any real number, and $P_g(\theta)$ denote the item characteristic curve (ICC) of item g $(= 1, 2, ..., n)$, or the conditional probability, given θ , with which the individual answers item g correctly. Thus

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

where U_g is a binary item score.

It is noted that the ICC in both the normal ogive model and in the logistic model, which includes Rasch model as its special case, is *point-symmetric*, with $(b_g, 0.5)$ as the point of symmetry. That is,

$$P_g(b_g + \alpha) = 1 - P_g(b_g - \alpha) \ , \tag{2}$$

where b_g is the item difficulty parameter or the value of θ at which $P_g(\theta) = 0.5$, and α is any arbitrary number. For brevity, any ICC that satisfies Eq. (2) will be called *symmetric ICC*.

Insert Table 1 About Here

A characteristic of a symmetric ICC is that it treats both correct and incorrect answers symmetrically, which results in a logical contradiction in ordering examinees on the ability scale. Table 1 shows the 32 possible response patterns arranged in the ascending order of the maximum likelihood estimates (MLE's) of ability based on a hypothetical test of 5 dichotomous items following the normal ogive model, with a common discrimination parameter 1.0 and equally spaced difficulty parameters $-3.0, -1.5, 0.0, 1.5, 3.0$. A close examination of Table 1 discloses that, if we divide the 32 response patterns into two subgroups by cutting the table between items 16 and 17, the response patterns of the second group are compliments of those of the first group arranged in the reversed order. This includes, for example, that:

1. the 5 response patterns, in each of which only one item is answered correctly, are arranged in the order of *difficulty* of the item that is answered correctly, and
2. the 5 response patterns in each of which 4 out of 5 items are answered correctly are arranged in the order of *easiness* of the item that is *not* answered correctly.

These two principles are contradictory to each other, for if we accept the first principle then we should expect that the response pattern which includes the correct answers to the 4 most difficult items should get the highest ability estimate, for example. Nonetheless these results are natural outcomes of symmetric ICC's, and the contradiction is intrinsic. This also implies that the order of the ability estimates is influenced by the number of items.

III. Logistic Positive Exponent Family

In solving the problem g, one must clear each of many sequential subprocesses. It is expected that the conditional probability, given θ, with which tougher and a larger number of sequential subprocesses are cleared becomes less. These differences in the conditional probability are expected to become more pronounced for different levels of ability as a larger number of subprocesses are cleared. For convenience, let us call this aspect of an item *item complexity*, as distinct from item difficulty. Taking this item complexity into account, it will be more appropriate to assume that the conditional distribution of a response tendency, given θ, be skewed, rather than symmetric. This skewness depends on the degree of complexity of the item, and provides an *asymmetric* ICC.

Due credit given to success in solving a complex item may be exemplified by marathon running. It is unlikely that some mediocre marathon runner would make a world record because he/she is "up", but if he/she is "down" he/she can be slower than any fellow runners, or even unable to finish the race. In cases like this, it will be more appropriate to adopt a model that is based on a positively skewed conditional distribution of the item response tendency and the resulting *drop* ratio from the symmetric ICC be affected by both the individual's competency level and the item.

Let $\tilde{P}_g(\theta)$ be a symmetric ICC that is provided by a specific mathematical model. The conditional drop ratio, given θ, may be provided by

$$1 - [\tilde{P}_g(\theta)]^c \tag{3}$$

where $c > 0$, which is strictly decreasing in θ representing the principle: *the higher the ability, the lower the drop ratio.* Then the ICC will become

$$P_g(\theta) = \tilde{P}_g(\theta) - [1 - \{\tilde{P}_g(\theta)\}^c]\,\tilde{P}_g(\theta) = [\tilde{P}_g(\theta)]^{1+c} = [\tilde{P}_g(\theta)]^{\xi_g} , \tag{4}$$

where

$$\xi_g = 1 + c > 1 . \tag{5}$$

Equation (4) represents a positive exponent family (Samejima, 1972), in which ξ_g (>

0) is called the acceleration parameter (Samejima, 1995). Note that when (5) is true any ICC given by (4) assumes less values than $\tilde{P}_g(\theta)$ for the entire range of θ .

If we change the interval of c to

$$-1 < c < 0 ,$$

Equation (3) will assume negative values, and (4) can be rewritten as

$$P_g(\theta) = \tilde{P}_g(\theta) + [\{\tilde{P}_g(\theta)\}^c - 1]\, \tilde{P}_g(\theta) = [\tilde{P}_g(\theta)]^{1+c} = [\tilde{P}_g(\theta)]^{\xi_g} .$$

Thus $[\{\tilde{P}_g(\theta)\}^c - 1]$ can be considered as the conditional *elevation* ratio, which is decreasing in θ , representing the principle, *lower the ability, greater the elevation ratio.* The acceleration parameter in (5) becomes

$$0 < \xi_g < 1 . \tag{6}$$

This second principle leads to penalization of failure in solving an easy item, that is, the reversed philosophy. Note that any ICC given by (4) with ξ_g that satisfies (6) assumes higher values than $\tilde{P}_g(\theta)$ for the entire range of θ .

The size of ξ_g is determined by the subprocesses that are required to clear in solving item g . Let w_{gi} $(i = 0, 1, 2, ..., t_g)$ be the i-th subprocess, and $\xi_{w_{gi}}$ (≥ 0) be the subprocess acceleration parameter, which assumes a high positive value if clearing the subprocess w_{gi} is tough, and vice versa. Since everyone can be at the starting point regardless of his/her ability level and no toughness is involved, it is reasonable to set

$$\xi_{w_{g0}} = 0 ,$$

and (4), with the replacement of ξ_g by $\xi_{w_{g0}}$, becomes unity for all θ . The item acceleration parameter ξ_g can be written as

$$\xi_g = \sum_{i=0}^{t_g} \xi_{w_{gi}} .$$

The ICC of the logistic positive exponent family (LPEF) is defined by replacing $\tilde{P}_g(\theta)$ in (4) by $\Psi_g(\theta)$ such that

$$\Psi_g(\theta) = \frac{1}{1 + \exp[-Da_g(\theta - b_g)]} , \tag{7}$$

that is, the ICC in the logistic model. Thus from (4) and (7) the ICC in LPEF is given by

$$P_g(\theta) = [\Psi_g(\theta)]^{\xi_g} \qquad \xi_g > 0 \ , \tag{8}$$

and from (8) its first and second partial derivatives with respect to θ become

$$P_g'(\theta) = \frac{\partial}{\partial\theta} P_g(\theta) = \xi_g D a_g \, [\Psi_g(\theta)]^{\xi_g} \, [1 - \Psi_g(\theta)] \tag{9}$$

and

$$\frac{\partial^2}{\partial\theta^2} P_g(\theta) = \xi_g D^2 a_g^2 \, [\Psi_g(\theta)]^{\xi_g} \, [1 - \Psi_g(\theta)] \, [\xi_g - (1+\xi_g)\Psi_g(\theta)] \ ,$$

respectively. It has been shown (Samejima, 1996a) that the point of θ at which the discrimination power of the conditional probability becomes maximal increases as more subprocesses are successfully cleared. Thus the eventual ICC depends on how many and how *tough* sequential subprocesses are involved in solving the problem. Note that the word *sequential* is used in a very broard sense. Subprocesses may be either *serial* or *parallel* (Samejima, 1995).

Insert Figure 1 About Here

Figure 1 presents seven examples of the ICC's of LPEF models with the common discrimination parameter $a_g = 1$ and the common difficulty parameter $b_g = 0$, and $\xi_g = 0.3, 0.5, 0.8, 1.0, 1.5, 2.0, 3.0$. When $\xi_g = 1.0$, $P_g(\theta)$ given by (8) becomes the ICC of the logistic model. Thus in LPEF the logistic model is treated as a *transition* from one principle to the other. Note that the ICC is *point-asymmetric* whenever $\xi_g \neq 1$.

Insert Figure 2 About Here

The first derivative of $P_g(\theta)$ given by (9) is negatively skewed when $\xi_g < 1$ and positively skewed when $\xi_g > 1$. Examples are shown in Figure 2 for the same 7 hypothetical items used in Figure 1. Note that setting

$$\theta^* \;=\; a_g(\theta - b_g) \;\;,$$

the left-hand-side of each curve from the vertical line $\theta^* = \theta_0^*$ indicates the $P_g(\theta)$ at

$$\theta \;=\; b_g + \frac{\theta_0^*}{a_g} \;\;.$$

It has been shown (Samejima, 1972) that any model in LPEF satisfies the *unique maximum condition*, as is the case with the positive exponent family of the normal ogive model, which assures uniqueness of the maximum likelihood estimate of θ for each and every response pattern. The two basic functions (Samejima, 1969, 1972) are specified from (7), (8) and (9) as

$$A_g(\theta) \;\equiv\; \frac{\partial}{\partial\theta}\log P_g(\theta) \;=\; \xi_g D a_g[1 - \Psi_g(\theta)] \;>\; 0$$

with

$$\begin{cases} \lim_{\theta\to-\infty} A_g(\theta) & = \;\xi_g D a_g \\ \lim_{\theta\to\infty} A_g(\theta) & = \;0 \;\;, \end{cases}$$

and

$$B_g(\theta) \;\equiv\; \frac{\partial}{\partial\theta}\log[1 - P_g(\theta)] \;=\; \frac{-\xi_g D a_g\,[\Psi_g(\theta)]^{\xi_g}[1 - \Psi_g(\theta)]}{1 - [\Psi_g(\theta)]^{\xi_g}} \;<\; 0$$

with

$$\begin{cases} \lim_{\theta\to-\infty} B_g(\theta) & = \;0 \\ \lim_{\theta\to\infty} B_g(\theta) & = \;-Da_g \;\;, \end{cases} \tag{10}$$

respectively. These asymptotes of the two basic functions are straight-forward, except for the lower asymptote of $B_g(\theta)$. To find this, it is sufficient to prove that

$$\lim_{\theta\to\infty} \;\frac{1 - \Psi_g(\theta)}{1 - [\Psi_g(\theta)]^{\xi_g}} = \frac{1}{\xi_g} \;\;. \tag{11}$$

When ξ_g is rational, we can write

$$\xi_g = \frac{r}{t} \qquad\qquad r > 0, \;\; t > 0 \;\;, \tag{12}$$

where r and t are integers. Substituting (12) into the term on the left hand side of (11) we obtain

$$\frac{1 - \Psi_g(\theta)}{1 - [\Psi_g(\theta)]^{\xi_g}} \;=\; \frac{1 - \Psi_g(\theta)}{1 - [\Psi_g(\theta)]^{r/t}} \;=\; \frac{\sum_{v=0}^{t-1}\,[\Phi_g(\theta)]^v}{\sum_{u=0}^{r-1}\,[\Phi_g(\theta)]^u} \;\;,$$

where

$$\Phi_g(\theta) = [\Psi_g(\theta)]^{1/t} . \tag{13}$$

Thus

$$\lim_{\theta\to\infty} \frac{1-\Psi_g(\theta)}{1-[\Psi_g(\theta)]^{\xi_g}} = \lim_{\theta\to\infty} \frac{1-\Psi_g(\theta)}{1-[\Psi_g(\theta)]^{r/t}} = \lim_{\theta\to\infty} \frac{\sum_{v=0}^{t-1} [\Phi_g(\theta)]^v}{\sum_{u=0}^{r-1} [\Phi_g(\theta)]^u} = \frac{t}{r} = \frac{1}{\frac{r}{t}} = \frac{1}{\xi_g} ,$$

and (11) has been proved. When ξ_g is irrational, we can always find r and t such that

$$\frac{r}{t} < \xi_g < \frac{r+1}{t} \qquad r>0, \quad t>0 , \tag{14}$$

and by increasing r, the interval width of

$$\left(\frac{r}{t}, \frac{r+1}{t}\right) \tag{15}$$

can be made as small as we wish, and

$$\frac{1-\Psi_g(\theta)}{1-[\Psi_g(\theta)]^{\xi_g}}$$

is given as the limiting case where the interval (15) becomes degenerated. Thus all we need to prove is

$$\lim_{\theta\to\infty} \frac{1-\Psi_g(\theta)}{1-[\Psi_g(\theta)]^{(r+1)/t}} = \frac{1}{\frac{r+1}{t}} .$$

This can be done by substituting $(r+1)/t$ for ξ_g on the left hand side of (11), to obtain

$$\lim_{\theta\to\infty} \frac{1-\Psi_g(\theta)}{1-\Psi_g(\theta)^{(r+1)/t}} = \lim_{\theta\to\infty} \frac{\sum_{v=0}^{t-1} [\Phi_g(\theta)]^v}{\sum_{u=0}^{r} [\Phi_g(\theta)]^u} = \frac{t}{r+1} = \frac{1}{\frac{r+1}{t}} .$$

Thus (11) has been proved, and we obtain the second line of (10).

Insert Figure 3 About Here

It is interesting to note that, unlike the upper asymptote of $A_g(\theta)$, the lower asymptote of $B_g(\theta)$ is not affected by the acceleration parameter ξ_g. Figure 3 presents the basic functions $A_g(\theta)$ and $B_g(\theta)$ for the seven hypothetical dichotomous items used in Figures 1 and 2.

The model/item feature function $S_g(\theta)$ (Samejima, 1972, 1997c) in LPEF can be written from (8) and (9) as

$$S_g(\theta) \equiv \frac{\frac{\partial}{\partial\theta}P_g(\theta)}{P_g(\theta)-[1-P_g(\theta)]} = \frac{\xi_g D a_g\,[1-\Psi_g(\theta)]}{1-\{\Psi_g(\theta)\}^{\xi_g}} \;.$$

Thus the partial derivative of $S_g(\theta)$ in LPEF becomes

$$\frac{\partial}{\partial\theta}\,S_g(\theta) = \frac{\xi_g\,D^2\,a_g^2\,\Psi_g(\theta)\,\{1-\Psi_g(\theta)\}\,[\xi_g\,\{\Psi_g(\theta)\}^{\xi_g-1}\,\{1-\Psi_g(\theta)\} - \{1-[\Psi_g(\theta)]^{\xi_g}\}]}{[1-\{\Psi_g(\theta)\}^{\xi_g}]^2}\;, \tag{16}$$

which equals zero and $S_g(\theta)$ becomes the constant $D\,a_g$ when $\xi_g = 1$.

When $\xi_g \neq 1$, the sign of $\frac{\partial}{\partial\theta}\,S_g(\theta)$ is determined by the last factor in parenthesis in the numerator of (16), that is, the sign of

$$\xi_g\,\{\Psi_g(\theta)\}^{\xi_g-1}\,\{1-\Psi_g(\theta)\} - \{1-[\Psi_g(\theta)]^{\xi_g}\} \;. \tag{17}$$

When ξ_g is a rational number expressed by (12), (17) can be written as

$$\xi_g\,\{\Psi_g(\theta)\}^{\xi_g-1}\,\{1-\Psi_g(\theta)\} - \{1-[\Psi_g(\theta)]^{\xi_g}\} \tag{18}$$

$$= \frac{\{1-\Phi_g(\theta)\}\,[r\,\{\Phi_g(\theta)\}^r\,\sum_{u=1}^{t}\{\Phi_g(\theta)\}^{u-1} - t\,\{\Phi_g(\theta)\}^t\,\sum_{v=1}^{r}\{\Phi_g(\theta)\}^{v-1}]}{t\,[\Phi_g(\theta)]^t}\;,$$

where $\Phi_g(\theta)$ is given by (13). From (18) it is obvious that (17) assumes a negative value when $r > t$, and a positive value when $r < t$, and so does $\frac{\partial}{\partial\theta}\,S_g(\theta)$ given by (16). When ξ_g is irrational, again we can always find r and t that satisfy (14), and the interval (15) can be made as small as we want. Following a similar process as we did in proving the inequality in the second line of (10), ξ_g can be treated as the limiting case where the interval (15) degenerates.

From the above observations, therefore, we obtain

$$\frac{\partial}{\partial\theta}\,S_g(\theta) = \begin{cases} < 0 & \text{if } \xi_g > 1 \\ > 0 & \text{if } \xi_g < 1\,. \end{cases}$$

Thus the model/item feature function $S_g(\theta)$ in LPEF is *strictly decreasing* in θ when $\xi_g > 1$, and *strictly increasing* in θ when $0 < \xi_g < 1$. This indicates that for an arbitrary pair of items g and h , with the rest of the response pattern fixed in any sequence, the principle of ordering $\hat{\theta}_v$'s is such that failure in solving the easier of the two items g and h is more penalized than that in solving the more difficult item whenever $\xi_g > 1$ and $\xi_h > 1$, whereas success in solving the more difficult item gets a higher credit than that in solving the easier item whenever $\xi_g < 1$ and $\xi_h < 1$, regardless of the rest of the response pattern.

Let $\alpha_n(\theta)$ and $\beta_n(\theta)$ be

$$\alpha_n(\theta) \;=\; \sum_{g=1}^{n} A_g(\theta)$$

and

$$\beta_n(\theta) \;=\; \sum_{g=1}^{n} B_g(\theta) \quad ,$$

respectively. Thus the likelihood equation can be written as

$$\begin{aligned} \frac{\partial}{\partial\theta} \log\, L_v(\theta) \;&=\; \sum_{g\in G} A_g(\theta) \;+\; \sum_{h\in\bar{G}} B_h(\theta) \\ &=\; \alpha_n(\theta) \;-\; \sum_{h\in\bar{G}} S_h(\theta) \;=\; \beta_n(\theta) \;+\; \sum_{g\in G} S_g(\theta) \;\equiv 0 \quad , \end{aligned} \tag{19}$$

where G is the subset of the n items to which the answers are correct, and $\bar{G}$ denotes its complement, or the subset of items to which the answers are not correct, and by local independence (Lord & Novick, 1968) the likelihood function $L_v(\theta)$ is given by

$$L_v(\theta) \;=\; \prod_{g=1}^{n} P_g(\theta)^{u_g}\, [1 - P_g(\theta)]^{1-u_g} \quad .$$

Equation (19) indicates that the set of item feature functions determines the value of $\hat{\theta}_v$ for a specific response pattern v .

Insert Figure 4 About Here

When the items share a common discrimination parameter and a common acceleration parameter, $S_g(\theta)$'s are identical in shape and placed alongside the abscissa in the order of their difficulty parameters. Figure 4 presents three examples of a set of five model/item feature functions with a common acceleration parameter in each and $a_g = 1$ for all the 5 items in each set, and with the 5 separate difficulty parameters $-3.0, -1.5, 0.0, 1.5, 3.0$, respectively. The common value of ξ_g is 2.0 in the upper graph, 1.0 in the middle graph, and 0.3 in the lower graph. In each graph, also presented are $\alpha_n(\theta)$ (dash and dot) and $\beta_n(\theta)$ multiplied by (-1) (dash and 2 dots), both of which are defined earlier and used in (19). In these graphs, the $\hat{\theta}_v$'s of the response patterns in which all but one item score are correct and those in which all but one item score are incorrect are indicated by arrows. For the first subgroup of 5 response patterns $\hat{\theta}_v$ is given as the value of θ at which the model/item feature function intercepts $-\beta_n(\theta)$, and for the second subgroup as the value of $\hat{\theta}_v$ at which the model/item feature function and $\alpha_n(\theta)$ cross each other.

It is obvious from the results in the upper graph of Figure 4 where $\xi_g = 2.0 \ (> 1.0)$ that with all other item scores being zero the person who has solved the most difficult item obtains the highest $\hat{\theta}_v$ and the person who has solved the easiest item gets the lowest $\hat{\theta}_v$, and between the two values the three $\hat{\theta}_v$'s are arranged in accordance with the difficulty parameters of the items for which $u_g = 1$. It is also obvious in this graph, though less visible, that the person who has solved the 4 most difficult items obtains the highest $\hat{\theta}_v$ and the person who has solved the 4 easiest items gets the lowest $\hat{\theta}_v$, and between the two values the three $\hat{\theta}_v$'s are arranged in accordance with the easiness of the failed item. Thus in these two orderings the same principle is followed consistently. The middle graph represents the logistic model, and, since the 5 items have a common discrimination parameter, the same $\hat{\theta}_v$ is given to all of the 5 response patterns with only one correct answer. All of the 5 response patterns with 4 correct answers also get the same $\hat{\theta}_v$. In the lower graph where $\xi_g = 0.3 \ (< 1.0)$, the rule follows the reversed principle that is used in the upper graph, and this principle is used consistently. Comparison of the three graphs in Figure 4 clarifies the meaning of the logistic model as a transition within LPEF.

Insert Table 2 About Here

Table 2 illustrates the same 32 response patterns as shown in Table 1, arranged in the order of the MLE's obtained by following the LPEF model with the same discrimination and difficulty parameters and a common acceleration parameter $\xi_g = 2$. Note that unlike in Table 1, the logic in ordering all possible response patterns (individuals) is consistent, as exemplified by the reversal of the order of the response patterns (01111), (10111), (11011), (11101) and (11110) in table 2.

The item response information function $I_{u_g}(\theta)$ (Samejima, 1969, 1972) in LPEF is given by

$$I_{u_g}(\theta; u_g = 0) = -\frac{\partial^2}{\partial\theta^2} \log[1 - P_g(\theta)] \tag{20}$$

$$= \frac{\xi_g D^2 a_g^2 \, [\Psi_g(\theta)]^{\xi_g} \, [1 - \Psi_g(\theta)] \, [\xi_g\{1 - \Psi_g(\theta)\} - \Psi_g(\theta)\{1 - [\Psi_g(\theta)]^{\xi_g}\}]}{(1 - [\Psi_g(\theta)]^{\xi_g})^2} > 0$$

and

$$I_{u_g}(\theta; u_g = 1) = -\frac{\partial^2}{\partial\theta^2} \log P_g(\theta) = \xi_g D^2 a_g^2 \, \Psi_g(\theta) \, [1 - \Psi_g(\theta)] > 0 \ , \tag{21}$$

respectively, for $u_g = 0$ and $u_g = 1$. The inequality in (21) is straight-forward. To obtain the inequality in (20), it is sufficient to prove that

$$\xi_g \, [1 - \Psi_g(\theta)] - \Psi_g(\theta) \, [1 - \{\Psi_g(\theta)\}^{\xi_g}] > 0 \ . \tag{22}$$

When ξ_g is rational, using (12) and (13) we can write

$$\begin{aligned} \xi_g \, [1 - \Psi_g(\theta)] \ &- \ \Psi_g(\theta) \, [1 - \{\Psi_g(\theta)\}^{\xi_g}] \\ &= \frac{r}{t} \, [1 - \Psi_g(\theta)] - \Psi_g(\theta) \, [1 - \{\Psi_g(\theta)\}^{r/t}] \\ &= \frac{r}{t} \, [1 - \{\Phi_g(\theta)\}^t] - [\Phi_g(\theta)]^t \, [1 - \{\Phi_g(\theta)\}^r] \\ &= \frac{1}{t} \, [1 - \Phi_g(\theta)][r \sum_{v=0}^{t-1} \{\Phi_g(\theta)\}^v - t\{\Phi_g(\theta)\}^t \sum_{u=0}^{r-1} \{\Phi_g(\theta)\}^u] \\ &= \frac{1}{t} \, [1 - \Phi_g(\theta)][r \sum_{v=0}^{t-1} \{\Phi_g(\theta)\}^v - t \sum_{u=0}^{r-1} \{\Phi_g(\theta)\}^{t+u}] > 0 \ . \end{aligned} \tag{23}$$

When ξ_g is irrational, again (14) holds, and the interval (15) can be made as small as we want. Substituting $(r+1)/t$ for ξ_g in (22), we obtain

$$\frac{r+1}{t} \, [1 - \Psi_g(\theta)] \ - \ \Psi_g(\theta) \, [1 - \{\Psi_g(\theta)\}^{(r+1)/t}] \tag{24}$$

$$= \frac{1}{t}\,[1-\Phi_g(\theta)][(r+1)\sum_{v=0}^{t-1}\{\Phi_g(\theta)\}^v - t\sum_{u=0}^{r}\{\Phi_g(\theta)\}^{t+u}] \;>\; 0 \;\;.$$

Thus, from (23) and (24), (22) has been proved, and, therefore, the inequality in (20) holds.

Insert Figure 5 About Here

From (20) and (21) the item information function $I_g(\theta)$ in LPEF can be written as

$$I_g(\theta) \;=\; E[I_{u_g}(\theta) \mid \theta] \;=\; \frac{\xi_g^2 D^2 a_g^2\,[\Psi_g(\theta)]^{\xi_g}\,[1-\Psi_g(\theta)]^2}{1-[\Psi_g(\theta)]^{\xi_g}} \;>\; 0 \;\;. \qquad (25)$$

Figure 5 presents the item information functions, which are given by (25), of the seven hypothetical items whose ICC's are shown in Figure 1, and also their square roots. These $I_g(\theta)$'s are asymmetric except for the case in which $\xi_g = 1$, that is, in the logistic model, negatively skewed when $\xi_g < 1$, and positively skewed when $\xi_g > 1$. The maximum amount of information increases with ξ_g , and so does the value of θ at which the item information is maximal. Note that the areas under the square root of the item information function for the seven items are the same value, which equals π as was pointed out by Samejima (1997b) in a more general case.

IV. Discussion: Scientific Importance

Researchers tend to accept existing models such as the normal ogive or logistic model without questioning the rationale behind them. We need to look into them, however, to select a substantively validated model. One of the most scientifically important accomplishments of the proposal of LPEF may be to treat a point-symmetric (logistic) model as one of the infinitely many models of the family.

Human behavior is more complex than, say, behavior of agricultural products. Thus the assumption of a linear regression and a conditional normality of the response tendency, given ability, with the normal ogive model, and hence the logistic model as its approximation, is based on may not be suitable, although these models can be used as working hypotheses. LPEF discussed in the present paper, which includes the logistic model as a special case and as a transition between two opposing principles of ordering individuals on the ability scale, provides more appropriate models for human behavior than those based on error distributions.

Item complexity can be illustrated by running marathon with different restrictions. Suppose that some job requires stamina for running a full marathon. If the time required for running does not matter, item difficulty represents the level of physical atamina for running 42.195 kilometers. In such a case each individual may use a strategy of his/her own choice. Some, for instance, may choose to run the whole course with even slow paces that fit them best, while others may choose to change paces many times. Thus there are varieties of ways of running 42.195 kilometers successfully, the union of which is regarded as the accomplishment of the task. This will enhance the probabilities of success for individuals with relatively low levels of stamina, and it will be reasonable to consider an acceleration parameter in the interval $0 < \xi_g < 1$. Thus penalization of failure in completing the race will be more emphasized than giving a credit to the accomplishment of the task. On the other hand, suppose that, to qualify for a running marathon in a prestiged contest, one is required to run a full marathon within 2 hours and 30 minutes. In this second situation, strategies that individuals can take will be narrowed, and the intersection of many elements will lead to success. Thus the probabilities of success for individuals on lower levels of competency will be more reduced, and it will be reasonable to consider an acceleration parameter in the interval $\xi_g > 1$. Credit to those who have pass the criterion will be more emphasized than penalizing failure to qualify.

The acceleration model proposed earlier (Samejima, 1995) is an heterogeneous expansion (Samejima, 1972) of LPEF for ordered polychotomous responses. The model was basically developed for sequential cognitive processes and especially for problem solving. Although in proposing LPEF sequential terminologies have also been adopted, the word is used in a very broad sense, as was pointed out earlier in this paper.

It is expected that direct estimation of the three item parameters in LPEF from a set of raw data should cause indeterminancy problems, as have been observed in the item parameter estimation of the three-parameter logistic model (3PL). Although LPEF does not include noise caused by random guessing which makes estimation of the third parameter, and hence of the other two parameters, in 3PL extremely difficult, indeterminancy problems still occur. To ameliorate the situation, it will be desirable to *uncover* asymmetricity, or the lack of it, in each ICC first, using one of the nonparametric estimation methods such as Levine's (Levine, 1984), Ramsay's (Ramsay, 1991), Samejima's (Samejima, 1997b), etc., and then to parameterize the results. A short-cut parameterization method was proposed by Samejima (1995). It is possible to develop more elaborated methods in the same line without difficulty.

References

[1] Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability*, (Contributed chapters in Lord, F. M. and Novick, M. R., *Statistical theories of mental test scores*, Chapters 17-20, Addison Wesley, Reading, MA.)

[2] Levine, M. (1984). *An introduction to multilinear formula scoring theory.* Office of Naval Research Report, 84-4, Champaign, IL: University of Chicago.

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

[4] Ramsay, J. O. (1991). Kernel smoothing approaches to nonparametric item characteristic curve estimation. *Psychometrika*, *56*, 611-630.

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

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

[7] Samejima, F. (1995). Acceleration model in the heterogeneous case of the general graded response model. *Psychometrika*, 60, 549-572.

[8] Samejima, F. (1996a). Evaluation of mathematical models for ordered polychotomous responses. *Behaviormetrika*, 23, 17-35.

[9] Samejima, F. (1996b). *Polychotomous responses and the test score.* Paper presented at the 1996 National Council of Measurement in Education, New York City, April 10, 1996.

[10] Samejima, F. (1997a). *Graded Response Model*, in W. J. van der Linden & R. Hambleton (Eds.), *Handbook of Modern Item Response Theory*, New York: Springer-Verlag.

[11] Samejima, F. (1997b). Rationale and actual procedures of efficient nonparametric approaches for estimating the operating characteristics of discrete item responses. *Psychometrika*, (in press).

[12] Samejima, F. (1997c). *Virtue of asymmetric item characteristic curves with the logistic model as a trandition.* (Submitted.)

AERA971.TEX
February 19, 1997

TABLE 1

The Maximum Likelihood Estimates of θ Based on 32 Response Patterns of 5 Dichotomous Items Following the Normal Ogive Model and the Logistic Model with the Item Parameters $a_g = 1.0$ for All Items and $b_g = -3.0, -1.5, 0.0, 1.5, 3.0$, Respectively, Arranged in the Ascending Order of Those in the Normal Ogive Model.

	Response Pattern	Normal Ogv.	Logistic
1	00000	neg.infinity	neg.infinity
2	10000	-2.28385	-2.28753
3	01000	-2.27016	-2.28753
4	00100	-1.84831	-2.28753
5	00010	-1.34811	-2.28753
6	01100	-1.15759	-0.75260
7	00001	-0.86577	-2.28753
8	11000	-0.75034	-0.75260
9	10100	-0.75021	-0.75260
10	01010	-0.75013	-0.75260
11	00110	-0.75011	-0.75260
12	00101	-0.36062	-0.75260
13	10010	-0.34310	-0.75260
14	01001	-0.27309	-0.75260
15	00011	-0.19116	-0.75260
16	01110	-0.15292	0.75260
17	10001	0.15292	-0.75260
18	00111	0.19116	0.75260
19	01101	0.27309	0.75260
20	10110	0.34310	0.75260
21	01011	0.36062	0.75260
22	10011	0.75011	0.75260
23	10101	0.75013	0.75260
24	11010	0.75021	0.75260
25	11100	0.75034	0.75260
26	01111	0.86577	2.28753
27	11001	1.15759	0.75260
28	10111	1.34811	2.28753
29	11011	1.84831	2.28753
30	11101	2.27016	2.28753
31	11110	2.28385	2.28753
32	11111	pos.infinity	pos.infinity

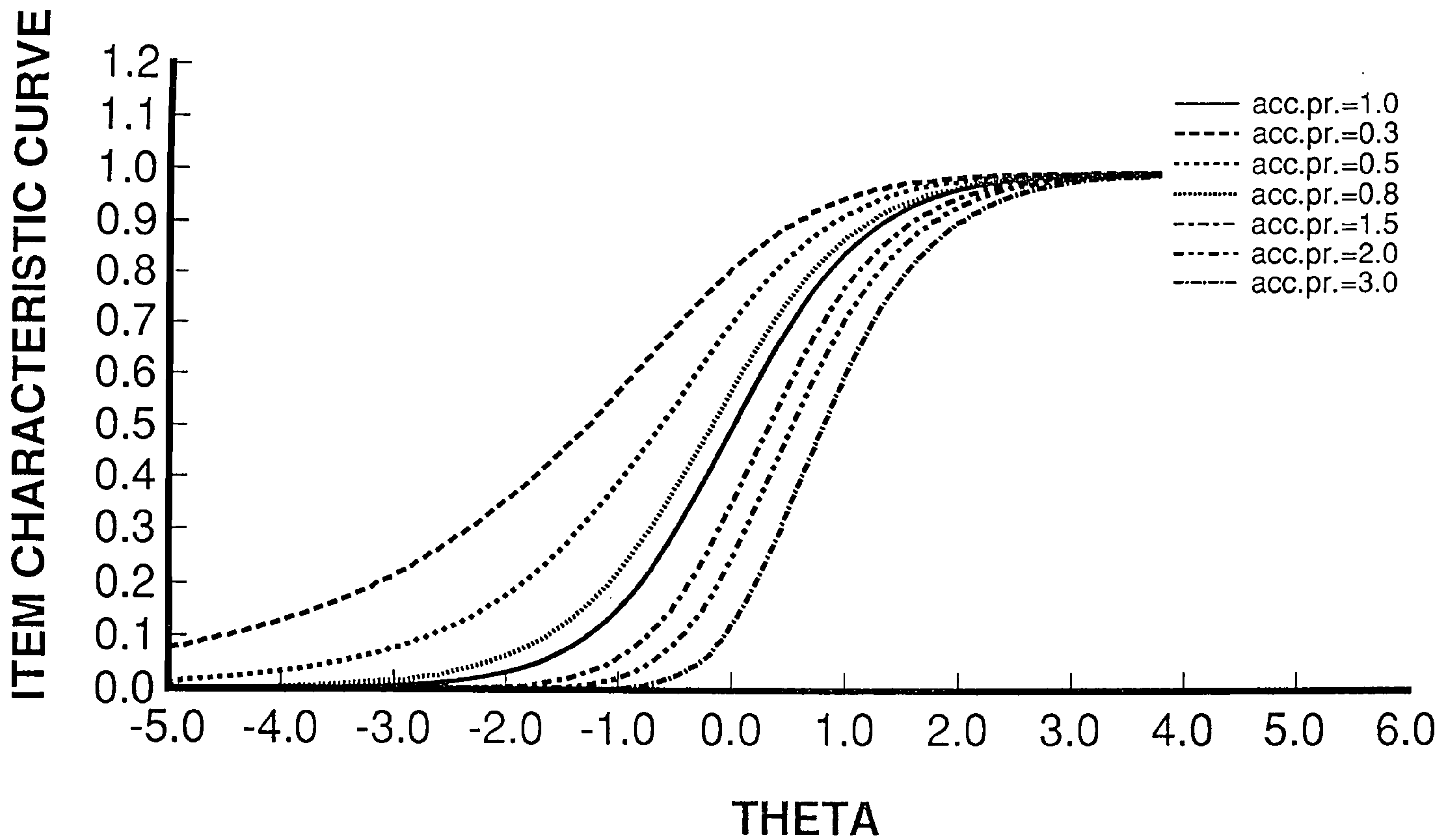
8701nmlog.rst12
8701rasch.rst12
8701nmlog1fort.8
(Taken from 9509n5.rst2)

TABLE 2

The Maximum Likelihood Estimates of θ Based on 32 Response Patterns of 5 Dichotomous Items Following the LPEF model with $\xi_g = 2$, the Normal Ogive Model and the Logistic Model with the Item Parameters $a_g = 1.0$ for All Items and $b_g = -3.0, -1.5, 0.0, 1.5, 3.0$, Respectively, Arranged in the Ascending Order of Those in the LPEF Model.

	Response Pattern	Lgst.Fam.	Normal Ogv.	Logistic
1	00000	neg.infinity	neg.infinity	neg.infinity
2	10000	-1.77109	-2.28385	-2.28753
3	01000	-1.40646	-2.27016	-2.28753
4	00100	-0.83936	-1.84831	-2.28753
5	00010	-0.44235	-1.34811	-2.28753
6	00001	-0.34778	-0.86577	-2.28753
7	11000	-0.24612	-0.75034	-0.75260
8	10100	0.10334	-0.75021	-0.75260
9	01100	0.13035	-1.15759	-0.75260
10	10010	0.66449	-0.34310	-0.75260
11	01010	0.67548	-0.75013	-0.75260
12	00110	0.77745	-0.75011	-0.75260
13	10001	1.06032	0.15292	-0.75260
14	01001	1.06796	-0.27309	-0.75260
15	00101	1.14590	-0.36062	-0.75260
16	11100	1.25580	0.75034	0.75260
17	00011	1.47795	-0.19116	-0.75260
18	11010	1.60421	0.75021	0.75260
19	10110	1.63116	0.34310	0.75260
20	01110	1.63323	-0.15292	0.75260
21	11001	2.16546	1.15759	0.75260
22	10101	2.17644	0.75013	0.75260
23	01101	2.17729	0.27309	0.75260
24	10011	2.27846	0.75011	0.75260
25	01011	2.27907	0.36062	0.75260
26	00111	2.28672	0.19116	0.75260
27	11110	2.76207	2.28385	2.28753
28	11101	3.11779	2.27016	2.28753
29	11011	3.14533	1.84831	2.28753
30	10111	3.14744	1.34811	2.28753
31	01111	3.14760	0.86577	2.28753
32	11111	pos.infinity	pos.infinity	pos.infinity

9512rasch.rst11f
8701nmlog.rst11



0.500 1.00 1.50 6.00 8.00

9601RSC1.DAT, IN9601IC, plotted by M. Foster

FIGURE 1

Seven examples of the item characteristic curves of models in the logistic positive exponent family with the common discrimination parameter $a_g = 1$ and the common difficulty parameter $b_g = 0$, and the seven different acceleration parameters $\xi_g = 0.3, 0.5, 0.8, 1.0, 1.5, 2.0, 3.0$.

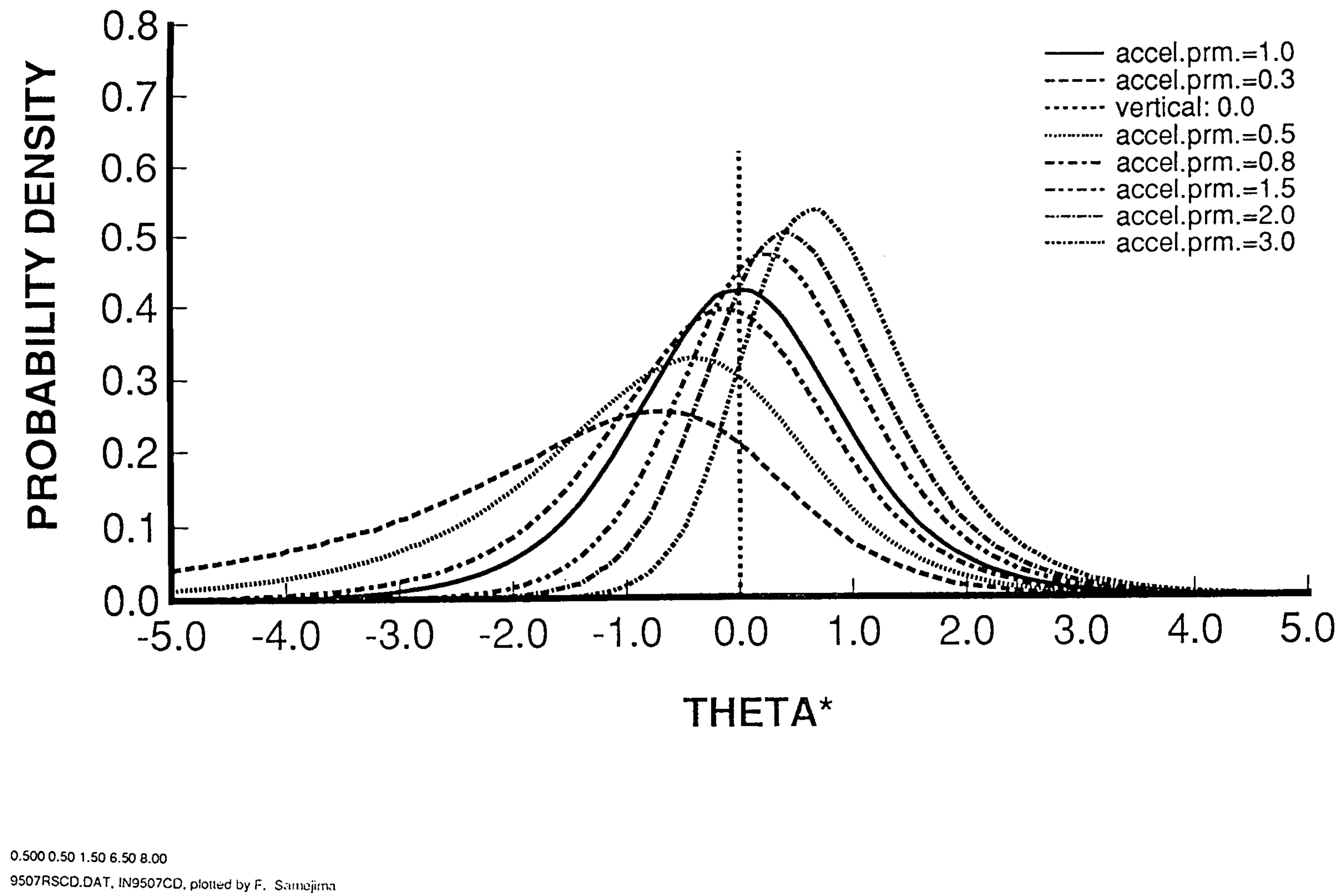
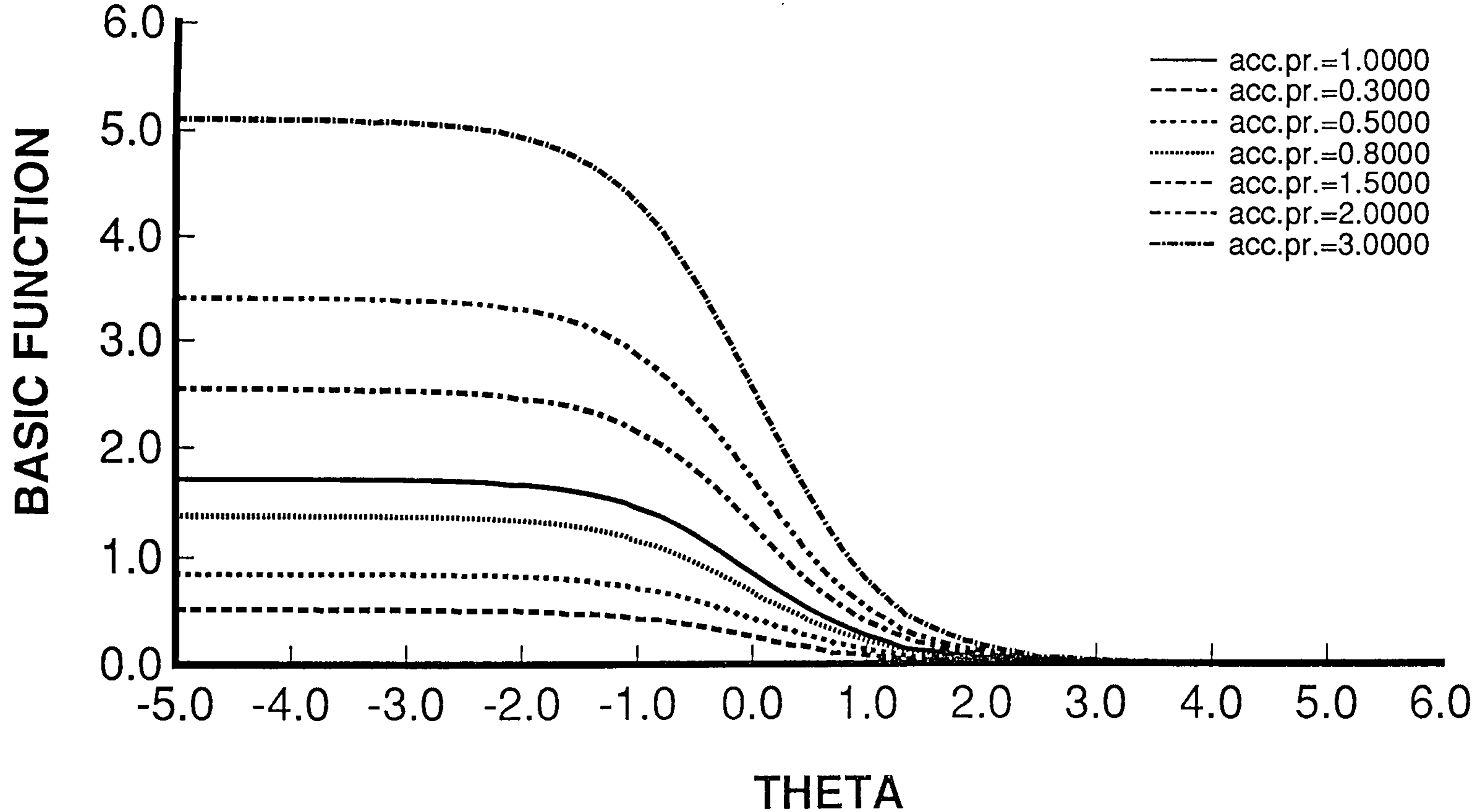


FIGURE 2

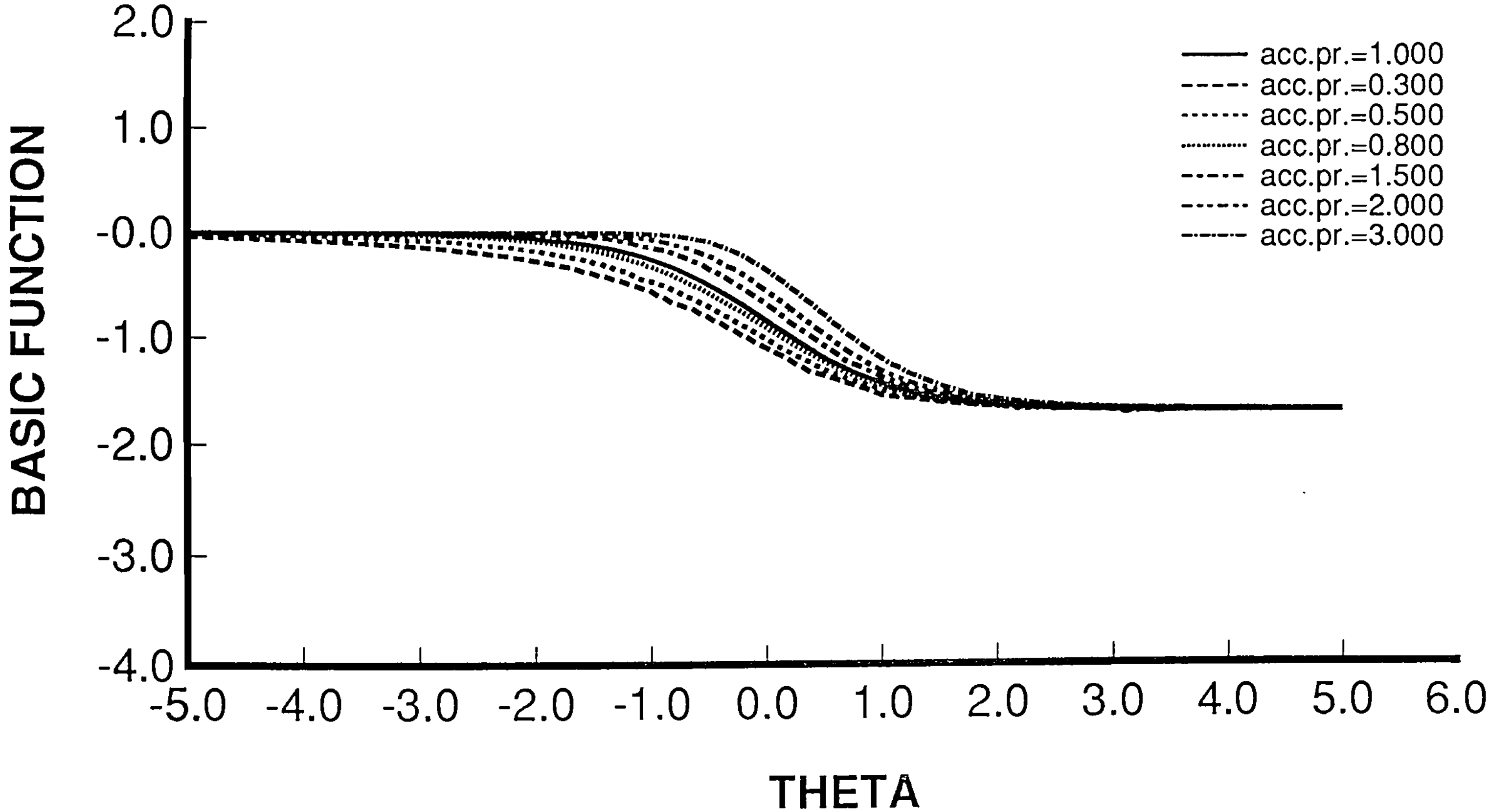
The first derivatives of the seven item characteristic curves shown in Figure 1.



0.500 1.00 1.50 6.00 8.00
9601RS1B.DAT, IN9601BS, plotted by F. Samejima

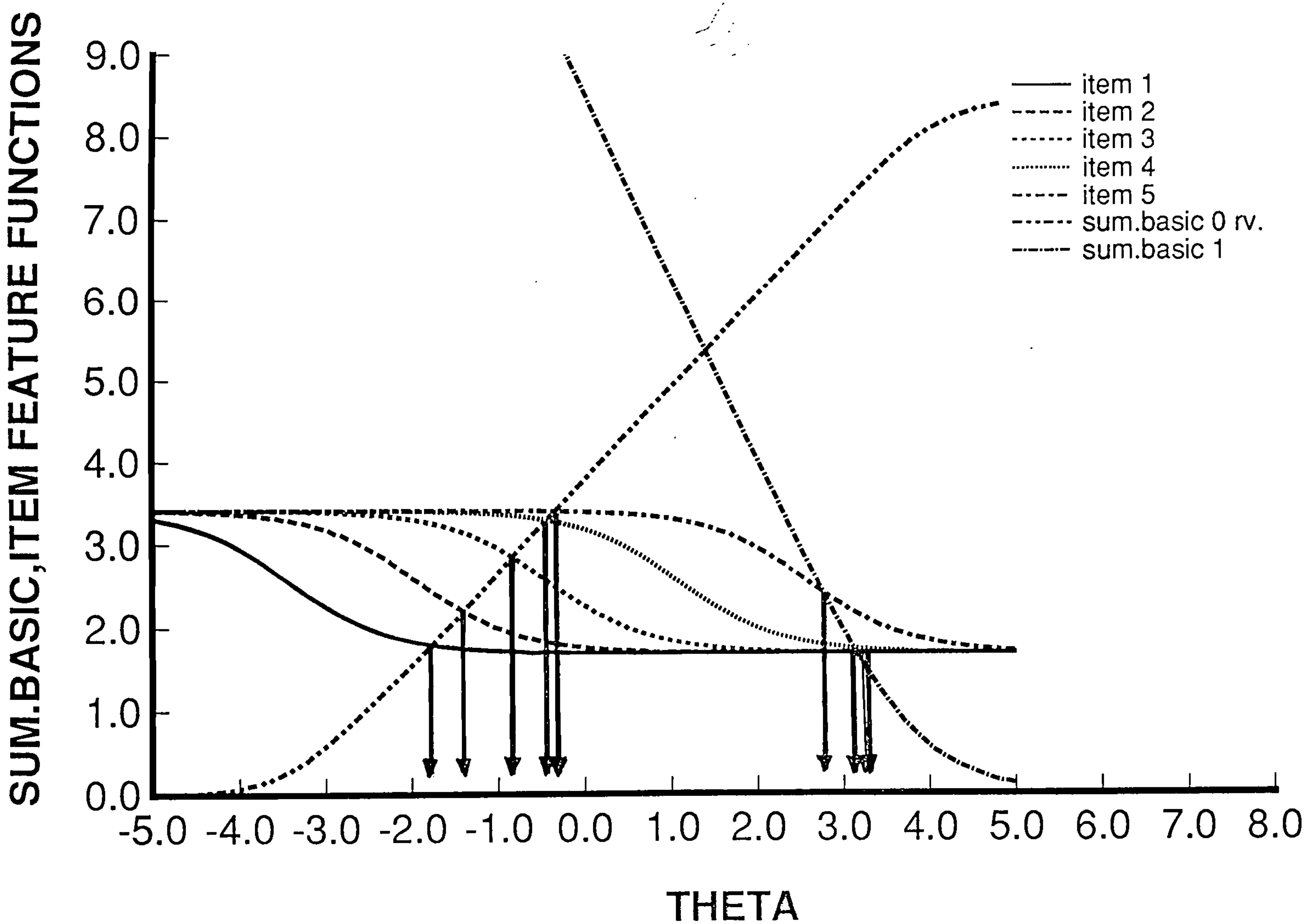
FIGURE 3

The basic functions for the correct answer (upper graph) and for the incorrect answer (lower graph), respectively, for the seven items whose item characteristic curves are shown in Figure 1.



0.500 1.00 1.50 6.00 8.00
9601RS0B.DAT, IN9601BS, plotted by F. Samejima

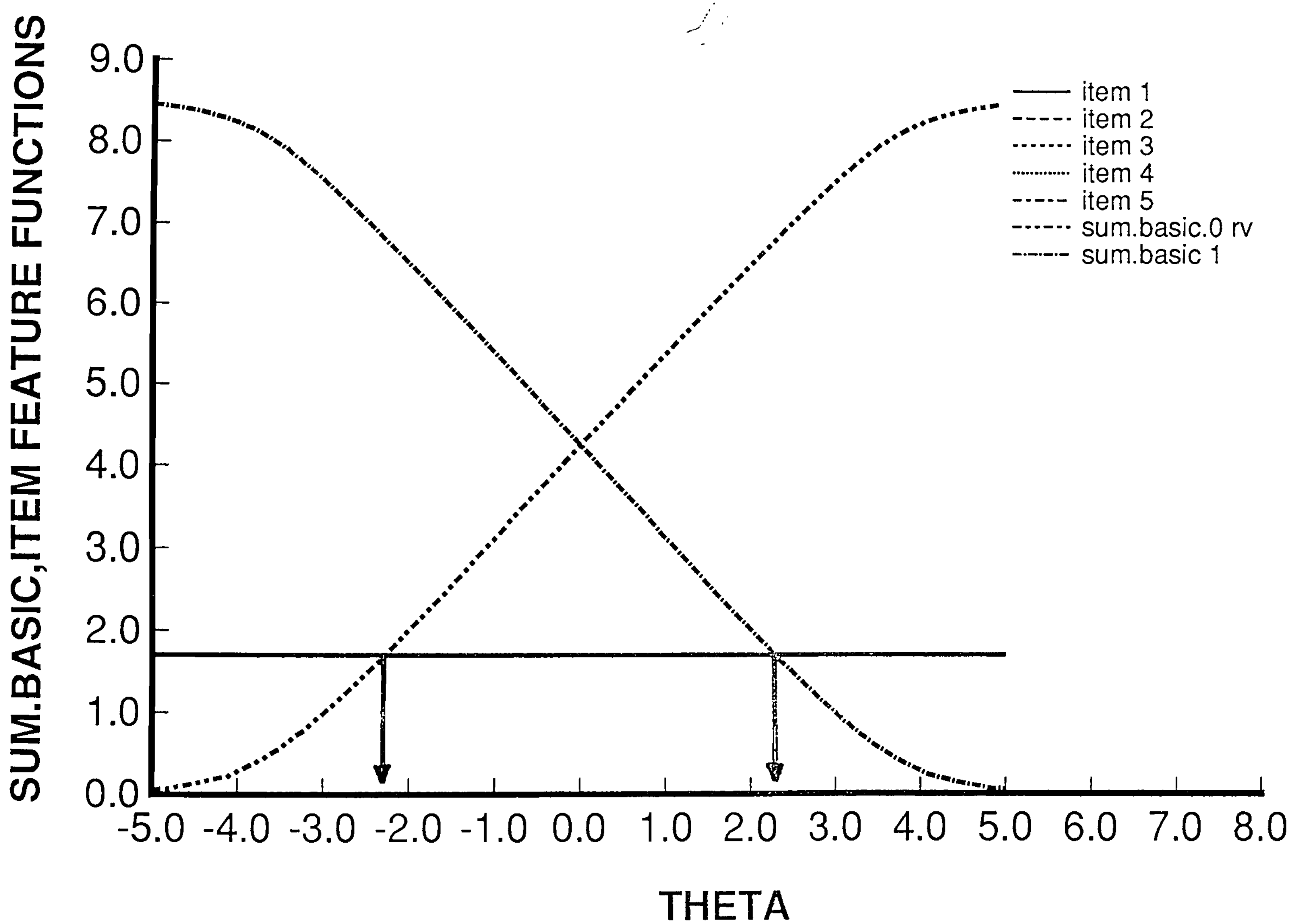
FIGURE 3 (Continued)



0.641 0.50 1.50 7.00 8.00
ITF95LA6.dat, INITF95A, plotted by F.W.Craig

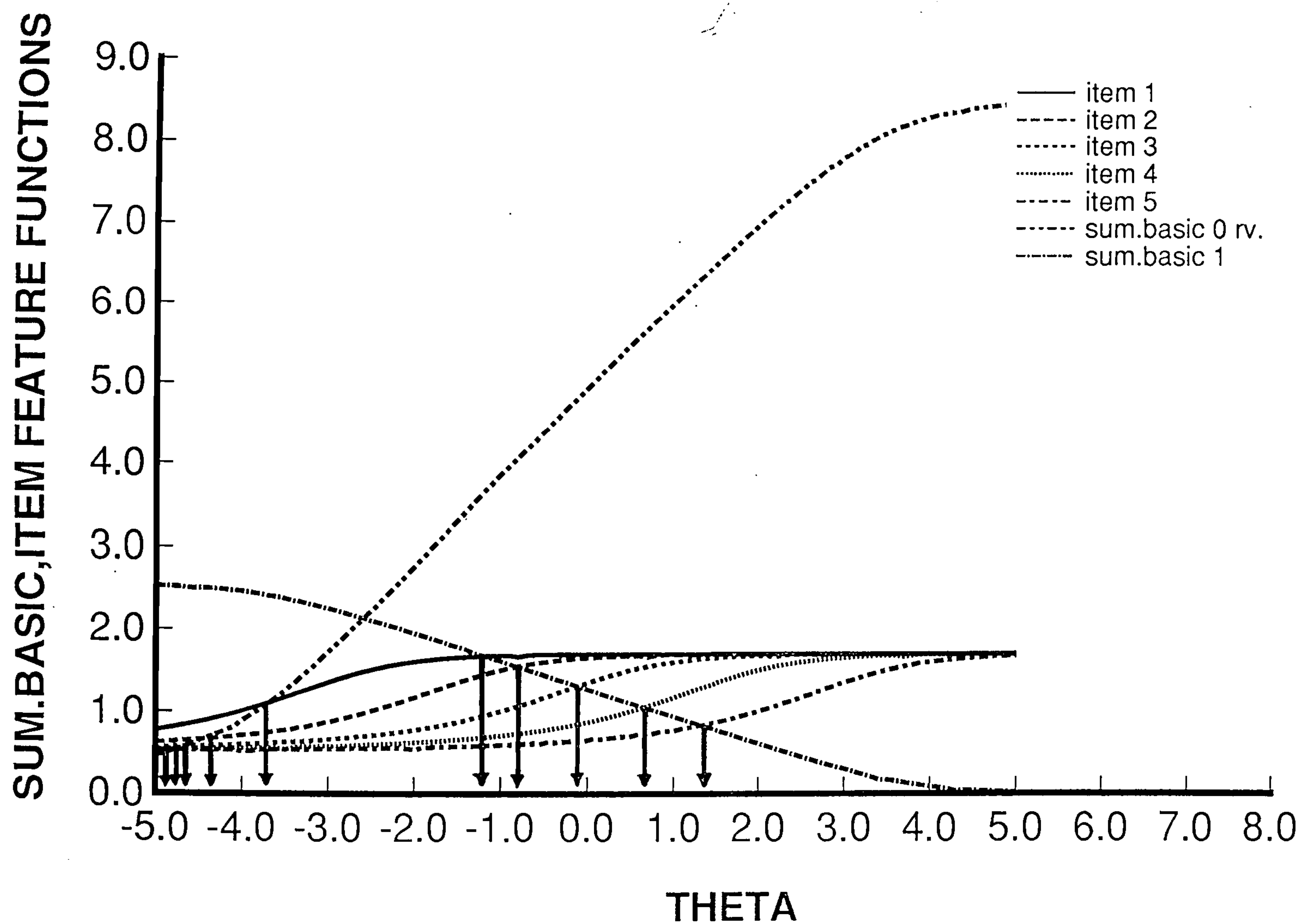
FIGURE 4

Three examples of a set of five model/item feature functions with common acceleration parameters $\xi_g = 2.0$ (upper graph), $\xi_g = 1.0$ (middle graph) and $\xi_g = 0.3$ (lower graph), respectively, with $a_g = 1$ and $b_g = -3.0, -1.5, 0.0, 1.5, 3.0$, together with $\alpha_n(\theta)$ (dash and dot) and $\beta_n(\theta)$ multiplied by (-1) (dash and 2 dots).



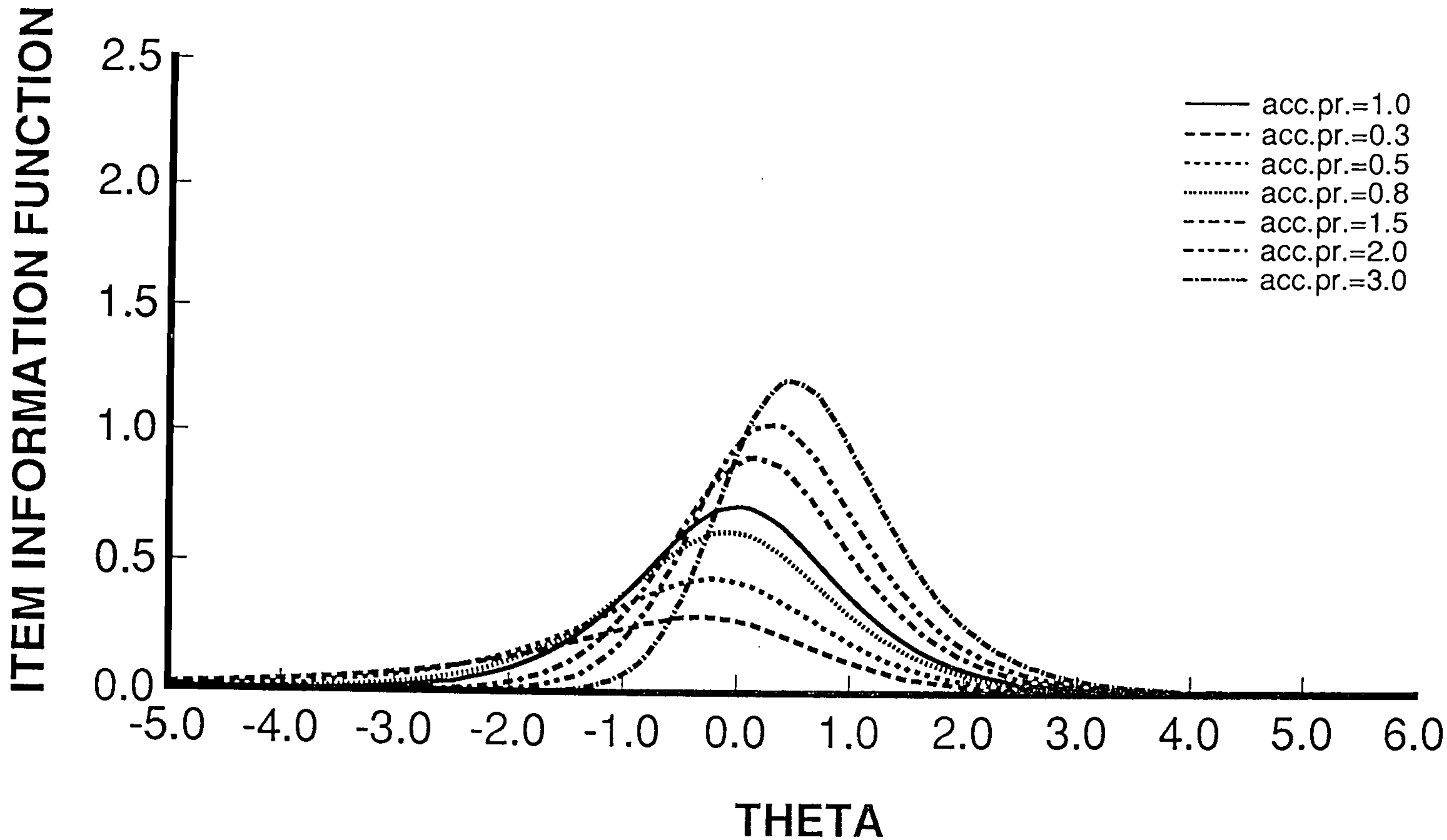
0.641 0.50 1.50 7.00 8.00
ITF95LA4.dat, INITF95A, plotted by F.W.Craig

FIGURE 4 (Continued)



0.641 0.50 1.50 7.00 8.00
ITF95LA1.dat, INITF95A, plotted by F.W.Craig

FIGURE 4 (Continued)



0.500 1.50 1.50 5.00 8.00

9601RSIF.DAT, IN9601IF, plotted by M. Foster

FIGURE 5

Item information functions (upper graph) and their square roots (lower graph), respectively, of the seven items whose item characteristic curves are shown in Figure 1.

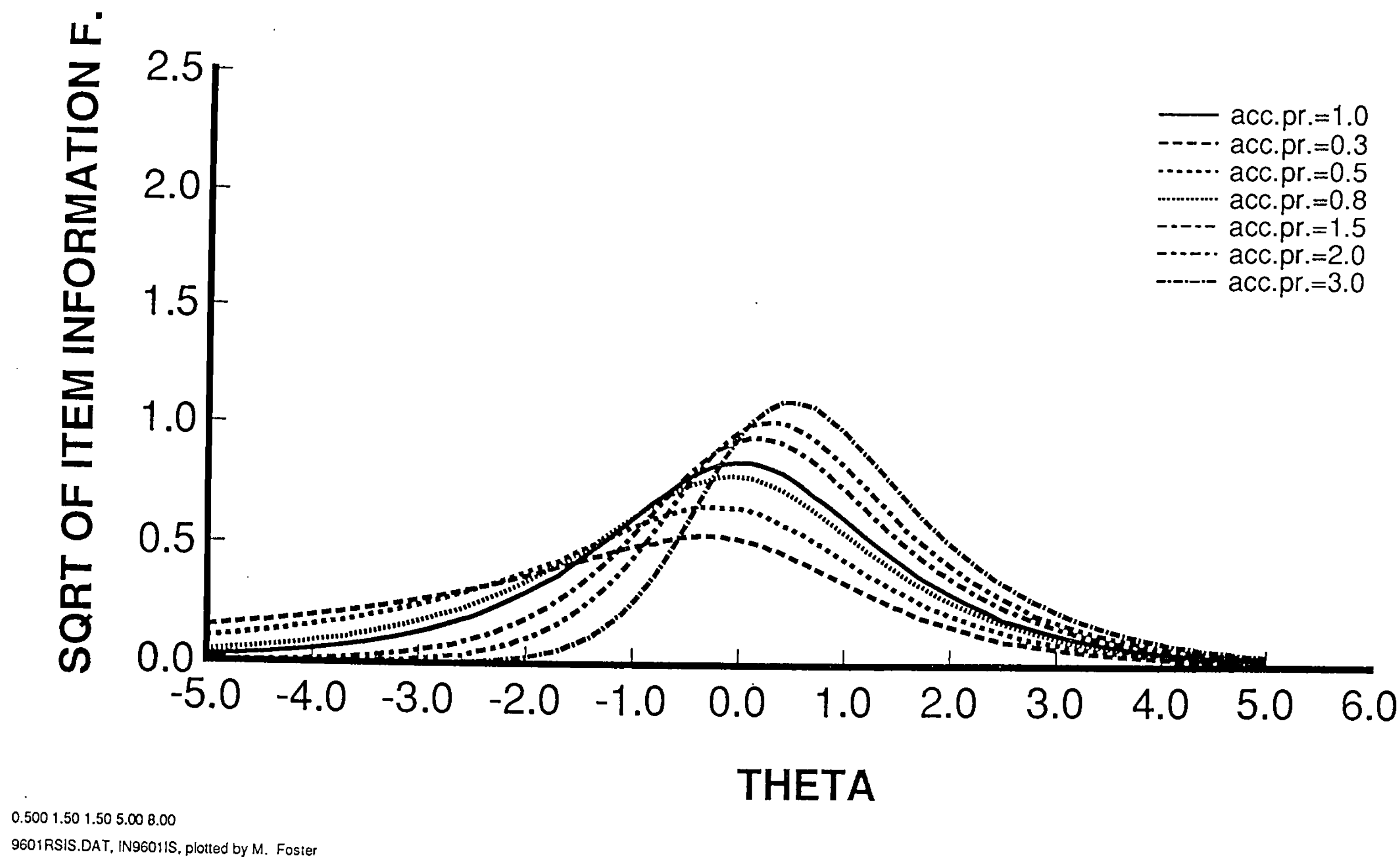


FIGURE 5 (Continued)

TM027571

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Ability Estimates that Order Individuals with Consistent Philosophies

Author(s): Fumiko Samejima

Corporate Source: University of Tennessee

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

Check here	The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2 documents	Check here
[X] **For Level 1 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	[ ] **For Level 2 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.
	Level 1	Level 2	

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: Fumiko Samejima

Printed Name/Position/Title: Fumiko Samejima, Professor

Organization/Address:
University of Tennessee
Department of Psychology
Knoxville, TN 37996-0900

Telephone: (423) 974-6843

FAX: (423) 974-3330

E-Mail Address: samejima@psych1.psych.utk.edu

Date: 2/2/97

CUA

THE CATHOLIC UNIVERSITY OF AMERICA
Department of Education, O'Boyle Hall
Washington, DC 20064

800 464-3742 (Go4-ERIC)

April 25, 1997

Dear AERA Presenter,

Hopefully, the convention was a productive and rewarding event. We feel you have a responsibility to make your paper readily available. If you haven't done so already, please submit copies of your papers for consideration for inclusion in the ERIC database. If you have submitted your paper, you can track its progress at http://ericae2.educ.cua.edu.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are soliciting all the AERA Conference papers and will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality.

Please sign the Reproduction Release Form on the back of this letter and stet **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can mail your paper to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to: AERA 1997/ERIC Acquisitions
The Catholic University of America
O'Boyle Hall, Room 210
Washington, DC 20064

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/E