DOCUMENT RESUME

ED 390 896 TM 024 178

AUTHOR	Muraki, Eiji
TITLE	Information Functions of the Generalized Partial Credit Model.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-93-27
PUB DATE	May 93
NOTE	47p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Item Response Theory; National Surveys; *Scores; *Scoring; *Writing Evaluation
IDENTIFIERS	Calibration; Dichotomous Responses; *Information Function (Tests); National Assessment of Educational Progress; *Partial Credit Model; Polytomous Scoring

ABSTRACT

The concept of information functions developed for dichotomous item response models is adapted for the partial credit model. The information function is explained in terms of the model parameters and scoring functions. The relationship between the item information function and the expected score function is also discussed. The information function is then used to investigate the effect of collapsing and recoding categories of polytomously-scored items of the National Assessment of Educational Progress (NAEP). Finally, the NAEP writing items are calibrated and the item and test information is used to discuss desirable properties of polytomous items. (Contains 1 table, 23 figures, and 15 references.) (Author)

RESEARCH REPORT

INFORMATION FUNCTIONS OF THE GENERALIZED PARTIAL CREDIT MODEL

Eiji Muraki

Educational Testing Service
Princeton, New Jersey
May 1993

Information Functions of the Generalized Partial Credit Model*

Eiji Muraki

Author's Mailing Address: Eiji Muraki, Educational Testing Service, Princeton, NJ 08541

*The author thanks Minhwei Wang for her valuable assistance in the data analyses.

Information Functions of the Generalized Partial Credit Model

ABSTRACT

The concept of information functions developed for dichotomous item response models is adapted for the partial credit model. The information function is explained in terms of the model parameters and scoring functions. The relationship between the item information function and the expected score function is also discussed. The information function is then used to investigate the effect of collapsing and recoding categories of polytomously-scored items of the National Assessment of Educational Progress (NAEP). Finally, the NAEP writing items are calibrated and the item and test information is used to discuss desirable properties of polytomous items.

Index terms: item response model
polytomous item response model
partial credit model
information function
NAEP

Tests have been basically constructed to measure the ability levels of test takers. Since the information function based on the item response theory provides valuable information about the precision of measurement at a specified ability level, the information function has been widely used for various applications that are central to the field of measurement (Hambleton and Swaminathan, 1985). The applications of the information function for dichotomously-scored test items have been extensively studied, but the investigation of the polytomous item response model using the concept of the information function has not been fully explored yet. The partial credit model is frequently used for analyses of polytomous cognitive and affective response data. In this paper, the advantages of the applications of the partial credit model are demonstrated and compared to the dichotomous item response model by applying the information function. The cautionary remarks about the uncritical applications of the partial credit model are also mentioned.

The Generalized Partial Credit Model

The generalized partial credit model (Muraki, 1992) is formulated on the assumption that the probability of choosing the kth category over the k-1th category is governed by the logistic dichotomous response model, that is,

$$P_{jk|k-1,k}(\theta) = \frac{P_{jk}(\theta)}{P_{j,k-1}(\theta)+P_{jk}(\theta)} = \frac{\exp[Da_j(\theta-b_{jk})]}{1+\exp[Da_j(\theta-b_{jk})]} \tag{1}$$

where $k=2, 3, \ldots, m_j$. The generalized partial credit model is, then, written as

$$P_{jk}(\theta) = \frac{\exp[\sum_{v=1}^{k} Z_{jv}(\theta)]}{\sum_{c=1}^{m_j}\exp[\sum_{v=1}^{c} Z_{jv}(\theta)]} \tag{2}$$

and

$$Z_{jk}(\theta) = Da_j(\theta-b_{jk}) = Da_j(\theta-b_j+d_k) \tag{3}$$

where D is a scaling constant that puts the θ scale in the same metric as the normal ogive model ($D=1.7$), a_j is a slope parameter, b_{jk} is an item-category, b_j is an item location parameter, and d_k is a category parameter.

If the number of response categories is m_j, only m_j-1 category parameters can be estimated. Any one of the m_j category threshold parameters can be defined as any value. This is because the term including the parameter is canceled out from both the numerator and denominator of the model (Muraki, 1992). We arbitrarily define $d_1 = 0$.

If a single Likert scale is used to evoke categorical responses for a set of items, the hypothesis that the set of items share common category parameters can be tested. If the assumption of a common set of category parameters for those items is not found to be appropriate, we fit a model in which category parameters differ between items. For both models, the notation d_k instead of d_{jk} can usually be used without confusion. A step-wise application of these two models to data was demonstrated by Muraki (1992). The partial credit model with a constant slope parameter was introduced by Masters (1982). This model is a special case of the generalized partial credit model in Equation 2. The comparison between these models was also discussed by Muraki (1992).

For the generalized partial credit model, there is an indeterminacy in the set of category parameters and location parameter. The following constraint, called a location constraint, is imposed on the category parameters within a categorical scale:

$$\sum_{k=2}^{m_j} d_k = 0 \tag{4}$$

The partial credit model contains the element $Z_{jk}^{+}(\theta)$, that is,

$$Z_{jk}^{+}(\theta) = \sum_{v=1}^{k} Z_{jv}(\theta) \tag{5}$$

where $Z_{jv}(\theta)$ is defined by Equation 3. The sum of $Z_{jv}(\theta)$ in Equation 5 can be written as

$$Z_{jk}^{+}(\theta) = Da_j[(k(\theta-b_j)+\sum_{v=1}^{k} d_v] \tag{6}$$

and the model above can be rewritten as

$$Z_{jk}^{+}(\theta) = Da_j[T_k(\theta-b_j)+K_k] \tag{7}$$

Andrich (1978) calls T_k and K_k the scoring function and the category coefficient, respectively. For the partial credit model, the scoring function T_k is a linear integer scoring function, that is, $\underline{T}$=(1, 2, 3, ..., m_j) where m_j is the number of categories of item j.

The log-odds of the model probabilities $P_{j,k-1}(\theta)$ and $P_{jk}(\theta)$ can be expressed as

$$\lambda_{j,k|k-1} = Da_j[(T_k - T_{k-1})(\theta - b_j) + d_k] \qquad (8)$$

Equation 8 shows that the log-odds becomes a monotonically increasing function of the latent trait, θ, only when the incremental change in the scoring function is used for successive categorical responses. The higher θ value a subject has, the more likely he/she responds to upper categories. In other words, the partial credit model becomes a model for ordered categorical responses only when the scoring function is increasing, that is, $T_k > T_{k-1}$, for any k and $a_j>0$. If the linear integer scoring function is used, the item-category characteristic curves (ICCCs) of $P_{j,k-1}(\theta)$ and $P_{jk}(\theta)$ intersect at the point b_{jk} in the θ scale. When a person's ability θ is greater b_{jk}, the person more likely chooses the kth category than the k-1th category. In this case, the log-odds in Equation 8 is positive. The log-odds increases as ability level increases. When a person's ability θ is less than b_{jk}, the person more likely chooses the k-1th category than the kth category. In this case, the log-odds is negative. The odds of choosing the k-1th category compared to the kth category increases as the ability level becomes lower.

The generalized partial credit model in Equation 2 can be rewritten, using the scoring function and the category coefficient, as

$$P_{jk}(\theta) = \frac{\exp[Da_j[T_k(\theta - b_j) + K_k]]}{\sum_{c=1}^{m_j} \exp[Da_j[T_c(\theta - b_j) + K_c]]} \qquad (9)$$

The model expressed by Equation 9 is similar to the nominal response model (Bock, 1972). The relationship to the generalized partial credit model is discussed by Thissen and Steinberg (1986). The scoring function also provides a convenient notation for collapsing or recoding categorical responses. For example, if the number of categorical responses of an item is five, then a scoring function $\underline{T}$ can be specified as $\underline{T}$=(1,2,3,4,5). If the original response categories are collapsed by combining the first and second categories into one category, the modified scoring function $\underline{T}$' can be written as $\underline{T}$'=(1,1,2,3,4). If this modification of the response categories is recoded by treating the original fourth category as the fifth and the original fifth as the fourth, the scoring function can be further modified to $\underline{T}$''=(1,1,2,4,3).

Information Function

The item information function, $I_j(\theta)$, represents the information contributed by a specific item j across the range of θ. The item information for the polytomous item response model was proposed by Samejima (1974) as

$$
\begin{aligned}
I_j(\theta) &= \sum_{k=1}^{m_j} P_{jk}(\theta)\,[-\frac{\partial^2}{\partial\theta^2}\ln P_{jk}(\theta)] \\
&= \sum_{k=1}^{m_j} P_{jk}(\theta)\,[[\frac{\frac{\partial}{\partial\theta}P_{jk}(\theta)}{P_{jk}(\theta)}]^2 - \frac{\frac{\partial^2}{\partial\theta^2}P_{jk}(\theta)}{P_{jk}(\theta)}] \\
&= D^2a_j^2\sum_{k=1}^{m_j} P_{jk}(\theta)\,[\sum_{c=1}^{m_j} T_j^2P_{jc}(\theta) - [\sum_{c=1}^{m_j} T_cP_{jc}(\theta)]^2] \\
&= D^2a_j^2\,[\sum_{c=1}^{m_j} T_c^2P_{jc}(\theta) - [\sum_{c=1}^{m_j} T_cP_{jc}(\theta)]^2] \\
&= D^2a_j^2\sum_{c=1}^{m_j}\,[T_c-\bar{T}_j(\theta)]^2P_{jc}(\theta)
\end{aligned}
\tag{10}
$$

where

$$
\bar{T}_j(\theta) = \sum_{c=1}^{m_j} T_cP_{jc}(\theta) \tag{11}
$$

The value of $T_j(\theta)$ is called the expected score function. Thus, the item information is the expected variance of scoring functions based on the model probability along the θ level multiplied by $(Da_j)^2$. Notice that the second derivative of the logarithm of categorical probability, $P_{jc}(\theta)$, with respect to θ is invariant with the category k. For the case of dichotomous item responses, Equation 11 can be simplified to

$$
I_j(\theta) = D^2a_j^2(T_1-T_2)^2P_{j1}(\theta)P_{j2}(\theta) \tag{12}
$$

where $P_{j2}(\theta)=1-P_{j1}(\theta)$.

Bock (1972) proposed defining the information due to the response in category k of item j as a partition of the item information due to that category, that is,

$$
I_{jk}(\theta) = P_{jk}(\theta)\,I_j(\theta) \tag{13}
$$

Equation 13 may be called the item-category information function. The item-category information functions of $I_{j,k-1}(\theta)$ and $I_{jk}(\theta)$ intersect at the point of b_{jk} in the θ scale.

The item information function can also be expressed as the summation of the item-category information function:

$$I_j(\theta) = \sum_{k=1}^{m} I_{jk}(\theta) \tag{14}$$

Finally, the test information function is defined as the summation of item information functions:

$$I(\theta) = \sum_{j=1}^{n} I_j(\theta) \tag{15}$$

Plotting the ICCCs of $P_{jk}(\theta)$ for k=1, 2, ..., m_j and their information functions is an essential step to interpretation of the parameters of the polytomous item response model. Figures 1,2, 3, 4, and 5 show plots of ICCCs of the generalized partial credit model with three categorical responses, where a_j=1 and $\underline{T}$=(1, 2, 3). The only differences among these items are the values of the item-category parameters, $\underline{b}_j$. Figures 1 through 5 show five items with $\underline{b}_1$=(0., 2., -4.), $\underline{b}_2$=(0., 2., -2.), $\underline{b}_3$=(0., 2., 0.), $\underline{b}_4$=(0., 2., 2.), and $\underline{b}_5$=(0., 2., 4.), respectively.

****
Insert Figures 1, 2, 3, 4, and 5 about here
****

As shown in Figures 1 through 5, the third ICCC, $P_{j3}(\theta)$, shifts to the left along the θ scale as the item-category parameter b_{jk} changes from -4.0 to 4.0 because the intersection of the second and third ICCCs, $P_{j2}(\theta)$ and $P_{j3}(\theta)$, moves from 4.0 to -4.0. Since the third category probability, P_{j3} becomes dominant over the θ axis from -6.0 to 6.0, the expected frequency of the third categorical response increases, assuming the latent trait is distributed normally with the mean zero. As the third category probability moves from the right to the left, the second category probability is being pushed downward. The expected frequency of the second category decreases as the distance between the second and third item-category parameters, b_{j2} and b_{j3}, decreases. The second and third item-category parameters are equal for item 4, and then the order of these parameters are interchanged for item 5. As shown in Figure 5, the polytomous responses for item 5 become essentially dichotomous item responses since the second category becomes a category not likely occurring as a response.

Item-category information functions and item information functions of items 1, 2, 3, 4, and 5 are shown in Figures 6, 7, 8, 9, and 10, respectively. The item information of polytomous item responses is not necessarily unimodal like that of dichotomous item responses. If the distance between the two adjacent item-category parameters is large, like b_{j2} and b_{j3} in Figure 6, the information becomes lower at the middle range of the θ scale. The loss of information over the middle range of the θ level becomes less noticeable as the distance between these parameters decreases, as shown in Figure 6. When the distance between b_{j2} and b_{j3} is 2.0, the item information function looks unimodal, as shown in Figure 8. This item is the most desirable if the item is designed to cover a wide range of θ for a group of subjects whose abilities are assumed to be normally distributed. If the order of these parameters is interchanged, the item information peaks over a very short range of the lower θ axis. Information at the higher ability levels is essentially lost. As shown in Figure 10, the shape of the item information function resembles that of dichotomous item responses. The plots of ICCCs and item and item-category information functions of polytomously-scored items provide valuable item analysis information.

****
Insert Figures 6, 7, 8, 9, and 10 here
****

The first derivative of the expected score function with respect to θ is

$$
\begin{aligned}
\frac{\partial \bar{T}_j(\theta)}{\partial\theta} &= Da_j \sum_{c=1}^{m_j} T_c(T_c - \bar{T}_j(\theta))P_{jc}(\theta) \\
&= Da_j \sum_{c=1}^{m_j} [T_c - \bar{T}_j(\theta)]^2 P_{jc}(\theta) \\
&= \frac{I_j(\theta)}{Da_j} \\
&> 0
\end{aligned}
\qquad (16)
$$

Because the first derivative of the expected score function with respect to θ in Equation 16 is always positive unless the slope parameter is negative, the expected score function is a strictly increasing function of θ. The expected score functions of items 1, 2, 3, 4, and 5 are plotted in Figure 11. The expected function becomes steeper if the slope parameter value or the item information increases. The slope of the expected score function is largest at the point of θ where the item information is a maximum. This is the reason that the expected score functions of the five items are ordered along the θ axis in Figure 11. It should also be pointed out that item 1 appears to have two extreme slopes because its item information has two modes.

Unlike the category parameters of the graded item response model (Samejima, 1969, 1972; Muraki, 1990), the item-category or category parameters are not necessarily ordered. However, if the order of the parameters is interchanged as in item 5, the category is depressed and the category becomes useless in terms of the contribution of the item-category information to the total item information. Such items are usually undesirable.

****
Insert Figure 11
****

Information Function and Collapsing and Reordering Categorical Responses

The item-category information functions can be used to investigate the appropriateness of collapsing or recoding of categorical responses. One item from the 1990 mathematics cross-sectional assessment and two items from the 1990 science cross-sectional assessment of the National Assessment of Educational Progress (NAEP) are used to illustrate the behaviors of item-category information functions with respect to the collapsing and recoding of categorical responses. Item parameters were calibrated by using the PARSCALE program (Muraki and Bock, 1991).

The first example is a mathematics item with seven categories. The scoring function is denoted as $\underline{T}_7=(1,2,3,4,5,6,7)$. The item parameters of this item were estimated with other items in the assessment. Then, the sixth and seventh categorical responses were combined and the parameters were estimated again. The scoring function after this collapsing is denoted as $\underline{T}_6=(1,2,3,4,5,6,6)$. The collapsing process was continued until only two categories remained. The scoring function of this collapsing is denoted as $\underline{T}_2=(1,2,2,2,2,2,2)$. After each collapsing, the model parameters were estimated and the item information functions were computed. They are plotted in Figure 12. Since the highest category was collapsed with the adjacent categories in sequence, the peak of the item information function moves to the left along the θ scale. In other words, by collapsing higher categorical responses, the information about subjects with higher θ values decreased. At the same time, the maximum information is decreasing, except in changing from $\underline{T}_3$ to $\underline{T}_2$.

****

Insert Figure 12 about here

****

The second example is a Life Science item with six categorical responses. This item was administered to two age groups. The item response information functions based on the item parameters estimated with the scoring function $\underline{T}=(1,2,3,4,5,6)$ are plotted in Figure 13. The mean, v_k, of the total score distribution for the subgroup in category k (k=1, 2,..., 6) of the item was computed. The means are $\underline{v}=(6.3, 7.7, 7.9, 8.3, 8.9, 10.2)$ and $\underline{v}=(8.5, 8.8, 9.4, 9.4, 10.5, 11.9)$ for the first and second age groups, respectively. The means for the first and second categories for the second age group and the means of the third and fourth categories for both age groups were very close for both age groups. In other words, these categories did not seem to be differentiated. Thus, Allen (1992) decided to combine these categories forming four categorical responses. The scoring function for this collapsing can be denoted as $\underline{T}=(1,1,2,2,3,4)$, and the item response functions based on this scoring function are plotted in Figure 14. The item information functions in Figures 5 and 6 are labeled 1+2+3+4+5+6 and 1+2+3+4, respectively. The peak of the item information function becomes higher after the collapsing, and the amount of information increases over the θ scale except at its lower end, which may be explained by the collapsing effect of lower categories.

****

Insert Figures 13 and 14 about here

****

The third example is a Physical Science item with four categorical responses. This item was also administered to two age groups. The item response information functions based on the item parameters estimated with the scoring function, $\underline{T}=(1,2,3,4)$, are plotted in Figure 15. The mean vectors for the two age groups are $\underline{v}=(4.8, 6.4, 6.2, 7.5)$ and $\underline{v}=(9.3, 10.9, 10.6, 11.5)$. From these statistics, it can be suspected that the original codings of the second and third categories were inappropriate (Allen, 1992). Thus, the orders of the second and third categories were reversed, and the scoring function can be denoted as $\underline{T}'=(1,3,2,4)$. The item response functions based on this scoring function are plotted in Figure 16. The information increases considerably over the range of θ values from -4.0 to 4.0. The chi-square fit statistic also improved from 370.002 with 87 degrees of freedom to 151.249 with 82 degrees of freedom. The difference, 218.753, with 5 degrees of freedom is a significant improvement for the model fit.

****
Insert Figures 15 and 16 about here
****

Analysis of the National Assessment of Educational Progress - Writing Assessment

The first NAEP writing assessment was conducted in 1984 (Grima and Johnson, 1991). Since then, four NAEP writing assessments were conducted to assess the trend of students' writing performance over the years (1984, 1988, 1990, and 1992).

In the trend assessment, nationally representative samples of students in grades 4, 8, and 11 respond to a series of writing tasks (or items). A set of 12 writing tasks were prepared to examine students' abilities to engage in three types of writing: informative, persuasive, and imaginative. The twelve tasks were administered in a balanced incomplete block (BIB) design with six of the twelve tasks presented to each grade. Some writing tasks were unique to a specific grade, and others were utilized as linking items. Students' writirgs were scored by trained readers on the basis of students' success in accomplishing the specific purpose of each writing task (as measured by primary trait scoring), their relative writing fluency (as measured by holistic scoring), and their mastery of the conventions of writing English (as measured by their spelling, punctuation, and grammar). For this paper, only the primary trait scores of the 1988 and 1990 assessments were analyzed.

For two writing tasks, a four-point scoring scale (Unsatisfactory through Adequate) was used to evaluate students' writings. For the remaining four writing tasks, a five-point scoring scale (Unsatisfactory through Elaborate) was used. Omitted responses were not rated. In the analysis, they were treated as the lowest categorical response (Unsatisfactory). Item parameters were calibrated based on the combined response data of the 1988 and 1990 assessment. Since we could not assume that the latent trait distributions of these assessments were unchanged, a separate normal prior was used for each assessment year. Forty one quadrature points were used for both prior distributions. After each EM cycle, the means and standard deviations were computed for these distributions. After each EM cycle, the weighted mean and standard deviation of the combined distribution of these priors are adjusted to 0.0 and 1.0, respectively. Intermediate estimated values were also adjusted accordingly.

The sample sizes of the 1988 and 1990 assessments are 4878 and 5606, respectively. Forty-six EM cycles were needed for convergence, using a criterion of 0.001. Because there were no responses in the fifth category of the second task in the 1990 assessment, the fourth and fifth categories of the second task were combined for both years. Thus, the second task is treated as an item of four categories. The estimated parameter values for age 9 are presented in Table 1. The item information functions of these six items (or tasks) are plotted in Figures 17, 18, 19, 20, 21, and 22. The test information function is plotted in Figure 23.

****
Insert Table 1 about here
****

****
Insert Figures 17, 18, 19, 20, 21, and 22 about here
****

****
Insert Figure 23 about here
****

Since the slope parameter of the first item is low, the item information function is relatively flat over the range of the θ scale. In addition to a low slope parameter, item 3 also has widely-dispersed category parameters. Because of this combination, when compared to the other items, the item information function for this item is the lowest of the six items for all θ values. The slope parameter of item 4 is not extremely low compared to items 1 and 3, but the category parameters are considerably dispersed . Consequently, the amount of information increases over the θ scale, but the information curve is relatively flat. The best items among the six seem to be items 5 and 6. The amount of information for item 5 is large for the lower end of the θ scale because the location parameter of this item is low. Compared to item 5, the information function of item 6 is more symmetric and reasonably high over the range of -3 to 3 of the θ scale. The test information plot in Figure 23 shows that this set of six items produces a reasonable amount of information for the range of θ.

Conclusion

The concept of information functions developed for dichotomous item response models (Lord, 1980; Lord and Novick, 1968) is adapted for the partial credit model. Because of the complex relationships among the parameters of the partial credit model, the plots of the ICCCs should be a routine step for analyses of test items. Computing and plotting the item-category and item information functions based on the partial credit model is also an essential procedure of item analyses. Information functions with other conventional item statistics provide valuable information about how to collapse or reorder categorical responses. It was pointed out that increasing the number of categorical responses does not automatically increase the information about ability levels for the entire range of the θ scale. Assembling polytomous items into a desirable test is also facilitated by using the information functions. Careful investigation of each test item in terms of its information function leads to desirable test construction methodology for polytomously-scored items. Applications of the partial credit model to polytomous items discussed in this paper is only an initial endeavor for this new psychometric field. Further investigation of the partial credit model and related concepts is needed.

References

Allen, N. L. (1992). An application of BILOG/PARSCALE to the 1990 science cross-sectional items. A paper presented at Design and Analysis Committee of the National Assessment of Educational Progress, Alexandria, VA, February 20-22.

Andrich, D. (1978). A rating formulation for ordered response categories. Psychometrika, 43, 561-573.

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

Grima, A. M. & Johnson, E. G. (1991). Data analysis for the writing assessment. The NAEP 1990 Technical Report. Princeton, NJ: Educational Testing Service.

Hambleton, R. K. & Swaminathan, H. (1985). Item response theory: Principles and applications. Boston, MA: Kluwer-Nijhoff Publishing.

Lord, F. M. (1980). Application of item response theory to practical testing problems. Hillsdale, NJ: Erlbaum.

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

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

Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. Applied Psychological Measurement, 14, 59-71.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. Applied Psychological Measurement (in press).

Muraki, E., & Bock, R. D. (1991). PARSCALE: Parameter scaling of rating data [Computer program]. Chicago, IL: Scientific Software, Inc.

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

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

Samejima, F. (1974). Normal ogive model on the continuous response level in the multidimensional latent space. Psychometrika, 39, 111-121.

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

Table 1

Item Parameters of NAEP Writing Items:
1988 & 1990 Age 9

Item	Slope	Location	Category	Item-Category
1	0.370	-1.362	4.015	-5.377
			-4.835	3.473
			0.819	-2.181
2	0.617	-0.288	2.744	-3.032
			-0.747	0.459
			-1.997	1.709
3	0.507	1.053	4.628	-3.575
			2.083	-1.030
			-1.053	2.106
			-5.659	6.712
4	0.706	0.595	4.059	-3.464
			0.983	-0.388
			-1.351	1.946
			-3.691	4.286
5	0.969	-1.114	1.844	-2.958
			-0.252	-0.862
			-1.592	0.478
6	0.981	0.522	2.705	-2.183
			0.677	-0.156
			-0.441	0.963
			-2.942	3.463

Figure 1. Item-Category Characteristic Curves of the Partial Credit Model.

Item 1: $\underline{b}_1 = (0., 2., -4.)$

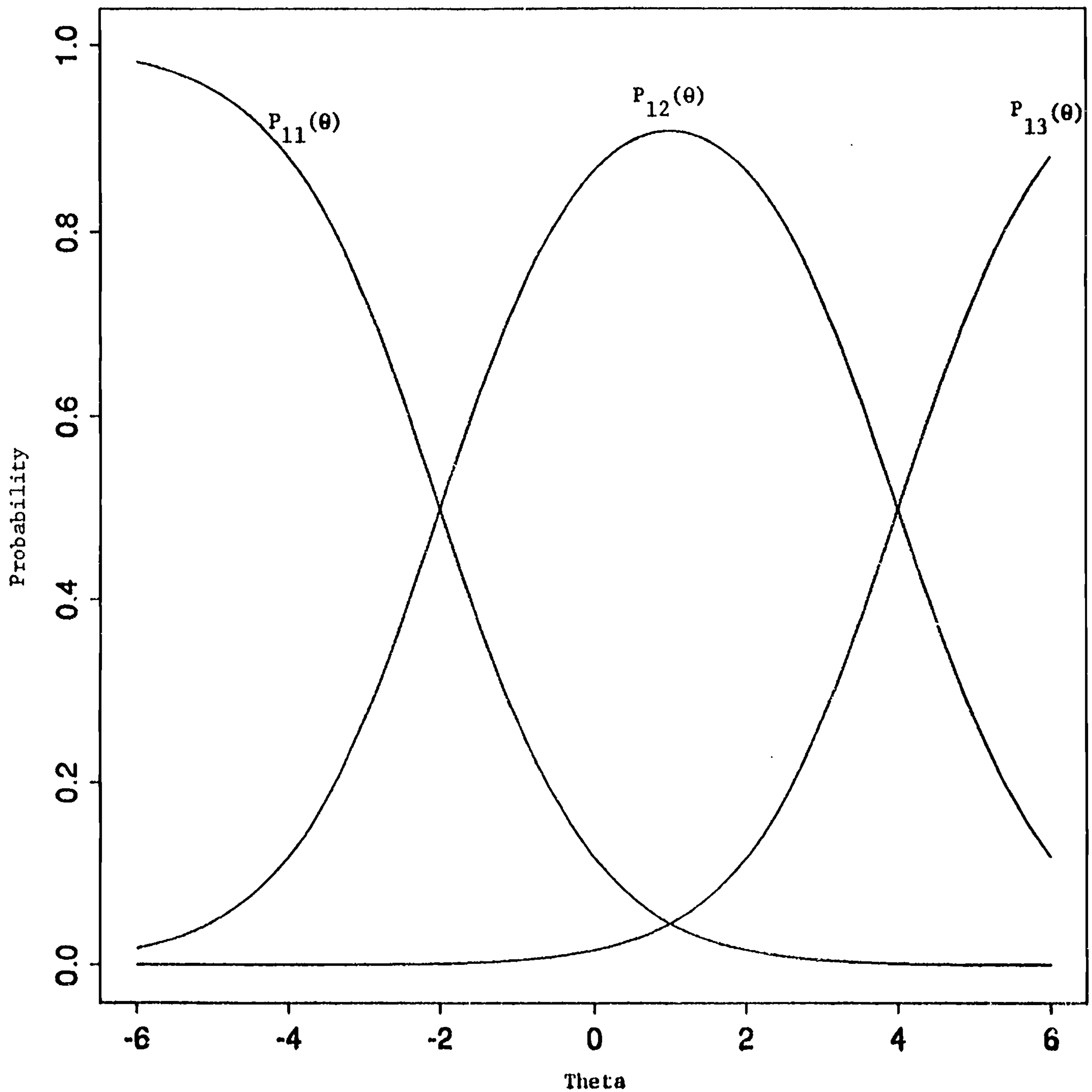

Figure 2. Item-Category Characteristic Curves of the Partial Credit Model.

Item 2: $\underline{b}_2$ = (0., 2., -2.)

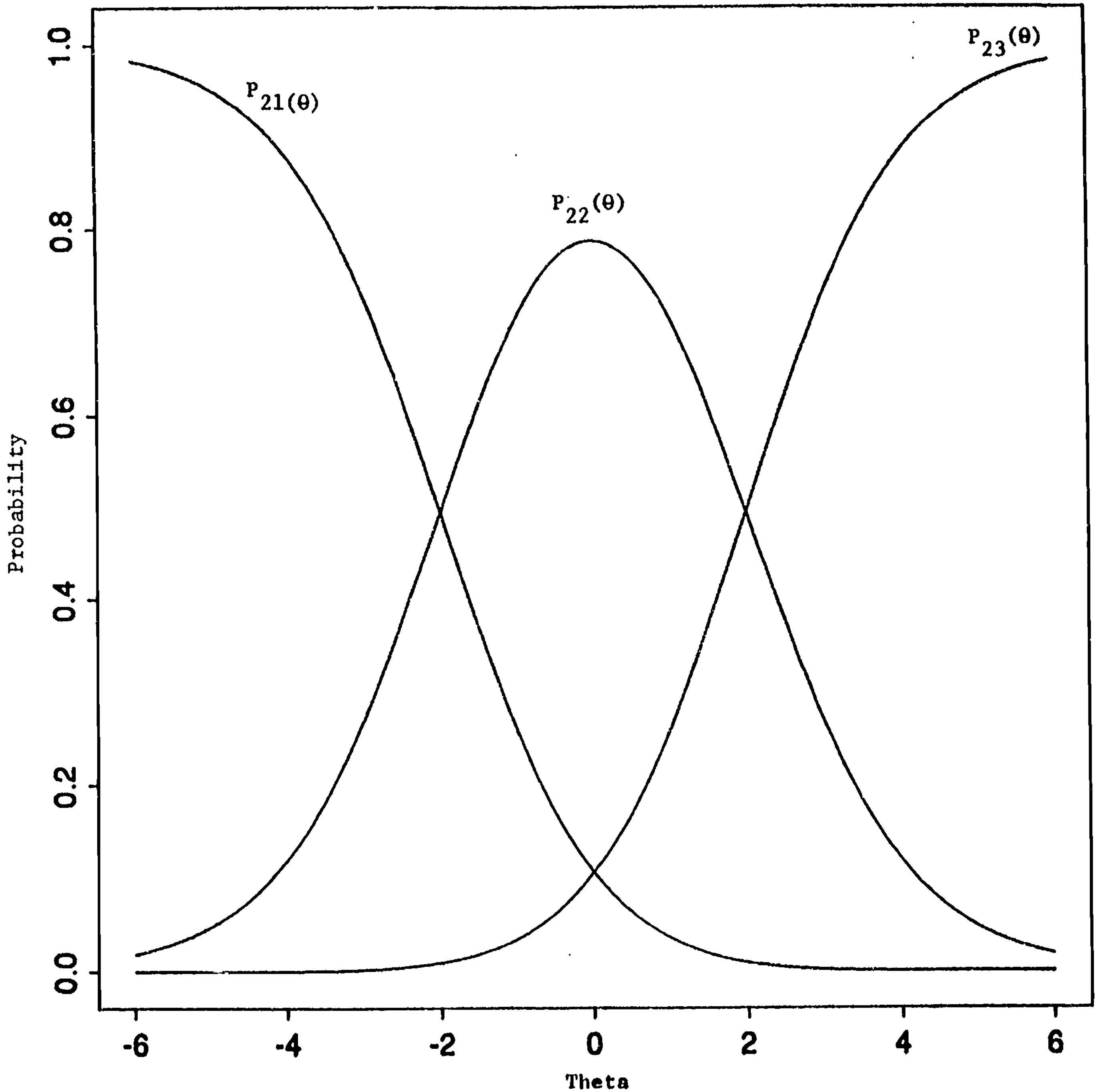

Figure 3. Item-Category Characteristic Curves of the Partial Credit Model.

Item 3: $\underline{b}_3(0., 2., 0.)$

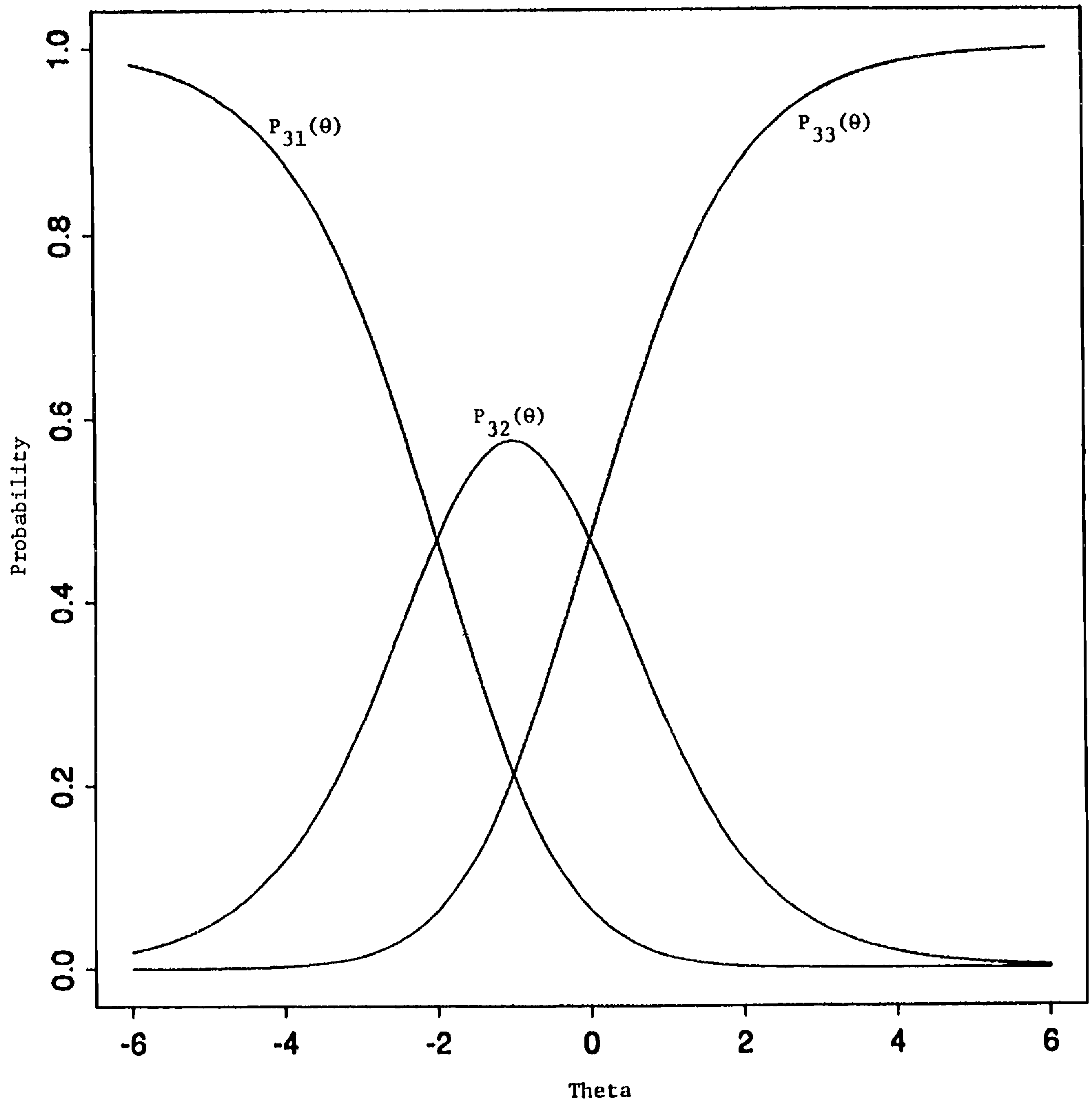

Figure 4. Item-Category Characteristic Curves of the Partial Credit Model.

Item 4: $\underline{b}_4$ = (0., 2., 2.)

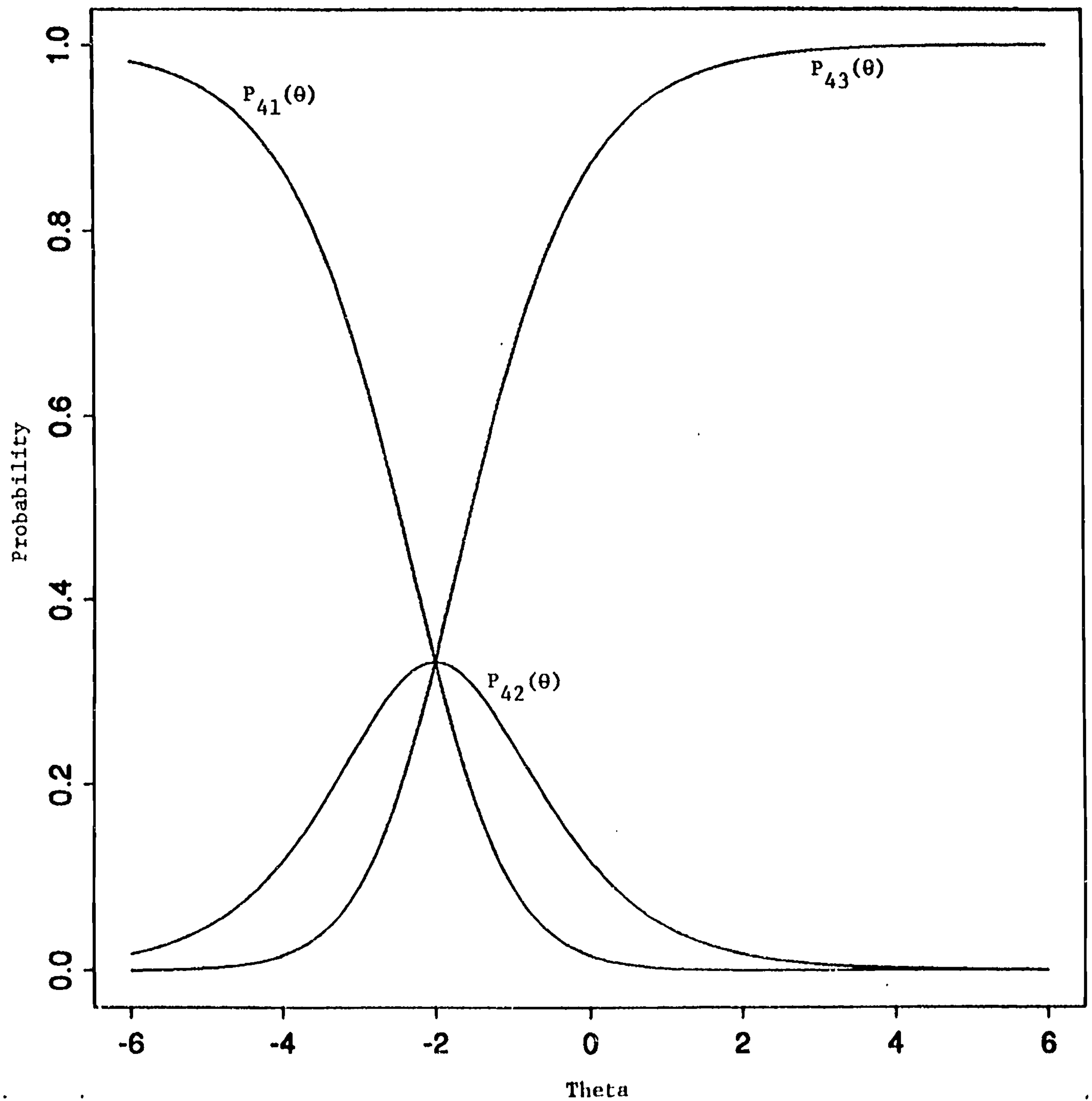

Figure 5. Item-Category Characteristic Curves of the Partial Credit Model.

Item 5: $\underset{\sim}{b}_5$ = (0., 2., 4.)

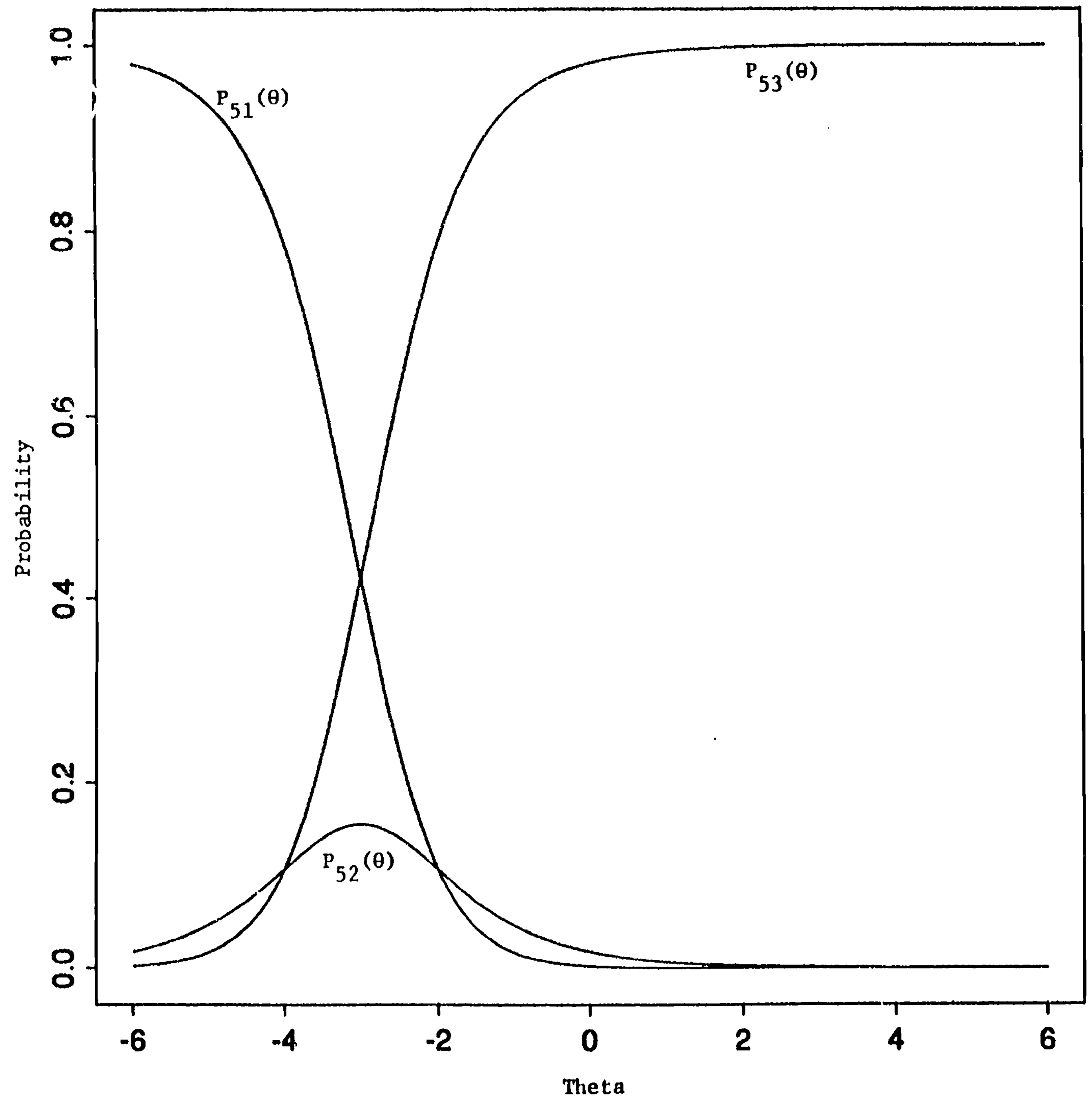

Figure 6. Item-Information Function and Item-Category Information Function.

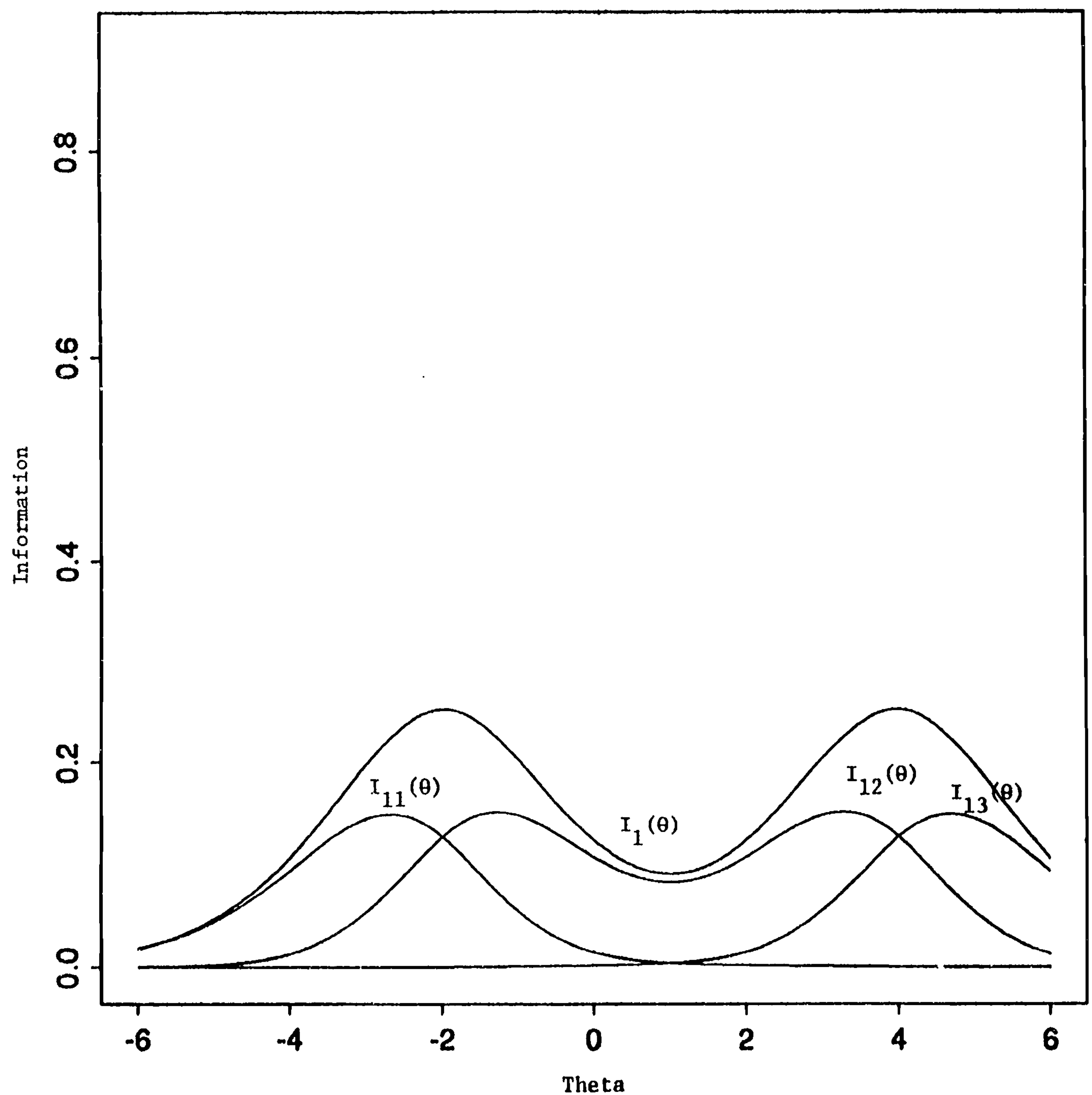

Figure 7. Item-Information Function and Item-Category Information Function.

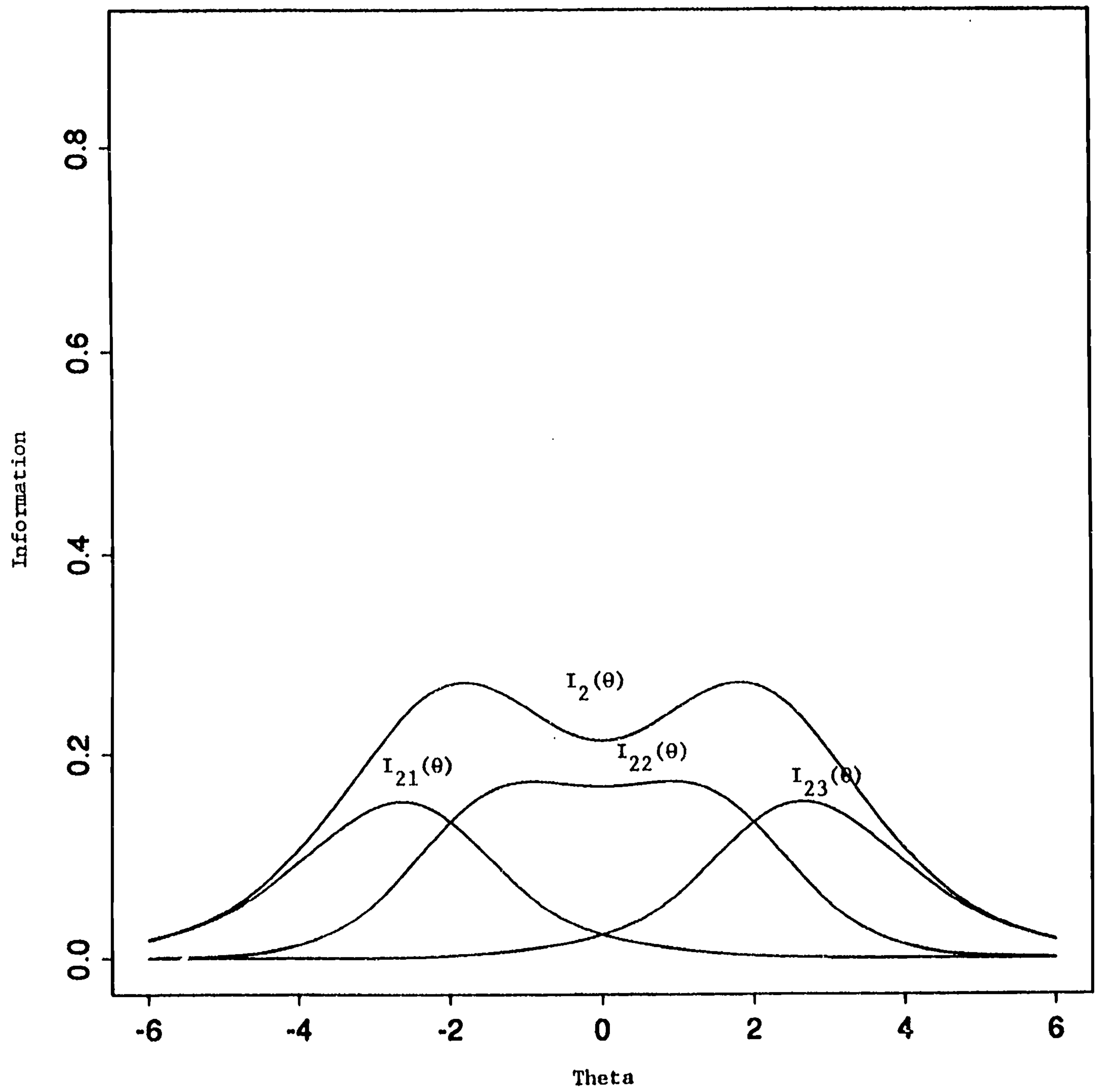

Figure 8. Item-Information Function and Item Category Information Function.

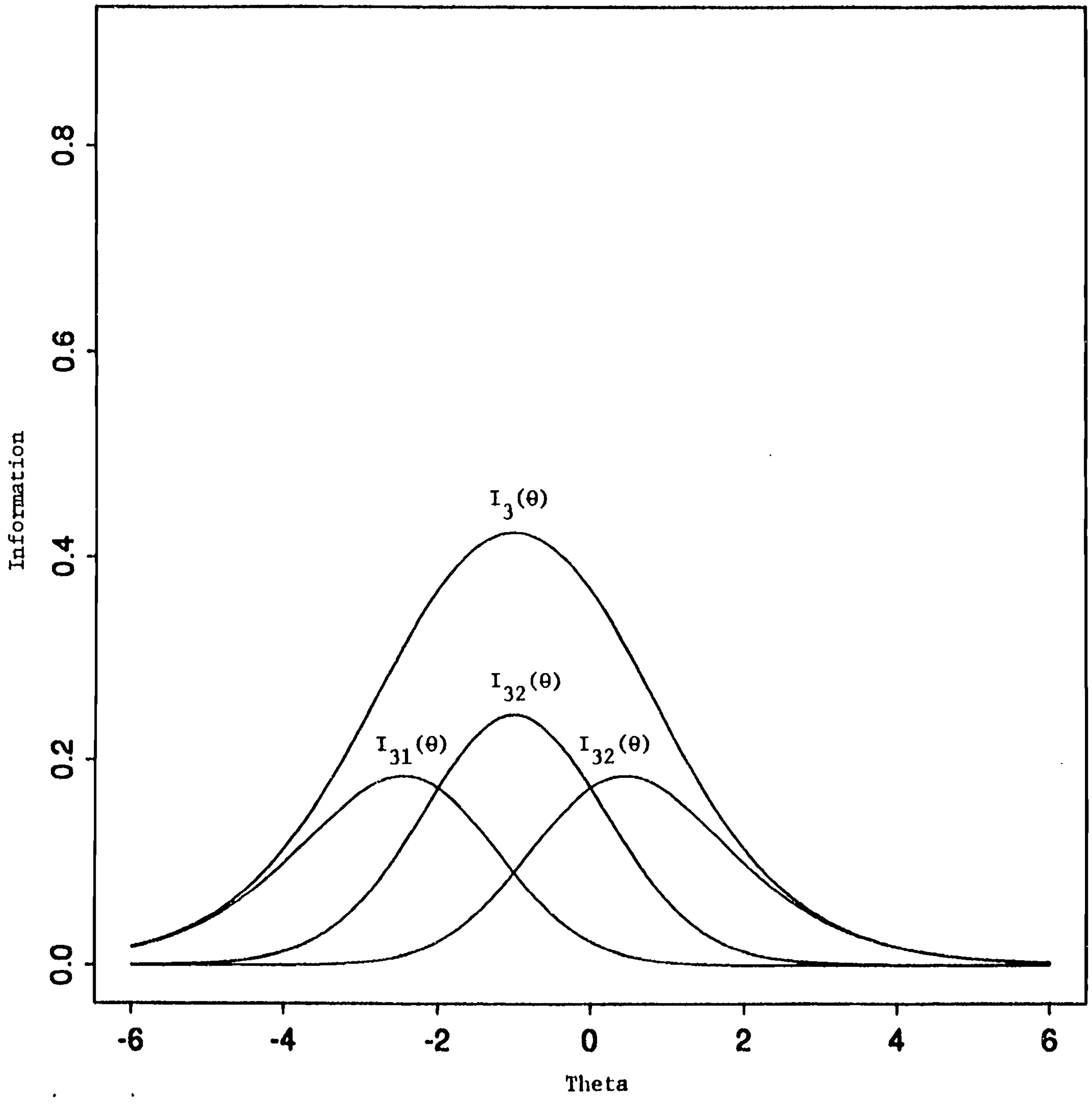

Figure 9. Item-Information Function and Item Category Information Function.

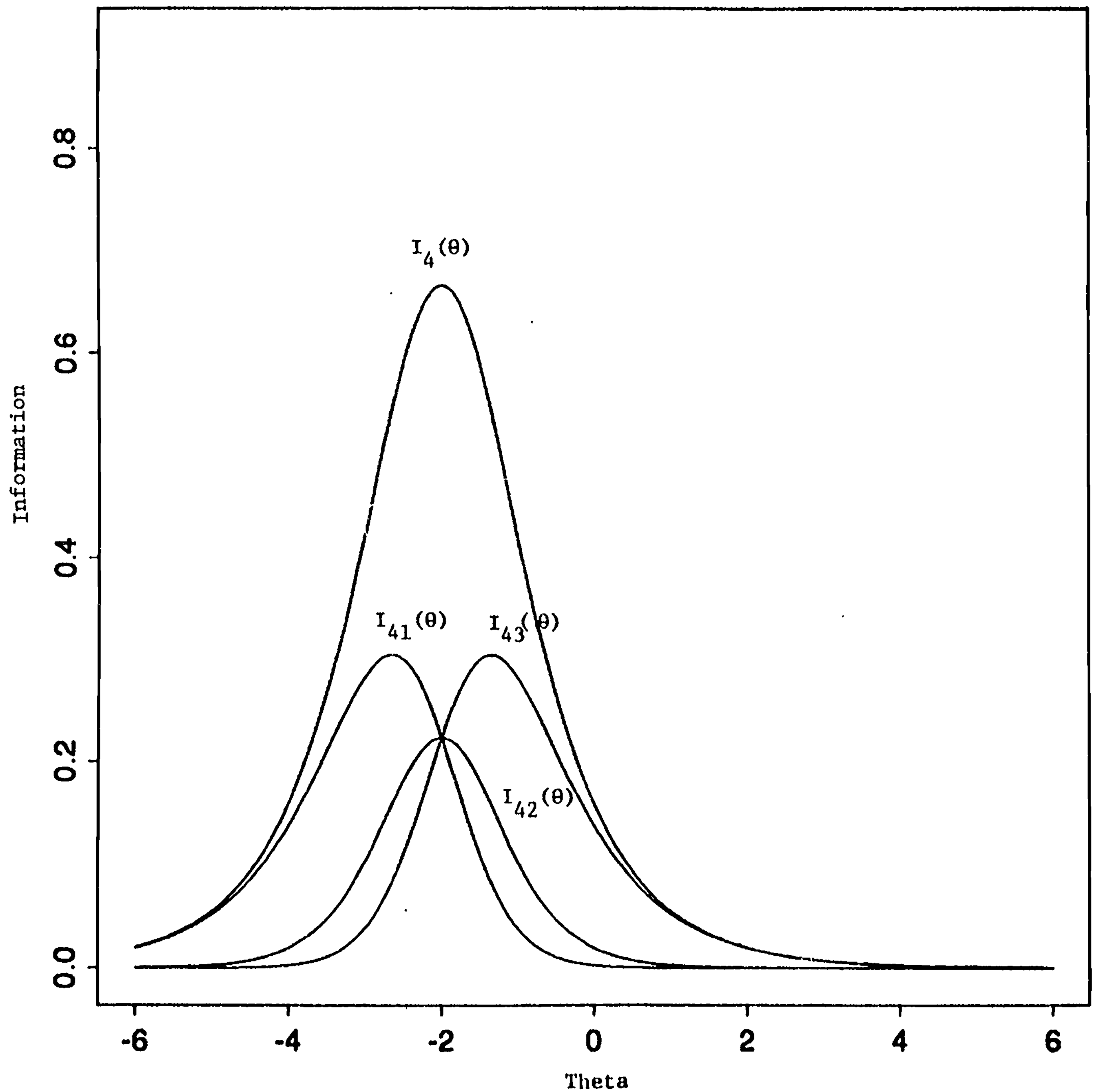

Figure 10. Item-Category Characteristic Curves of the Partial Credit Model.

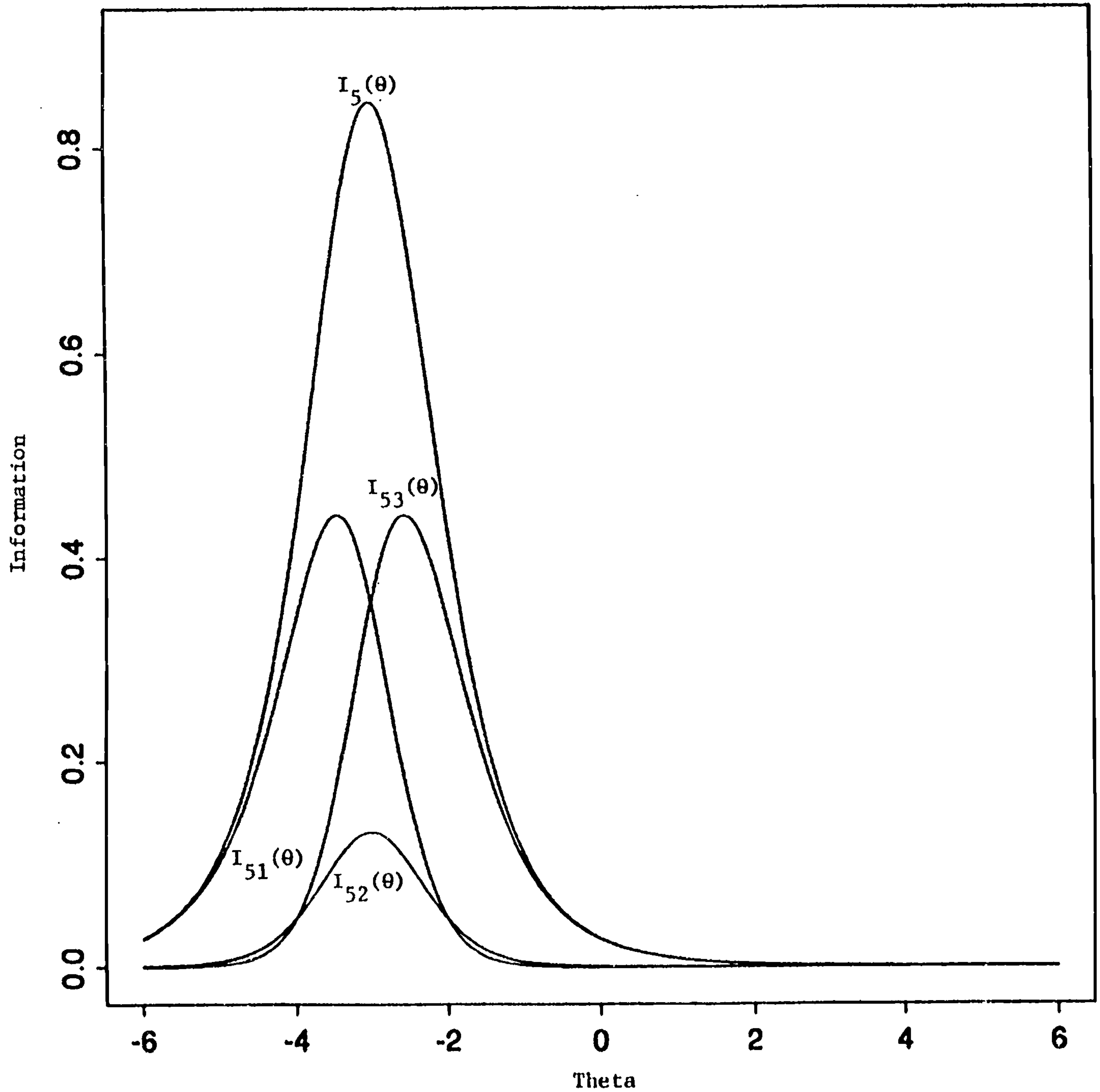

Figure 11. Expected Score Functions of Five Items.

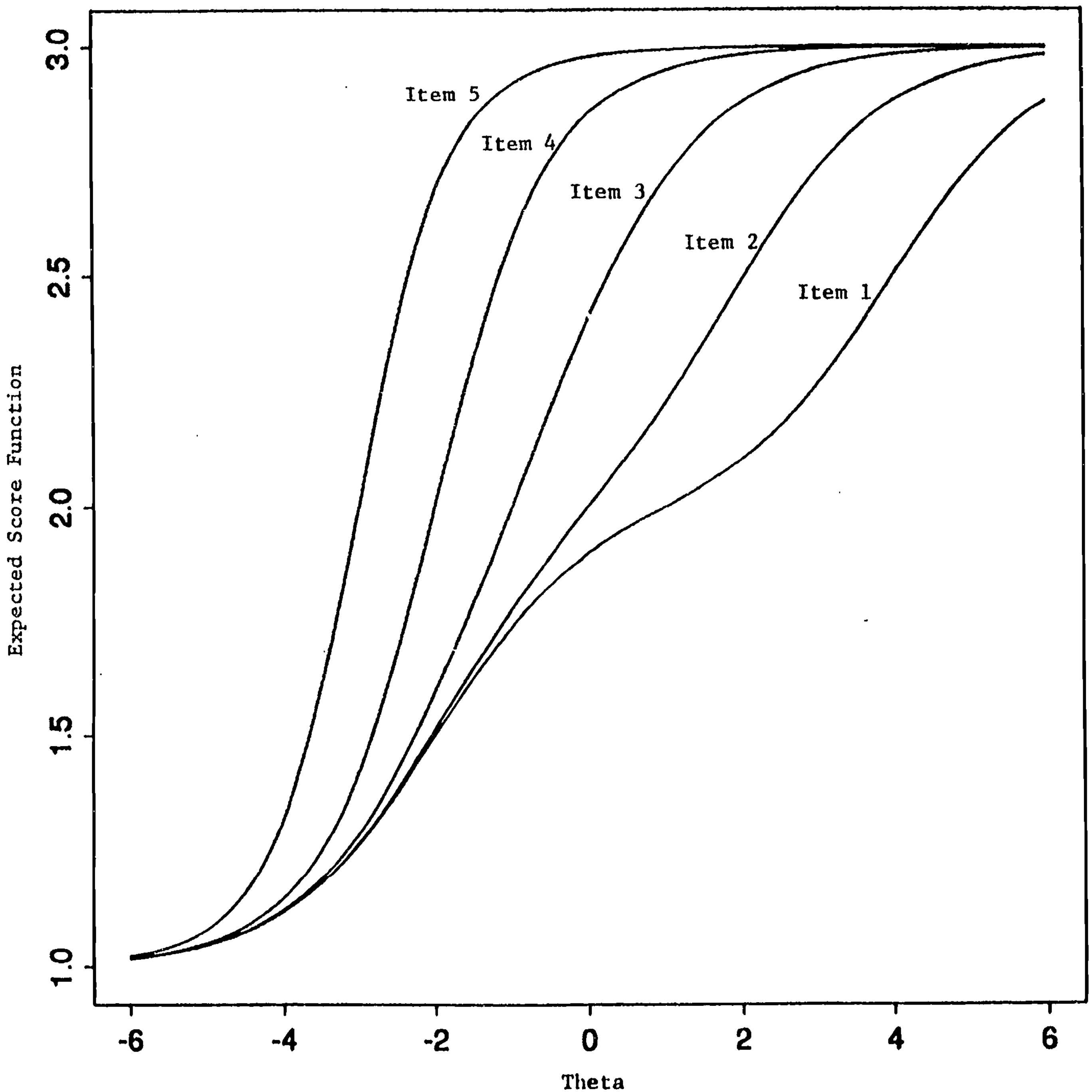

Figure 12

Item Information of Math Item

T7=(1,2,3,4,5,6,7), T6=(1,2,3,4,5,6,6), T5=(1,2,3,4,5,5,5)
T4=(1,2,3,4,4,4,4), T3=(1,2,3,3,3,3,3), T2=(1,2,2,2,2,2,2)

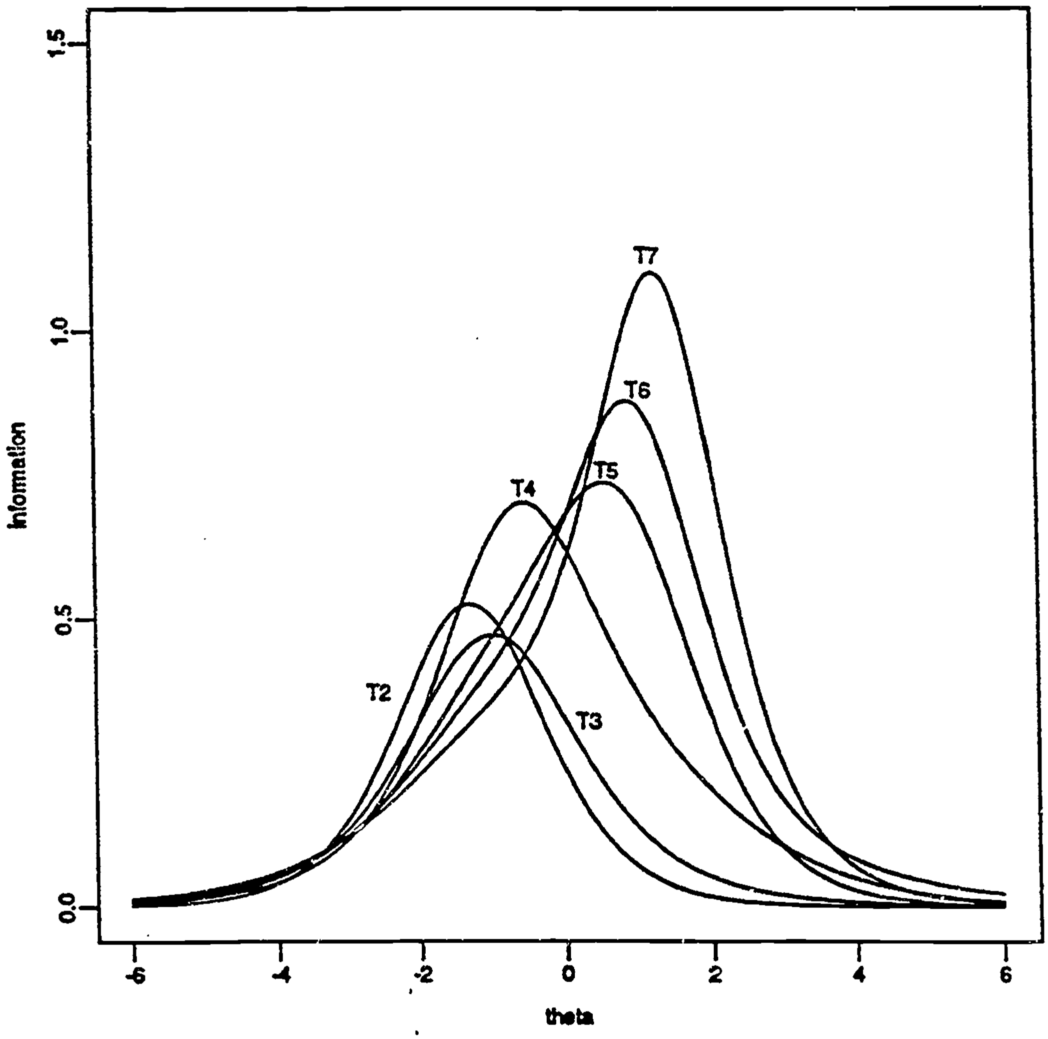

Figure 13

Item-Category Information : Life Science Item

Before Collapsing Categories - T=(1,2,3,4,5,6)

a=0.17, b=(0.0,-0.66,-1.72,0.31,1.93,-0.18)

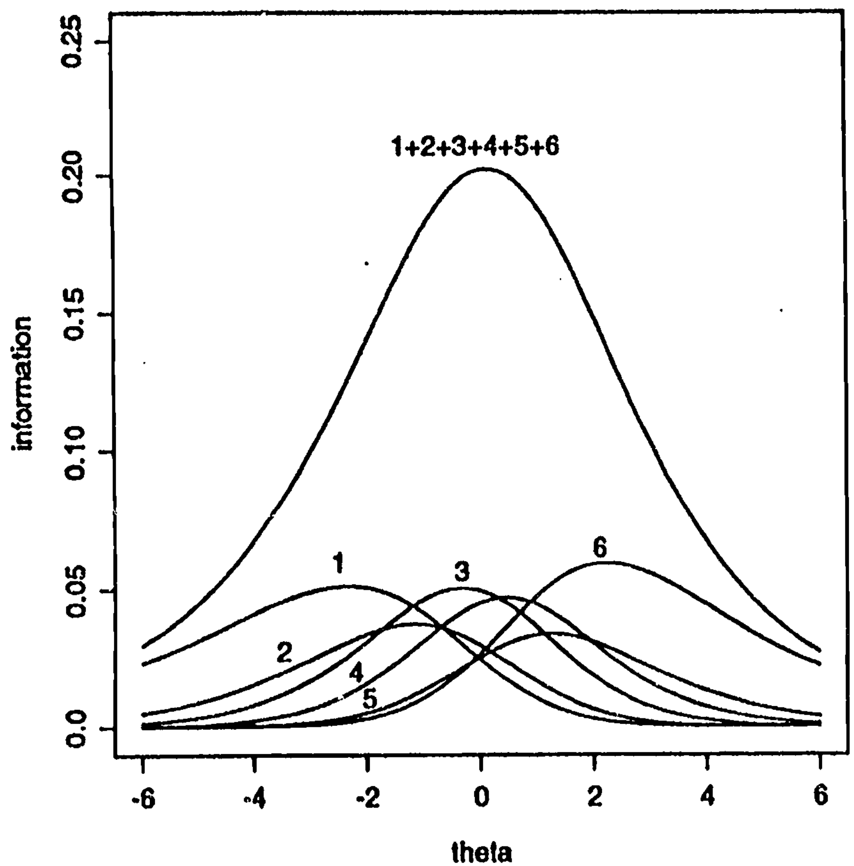

Figure 14

Item-Category Information : Life Science Item

After Collapsing Categories - T=(1,1,2,2,3,4)

a=0.28, b=(0.0,-1.20,2.76,0.21)

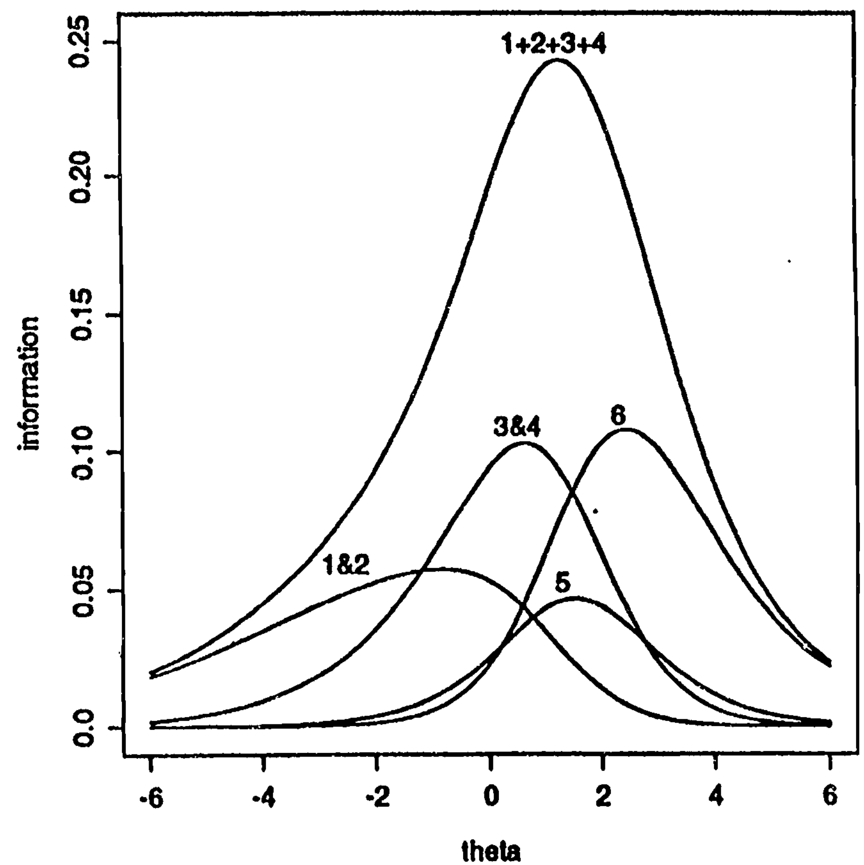

Figure 15

Item-Category Information : Physical Science Item

Before Recoding Categories - T=(1,2,3,4)

a=0.17, b=(0.0,-2.11,-1.76,5.19)

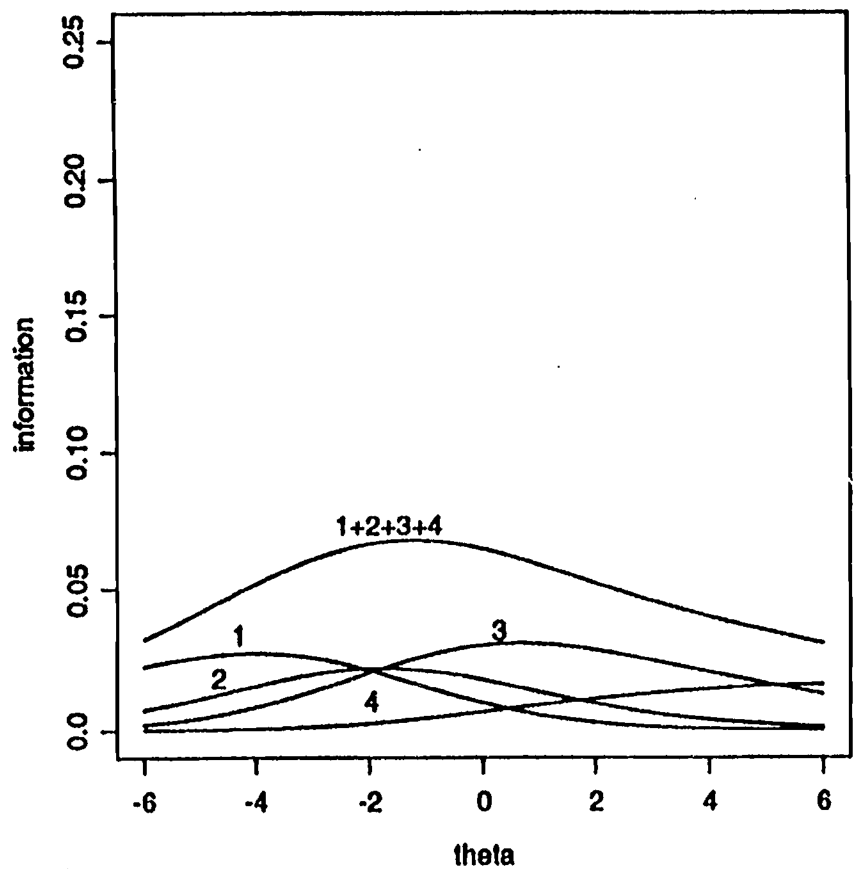

Figure 16

Item-Category Information : Physical Science Item

After Recoding Categories - T=(1,3,2,4)

a=0.32, b=(0.0,-1.56,0.60,2.95)

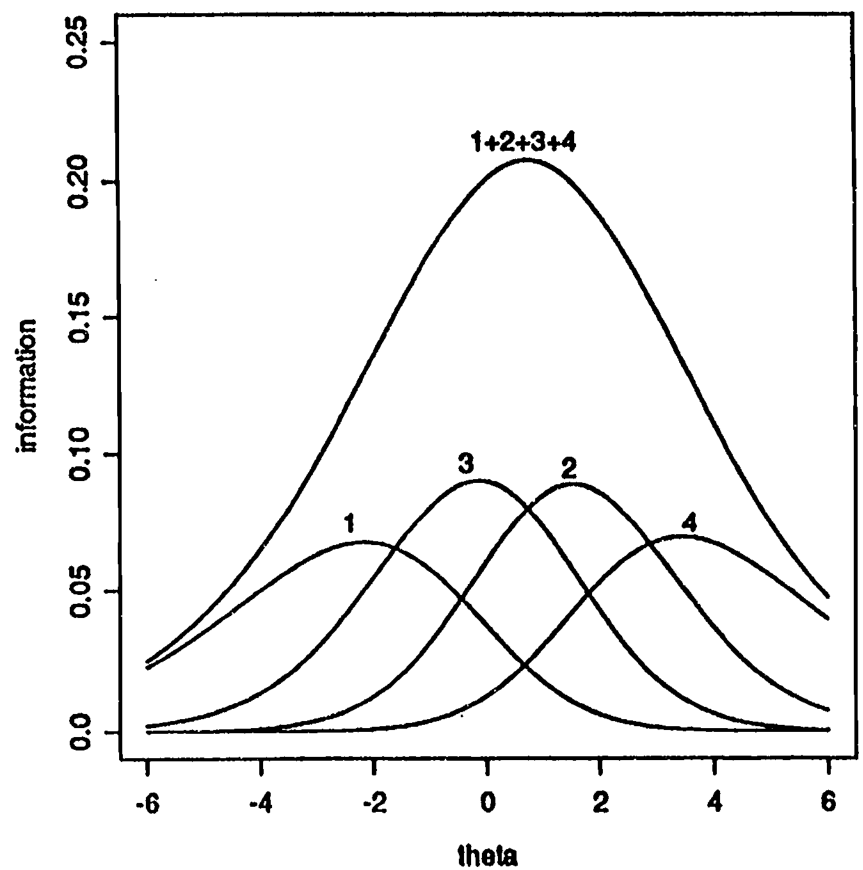

Figure 17 Item Information : Item 1 T=(1,2,3,4)

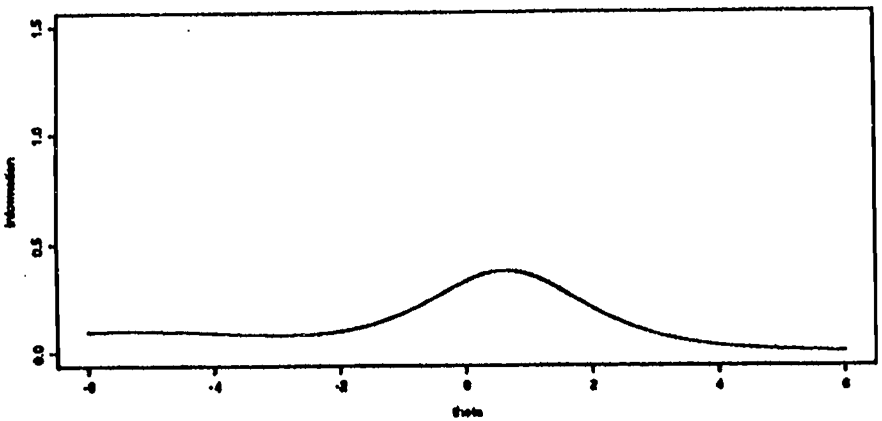

Figure 18 Item Information : Item 2 T=(1,2,3,4)

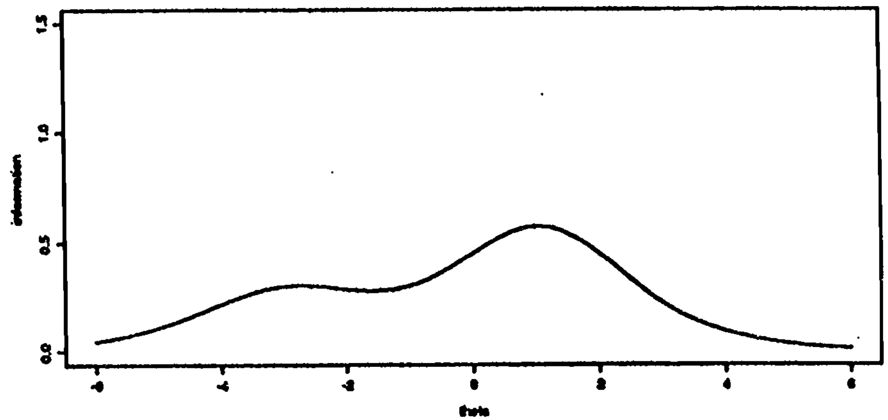

Figure 19 Item Information : Item 3 T=(1,2,3,4,5)

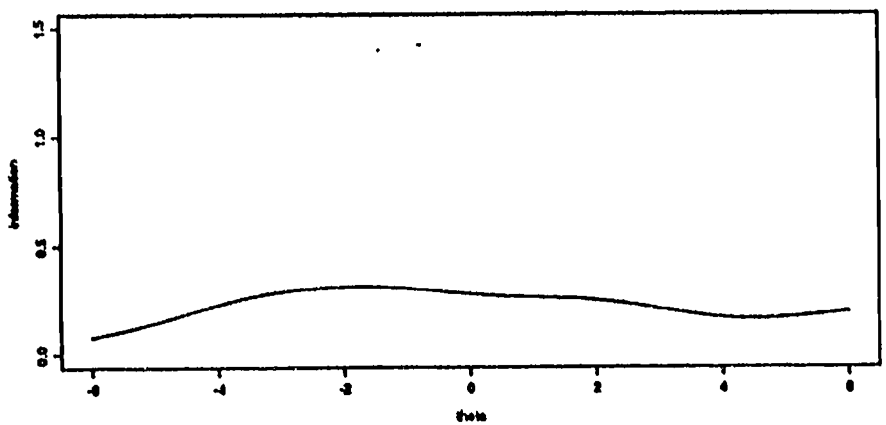

Figure 20 Item Information : Item 4 T=(1,2,3,4,5)

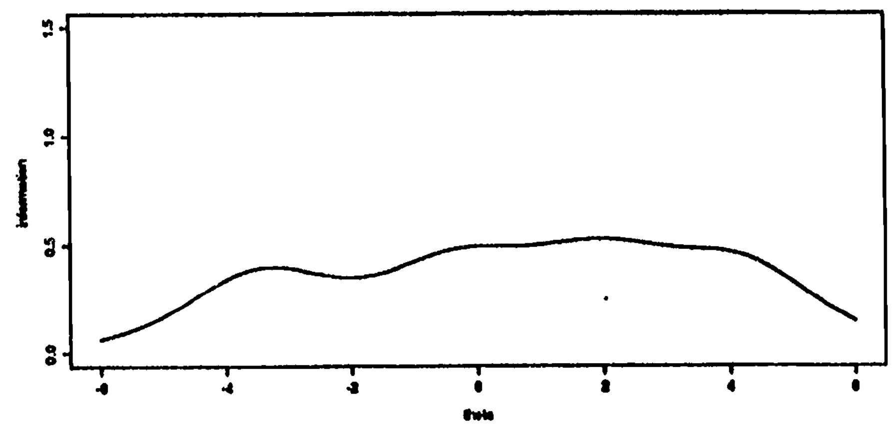

Figure 21 Item Information : Item 5 T=(1,2,3,4)

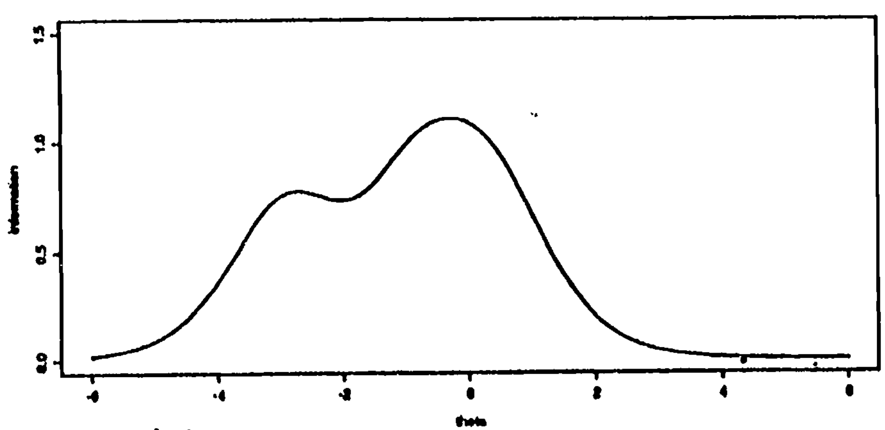

Figure 22 Item Information : Item 6 T=(1,2,3,4,5)

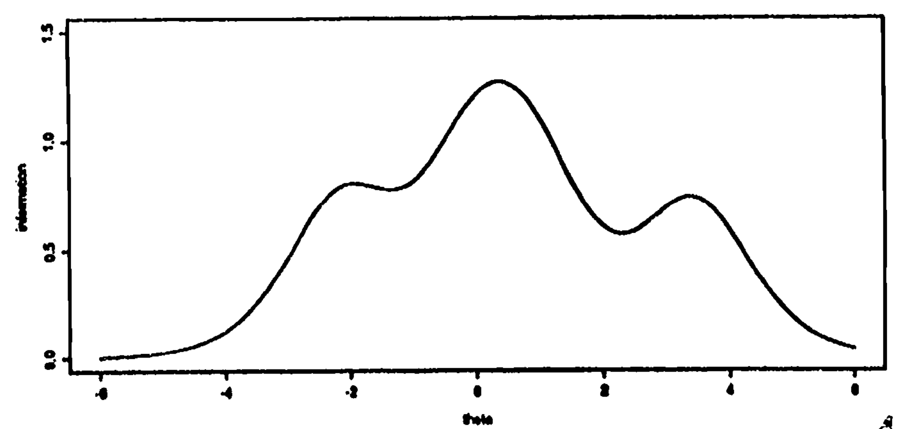

Figure 23

Test Information of the NAEP Writing Items