

AD-A183 811

BIAS FUNCTION OF THE MAXIMUM LIKELIHOOD ESTIMATE OF  
ABILITY FOR DISCRETE (U) TENNESSEE UNIV KNOXVILLE DEPT  
OF PSYCHOLOGY F SAMEJIMA 16 JUN 87 RR-87-1-ONR  
N00014-81-C-0569

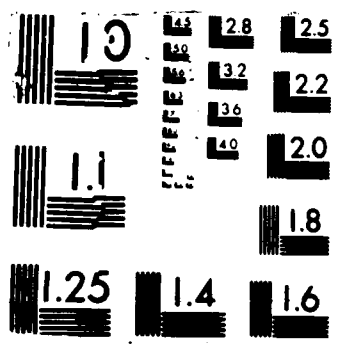
1/1

UNCLASSIFIED

F/G 12/3

NL


EVD  
1-27  
110



DTIC FILE COPY

2

# BIAS FUNCTION OF THE MAXIMUM LIKELIHOOD ESTIMATE OF ABILITY FOR DISCRETE ITEM RESPONSES

AD-A183 811

FUMIKO SAMEJIMA

UNIVERSITY OF TENNESSEE

KNOXVILLE, TENN. 37996-0900

JUNE, 1987

DTIC  
S AUG 07 1987 D  
A E

Prepared under the contract number N00014-81-C-0569  
NR 150-467 with the  
Personnel and Training Research Programs  
Psychological Sciences Division  
Office of Naval Research

Approved for public release; distribution unlimited.  
Reproduction in whole or in part is permitted for  
any purpose of the United States Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ONR/Research Report 87-1	2. GOVT ACCESSION NO. <b>AD-A183811</b>	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  Bias Function of the Maximum Likelihood Estimate of Ability for Discrete Item Responses	5. TYPE OF REPORT & PERIOD COVERED Technical Report	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s)  Dr. Fumiko Samejima	8. CONTRACT OR GRANT NUMBER(s)  N00014-81-C-0569	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research Arlington, VA 22217	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE: 61153N; PROJ: RR-042-04 TA: RR 042-04-01 WT: NR 150-467	
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE June 16, 1987	
	13. NUMBER OF PAGES 64	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States government.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Latent Trait Models Maximum Likelihood Estimation Bias		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  (Please see reverse side)		

The paper is concerned with the bias of the maximum likelihood estimate as a function of the latent trait, or ability, for the general case in which item responses are discrete. The rationale is presented, and observations are made with respect to the effects of the test information function, the item parameters, the number of items, the transformation of the latent variable, etc., on the amount of bias.

*mathematical modeling, psychometric, etc.*

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A1	

DTIC  
COPY  
INSPECTED  
6

## TABLE OF CONTENTS

	Page
<b>Abstract</b>	
<b>1 Introduction</b>	<b>1</b>
<b>2 Rationale</b>	<b>2</b>
<b>3 Dichotomous Response Level</b>	<b>8</b>
3.1 Normal Ogive Model . . . . .	8
3.2 Logistic Model . . . . .	9
3.3 Rasch Model . . . . .	18
3.4 Three-Parameter Logistic Model . . . . .	18
<b>4 Graded Response Level</b>	<b>30</b>
<b>5 Effect of the Discrimination Parameter</b>	<b>34</b>
<b>6 Effect of the Number of Items</b>	<b>43</b>
<b>7 Scale Transformation</b>	<b>44</b>
7.1 General Case of Discrete Responses . . . . .	44
7.2 Scale Transformation to Generate a Constant Test Information . . . . .	47
7.3 Equivalent Items on the Dichotomous Response Level . . . . .	53
<b>8 Adaptive Testing</b>	<b>57</b>
<b>9 Discussion and Conclusions</b>	<b>57</b>
<b>References</b>	<b>58</b>

---

The research was conducted at the principal investigator's laboratory, 405 Austin Peay Bldg., Department of Psychology, University of Tennessee, Knoxville, Tennessee. Those who worked as assistants for this research include David M. Immel, Richard D. Strouse, Christine A. Golik, J. Douglas Jenkins, Stacey M. Miller, Sherry C. Campbell, and Robert S. Hiers.

# 1 Introduction

Lord has proposed and discussed a bias function of the maximum likelihood estimate in the context of the three-parameter logistic model (cf. Lord, 1983). In so doing, he used Taylor's expansion of the likelihood equation and proceeded from there, obtained an equation which includes the conditional expectation of the discrepancy between the maximum likelihood estimate and the true ability, and ignored all terms of orders higher than  $n^{-1}$ , where  $n$  indicated the number of items.

Let  $\theta$  be ability, or latent trait, which assumes any real number. Let  $g$  ( $= 1, 2, \dots, n$ ) denote an item,  $k_g$  be a discrete response to item  $g$ , and  $P_{k_g}(\theta)$  denote the operating characteristic of the discrete response  $k_g$ , or the conditional probability, given  $\theta$ , with which the examinee responds to item  $g$  with  $k_g$ . The item response information function,  $I_{k_g}(\theta)$ , is defined by

$$(1.1) \quad I_{k_g}(\theta) = -\frac{\partial^2}{\partial \theta^2} \log P_{k_g}(\theta) ,$$

and the item information function  $I_g(\theta)$  is the conditional expectation of the item response information function, given  $\theta$ , so that we can write

$$(1.2) \quad I_g(\theta) = E[I_{k_g}(\theta) | \theta] = \sum_{k_g} I_{k_g}(\theta) P_{k_g}(\theta) .$$

On the dichotomous response level (Samejima, 1972), the set of operating characteristics for item  $g$  is represented by a single function, i.e., the operating characteristic of the positive response, which is called the item characteristic function, or item response function<sup>1</sup>.

Let  $P_g(\theta)$  be the item characteristic function in the three-parameter logistic model, which is given by,

$$(1.3) \quad P_g(\theta) = c_g + (1 - c_g)[1 + \exp\{-Da_g(\theta - b_g)\}]^{-1} ,$$

where  $a_g$ ,  $b_g$  and  $c_g$  are the item discrimination, difficulty and guessing parameters, and  $D$  is a scaling factor, which is set equal to 1.7 when the logistic model is used as a substitute for the normal ogive model. Lord's bias function  $B(\theta)$  can be written as

$$(1.4) \quad B(\theta) = D[I(\theta)]^{-2} \sum_{g=1}^n a_g I_g(\theta) [\Psi_g(\theta) - (1/2)] ,$$

where

$$(1.5) \quad \Psi_g(\theta) = [1 + \exp\{-Da_g(\theta - b_g)\}]^{-1} ,$$

and  $I_g(\theta)$  and  $I(\theta)$  are the item information function and the test information function, respectively, which can be given by

$$(1.6) \quad I_g(\theta) = [P'_g(\theta)]^2 [P_g(\theta)\{1 - P_g(\theta)\}]^{-1} ,$$

<sup>1</sup>The term, item response function, has been widely used in recent years by researchers who deal solely with the dichotomous response level. From the more comprehensive standpoint, however, this term is ambiguous and misleading, and not appropriate to use. On the graded response level, for example, there may be much more than two item response categories, or there may even be an infinite number of response categories, and the use of item response function for one of these many response categories is not justifiable. For this reason, throughout this paper, the original term, item characteristic function, will be used to indicate the conditional probability for the positive response, given latent trait, on the dichotomous response level.

and

$$(1.7) \quad I(\theta) = \sum_{g=1}^n I_g(\theta) ,$$

with  $P'_g(\theta)$  indicating the first derivative of  $P_g(\theta)$  with respect to  $\theta$ . The former of these two formulae can be given as a special case of the item information function given by (1.2), which is defined for the general case of discrete responses. (Incidentally, in Lord's paper,  $B_1(\hat{\theta})$  is used for this bias function. This is not appropriate, however, since it is a function of  $\theta$  itself, not of its maximum likelihood estimate  $\hat{\theta}$ .)

## 2 Rationale

A similar logic can be adopted for the general case, in which item responses are simply discrete. We assume that there are a finite or an enumerable number of  $k_g$ 's as possible responses to item  $g$ . Thus for the set of  $n$  items, we can write for the response pattern  $V$

$$(2.1) \quad V' = (k_1, k_2, \dots, k_g, \dots, k_n) .$$

We assume that the operating characteristic  $P_{k_g}(\theta)$  is, at least, three-times differentiable with respect to  $\theta$ . By virtue of local independence, we can write for the likelihood function

$$(2.2) \quad L_V(\theta) = P_V(\theta) = \prod_{k_g \in V} P_{k_g}(\theta) .$$

Thus the likelihood equation is given by

$$(2.3) \quad \frac{\partial}{\partial \theta} \log L_V(\theta) = \sum_{k_g \in V} \frac{\partial}{\partial \theta} \log P_{k_g}(\theta) \equiv 0 .$$

We define  $\Gamma_{sk_g}(\theta)$  such that

$$(2.4) \quad \Gamma_{sk_g}(\theta) = \frac{\partial^s}{\partial \theta^s} \log P_{k_g}(\theta)$$

for  $s = 1, 2, \dots$ . We notice, in particular, that

$$(2.5) \quad \Gamma_{1k_g}(\theta) = P'_{k_g}(\theta)[P_{k_g}(\theta)]^{-1} = A_{k_g}(\theta) ,$$

where  $A_{k_g}(\theta)$  is the basic function (Samejima, 1969), and

$$(2.6) \quad \Gamma_{2k_g}(\theta) = P''_{k_g}(\theta)[P_{k_g}(\theta)]^{-1} - [A_{k_g}(\theta)]^2$$

and

$$(2.7) \quad \Gamma_{3k_g}(\theta) = P'''_{k_g}(\theta)[P_{k_g}(\theta)]^{-1} - 3A_{k_g}(\theta)P''_{k_g}(\theta)[P_{k_g}(\theta)]^{-1} + 2[A_{k_g}(\theta)]^3 ,$$



where the superscripts ' , " and "' indicate the first, second and third partial derivatives of the function with respect to  $\theta$  , respectively. Thus from (2.3) and (2.5) we can write

$$(2.8) \quad \sum_{k_g \in V} \Gamma_{1k_g}(\hat{\theta}_V) = \sum_{k_g \in V} A_{k_g}(\hat{\theta}_V) = 0 .$$

Let  $\Gamma_{sV}(\theta)$  be defined by

$$(2.9) \quad \Gamma_{sV}(\theta) = \sum_{k_g \in V} \Gamma_{sk_g}(\theta)$$

for  $s = 1, 2, \dots$ . For a fixed value of  $\theta$  we can write by Taylor's formula

$$(2.10) \quad \begin{aligned} \Gamma_{1V}(\hat{\theta}_V) &= \Gamma_{1V}(\theta) + (\hat{\theta}_V - \theta)\Gamma_{2V}(\theta) + (1/2)(\hat{\theta}_V - \theta)^2\Gamma_{3V}(\theta) \\ &+ (1/6)(\hat{\theta}_V - \theta)^3\Gamma_{4V}(\theta) + (1/24)(\hat{\theta}_V - \theta)^4\Gamma_{5V}(\xi) = 0 . \end{aligned}$$

where  $\xi$  is some value between  $\theta$  and  $\hat{\theta}_V$  .

Since we have

$$(2.11) \quad \sum_{k_g} P_{k_g}(\theta) = 1 ,$$

we obtain

$$(2.12) \quad \sum_{k_g} \frac{\partial^s}{\partial \theta^s} P_{k_g}(\theta) = 0$$

for  $s = 1, 2, \dots$ . Equation (2.12) will be helpful in following the mathematical derivations which are needed in obtaining the bias function. The response pattern information function,  $I_V(\theta)$  , is defined by

$$(2.13) \quad I_V(\theta) = -\frac{\partial^2}{\partial \theta^2} \log P_V(\theta) = \sum_{k_g \in V} I_{k_g}(\theta) ,$$

and the test information function  $I(\theta)$  is the conditional expectation of  $I_V(\theta)$  , given  $\theta$  , for which we can write

$$(2.14) \quad I(\theta) = E[I_V(\theta) | \theta] = \sum_V I_V(\theta) P_V(\theta) .$$

Let  $f_{k_g}(\theta)$  be any function of  $\theta$  defined for a specific discrete response  $k_g$  . We can write

$$(2.15) \quad \begin{aligned} \sum_V \sum_{k_g \in V} f_{k_g}(\theta) P_V(\theta) &= \sum_V \sum_{k_g \in V} f_{k_g}(\theta) P_{k_g}(\theta) P_{V-k_g}(\theta) \\ &= \sum_{g=1}^n \sum_{k_g} f_{k_g}(\theta) P_{k_g}(\theta) , \end{aligned}$$

by virtue of the fact that

$$(2.16) \quad \sum_{V_{-g}} P_{V_{-g}}(\theta) = 1$$

where  $V_{-g}$  is the response pattern of  $(n-1)$  discrete item scores obtained by deleting  $k_g$  from  $V$ . Replacing  $f_{k_g}(\theta)$  by  $I_{k_g}(\theta)$  in (2.15) and using this result, (1.2), (2.13) and (2.14), we can obtain the same equation as (1.7).

Let  $\gamma_{sg}(\theta)$  be the conditional expectation of  $\Gamma_{sk_g}(\theta)$ , given  $\theta$ , which can be written as

$$(2.17) \quad \gamma_{sg}(\theta) = E[\Gamma_{sk_g}(\theta) | \theta] = \sum_{k_g} \Gamma_{sk_g}(\theta) P_{k_g}(\theta) .$$

In particular, we have from (2.5), (2.6), (2.7) and (2.12)

$$(2.18) \quad \gamma_{1g} = \sum_{k_g} P'_{k_g}(\theta) = 0 ,$$

$$(2.19) \quad \gamma_{2g}(\theta) = - \sum_{k_g} [P'_{k_g}(\theta)]^2 P_{k_g}(\theta)^{-1}$$

and

$$(2.20) \quad \gamma_{3g}(\theta) = 2 \sum_{k_g} [A_{k_g}(\theta)]^2 P'_{k_g}(\theta) - 3 \sum_{k_g} A_{k_g}(\theta) P''_{k_g}(\theta) .$$

It is noted from (1.1), (1.2), (2.12) and (2.17) that we can also write

$$(2.21) \quad \gamma_{2g}(\theta) = -I_g(\theta) .$$

We further define  $\gamma_s(\theta)$  such that

$$(2.22) \quad \gamma_s(\theta) = (1/n) \sum_{g=1}^n \gamma_{sg}$$

for  $s = 1, 2, \dots$ . In particular, we have

$$(2.23) \quad \gamma_1(\theta) = 0 ,$$

$$(2.24) \quad \gamma_2(\theta) = -(1/n) \sum_{g=1}^n I_g(\theta) = -(1/n) I(\theta)$$

and

$$(2.25) \quad \gamma_3(\theta) = (2/n) \sum_{g=1}^n \sum_{k_g} [A_{k_g}(\theta)]^2 P'_{k_g}(\theta) - (3/n) \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta) P''_{k_g}(\theta) .$$

Let  $\epsilon_{sk_g}(\theta)$  and  $\epsilon_{sV}(\theta)$  be

$$(2.26) \quad \epsilon_{sk_g}(\theta) = \Gamma_{sk_g}(\theta) - \gamma_{sg}(\theta)$$

and

$$(2.27) \quad \epsilon_{sV}(\theta) = (1/n) \sum_{k_g \in V} \epsilon_{sk_g}(\theta) ,$$

respectively. In particular, we can write from these definitions, (1.7), (2.5), (2.6), (2.18) and (2.21)

$$(2.28) \quad \epsilon_{1k_g}(\theta) = \Gamma_{1k_g}(\theta) = A_{k_g}(\theta) ,$$

$$(2.29) \quad \epsilon_{2k_g}(\theta) = P''_{k_g}(\theta)[P_{k_g}(\theta)]^{-1} - [A_{k_g}(\theta)]^2 + I_{\theta}(\theta) ,$$

$$(2.30) \quad \epsilon_{1V}(\theta) = (1/n) \sum_{k_g \in V} A_{k_g}(\theta)$$

and

$$(2.31) \quad \epsilon_{2V}(\theta) = (1/n) \sum_{k_g \in V} P''_{k_g}(\theta)[P_{k_g}(\theta)]^{-1} - (1/n) \sum_{k_g \in V} [A_{k_g}(\theta)]^2 + (1/n)I(\theta) .$$

We can also obtain from (2.15), (2.17), (2.22), (2.26) and (2.27) for the conditional expectation of  $\epsilon_{sV}(\theta)$ , given  $\theta$ ,

$$(2.32) \quad E[\epsilon_{sV}(\theta) | \theta] = \sum_V \epsilon_{sV}(\theta) P_V(\theta) = \gamma_s(\theta) - \gamma_s(\theta) = 0 .$$

With these definitions of  $\gamma_s(\theta)$  and  $\epsilon_{sV}(\theta)$  and from (2.10) we have

$$(2.33) \quad \begin{aligned} \epsilon_{1V}(\theta) + (\hat{\theta}_V - \theta)[\gamma_2(\theta) + \epsilon_{2V}(\theta)] + (1/2)(\hat{\theta}_V - \theta)^2[\gamma_3(\theta) + \epsilon_{3V}(\theta)] \\ + (1/6)(\hat{\theta}_V - \theta)^3[\gamma_4(\theta) + \epsilon_{4V}(\theta)] + (1/24)(\hat{\theta}_V - \theta)^4\Gamma_{5V}(\theta) \doteq 0 , \end{aligned}$$

and proceeding from here by taking the conditional expectation of each term in (2.33) with respect to  $V$ , given  $\theta$ , and ignoring all terms whose orders are higher than  $n^{-1}$ , we obtain

$$(2.34) \quad E[\epsilon_{1V}(\theta) | \theta] + \gamma_2(\theta)E[\hat{\theta}_V - \theta | \theta] + E[(\hat{\theta}_V - \theta)\epsilon_{2V}(\theta) | \theta] + (1/2)\gamma_3(\theta)E[(\hat{\theta}_V - \theta)^2 | \theta] \doteq 0 .$$

It is obvious from (2.32) that the first term on the left hand side of (2.34) disappears. As for the fourth and last term in (2.34), we can use the asymptotic variance of the distribution of the maximum likelihood estimate as the approximation to its last factor, i.e.,

$$(2.35) \quad E[(\hat{\theta}_V - \theta)^2 | \theta] \doteq [I(\theta)]^{-1} .$$

Since  $\gamma_2(\theta)$  and  $\gamma_3(\theta)$  are given by (2.24) and (2.25), respectively, all we need to do is to evaluate the third term on the left hand side of (2.34) in the general framework. In so doing we need to multiply (2.33) by  $\epsilon_{2V}(\theta)$ , take its expectation with respect to  $V$  and ignore all terms of  $o(n^{-1})$ , to obtain

$$(2.36) \quad E[(\hat{\theta}_V - \theta)\epsilon_{2V}(\theta) | \theta] = -[\gamma_2(\theta)]^{-1} E[\epsilon_{1V}(\theta)\epsilon_{2V}(\theta) | \theta] .$$

Thus the remaining task is to evaluate the second factor of the right hand side of (2.36). From (2.30) and (2.31) we have

$$\begin{aligned}
(2.37) \quad E[\epsilon_{1V}(\theta)\epsilon_{2V}(\theta) | \theta] &= \sum_V \epsilon_{1V}(\theta)\epsilon_{2V}(\theta)P_V(\theta) \\
&= (1/n^2) \sum_V \sum_{k_g \in V} A_{k_g}(\theta) \sum_{k_h \in V} P''_{k_h}(\theta)[P_{k_h}(\theta)]^{-1}P_V(\theta) \\
&\quad - (1/n^2) \sum_V \sum_{k_g \in V} A_{k_g}(\theta) \sum_{k_h \in V} [A_{k_h}(\theta)]^2 P_V(\theta) \\
&\quad + [(1/n) \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta)P'_{k_g}(\theta)]E[\epsilon_{1V}(\theta) | \theta] .
\end{aligned}$$

It is obvious from (2.32) that the third term on the right hand side of (2.37) disappears. We can write by virtue of (2.15)

$$\begin{aligned}
(2.38) \quad \sum_V \sum_{k_g \in V} A_{k_g}(\theta) \sum_{k_h \in V} P''_{k_h}(\theta)[P_{k_h}(\theta)]^{-1}P_V(\theta) \\
= \sum_V \sum_{k_g \in V} A_{k_g}(\theta)P''_{k_g}(\theta)[P_{k_g}(\theta)]^{-1}P_V(\theta) \\
\quad + \sum_V \sum_{k_g \in V} A_{k_g}(\theta) \sum_{\substack{k_h \in V \\ h \neq g}} P''_{k_h}(\theta)[P_{k_h}(\theta)]^{-1}P_V(\theta) \\
= \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta)P''_{k_g}(\theta) \\
\quad + \sum_V \sum_{k_g \in V} A_{k_g}(\theta)P_{k_g}(\theta) \sum_{\substack{k_h \in V \\ h \neq g}} P''_{k_h}(\theta)[P_{k_h}(\theta)]^{-1}P_{V-g}(\theta) .
\end{aligned}$$

It is also obvious from (2.5), (2.12) and (2.15) that we can further rewrite the second term of the rightest hand side of (2.38) in such a way that

$$\begin{aligned}
(2.39) \quad \sum_V \sum_{k_g \in V} A_{k_g}(\theta)P_{k_g}(\theta) \sum_{\substack{k_h \in V \\ h \neq g}} P''_{k_h}(\theta)[P_{k_h}(\theta)]^{-1}P_{V-g}(\theta) \\
= \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta)P_{k_g}(\theta) \sum_{V-g} \sum_{k_h \in V-g} P''_{k_h}(\theta)[P_{k_h}(\theta)]^{-1}P_{V-g}(\theta) \\
= \sum_{g=1}^n \sum_{k_g} P'_{k_g}(\theta) \sum_{h \neq g} \sum_{k_h} P''_{k_h}(\theta) = 0 .
\end{aligned}$$

Following a similar process, we have

$$\begin{aligned}
(2.40) \quad \sum_V \sum_{k_g \in V} A_{k_g}(\theta) \sum_{k_h \in V} [A_{k_h}(\theta)]^2 P_V(\theta) &= \sum_V \sum_{k_g \in V} [A_{k_g}(\theta)]^2 \sum_{k_h \in V} A_{k_h}(\theta)P_V(\theta) \\
&= \sum_V \sum_{k_g \in V} [A_{k_g}(\theta)]^3 P_V(\theta)
\end{aligned}$$

$$\begin{aligned}
& + \sum_V \sum_{k_g \in V} [A_{k_g}(\theta)]^2 \sum_{\substack{k_h \in V \\ h \neq g}} A_{k_h}(\theta) P_V(\theta) \\
& = \sum_{g=1}^n \sum_{k_g} [A_{k_g}(\theta)]^2 P'_{k_g}(\theta) .
\end{aligned}$$

Substituting these results into (2.37) and rearranging, we obtain

$$(2.41) \quad E[\epsilon_{1V}(\theta)\epsilon_{2V}(\theta) | \theta] = (1/n^2) \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta) [P''_{k_g}(\theta) - A_{k_g}(\theta) P'_{k_g}(\theta)] .$$

Thus we can write from this result, (2.24) and (2.36)

$$(2.42) \quad E[(\hat{\theta}_V - \theta)\epsilon_{2V}(\theta) | \theta] = (1/n) [I(\theta)]^{-1} \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta) [P''_{k_g}(\theta) - A_{k_g}(\theta) P'_{k_g}(\theta)] ,$$

where  $P'_{k_g}(\theta)$  and  $P''_{k_g}(\theta)$  indicate the first and second derivatives of  $P_{k_g}(\theta)$  with respect to  $\theta$ , respectively. Substituting (2.21), (2.22), (2.35) and (2.42) into (2.34) and rearranging, we obtain for the bias function,  $B(\theta)$ , of the maximum likelihood estimate

$$(2.43) \quad \begin{aligned} B(\theta) = E[\hat{\theta}_V - \theta | \theta] &= -(1/2) [I(\theta)]^{-2} \sum_{g=1}^n \sum_{k_g} A_{k_g}(\theta) P''_{k_g}(\theta) \\ &= -(1/2) [I(\theta)]^{-2} \sum_{g=1}^n \sum_{k_g} P'_{k_g}(\theta) P''_{k_g}(\theta) [P_{k_g}(\theta)]^{-1} . \end{aligned}$$

It is obvious from this result that the bias of the maximum likelihood estimate on the discrete response level has the negative relationship with the amount of test information, i.e., we can expect a small amount of bias when the amount of test information is large, and vice versa. The relationship is rather complicated, however, because of the numerator of the rightest hand side of (2.43), which includes  $P_{k_g}(\theta)$  and its first and second derivatives with respect to  $\theta$ .

On the graded response level, where item score  $x_g$  assumes successive integers, 0 through  $m_g$ , each  $k_g$  in (2.43) must be replaced by  $x_g$ . On the dichotomous response level, it can be reduced to the form

$$(2.44) \quad B(\theta) = E[\hat{\theta}_V - \theta | \theta] = (-1/2) [I(\theta)]^{-2} \sum_{g=1}^n P'_g(\theta) P''_g(\theta) [P_g(\theta) Q_g(\theta)]^{-1} ,$$

where

$$(2.45) \quad Q_g(\theta) = 1 - P_g(\theta) ,$$

with  $P''_g(\theta)$  indicating the second derivative of  $P_g(\theta)$  with respect to  $\theta$ . When  $P'_g(\theta)$  is nonzero throughout the entire range of  $\theta$ , we can also write

$$(2.46) \quad B(\theta) = E[\hat{\theta}_V - \theta | \theta] = (-1/2) [I(\theta)]^{-2} \sum_{g=1}^n I_g(\theta) P''_g(\theta) [P'_g(\theta)]^{-1} .$$

This includes Lord's bias function in the three-parameter logistic model as a special case.

### 3 Dichotomous Response Level

On the dichotomous response level where we deal only with two item score categories, as is exemplified by such a pair as "right" and "wrong", or "agree" or "disagree", the most commonly used family of models may be the one in which the item characteristic function is strictly increasing in  $\theta$ . In such a case,  $P'_g(\theta)$ , the first derivative of the item characteristic function with respect to  $\theta$ , is nonnegative. If, in addition,  $P'_g(\theta)$  is unimodal, as is the case with many commonly used mathematical models, the second derivative,  $P''_g(\theta)$ , assumes positive values up to the modal point, and then it has negative values. A close examination of (2.46) reveals that, in such a model, the direction of bias is positive for very high levels of  $\theta$ , and it is negative for very low levels of  $\theta$ . In other words, individuals of very high levels of ability tend to be overevaluated, and those of very low levels of ability tend to be underevaluated.

Now we shall observe the bias functions in some specified models which belong to this family.

#### 3.1 Normal Ogive Model

In the normal ogive model, the item characteristic function is given by

$$(3.1) \quad P_g(\theta) = (2\pi)^{-1/2} \int_{-\infty}^{a_g(\theta - b_g)} e^{-u^2/2} du ,$$

where  $a_g$  and  $b_g$  are the item discrimination and difficulty parameters, respectively. From (3.1), we can write for the first and second derivatives of  $P_g(\theta)$  with respect to  $\theta$

$$(3.2) \quad P'_g(\theta) = a_g(2\pi)^{-1/2} e^{-a_g^2(\theta - b_g)^2/2}$$

and

$$(3.3) \quad P''_g(\theta) = -a_g^2(\theta - b_g)P'_g(\theta) ,$$

respectively. Substituting (3.2) and (3.3) into (2.46) and rearranging, we obtain for the bias function

$$(3.4) \quad B(\theta) = (1/2)[I(\theta)]^{-2} \sum_{g=1}^n a_g^2(\theta - b_g)I_g(\theta) .$$

It is obvious from its definition, which is given by (1.6), that the item information function,  $I_g(\theta)$ , is nonnegative regardless of the mathematical model. From this, we can see that for a fixed value of  $\theta$  the sign of the term under the summation sign in (3.4) depends upon the value of the difficulty parameter  $b_g$  of each item  $g$ . We can also see that, if all the  $n$  items have the same values of difficulty parameter, i.e.,  $b = b_1 = b_2 = \dots = b_n$ , then the bias function is strictly increasing in  $\theta$ , and equals zero only at  $\theta = b$ , with positive and negative infinities as its two asymptotes. Moreover, since  $P_g(\theta)$  is point-symmetric with  $(b, 0.5)$  as the point of symmetry, the bias function is also point-symmetric with the same point of symmetry. In this situation, generally speaking, we should expect a substantial amount of bias as we depart from  $\theta = b$ .

In many situations of practical importance, however, it is desirable to have a test whose bias function practically assumes zero for a wide range of  $\theta$ . Equation (3.4) suggests that, in order to materialize such a test, we must develop a set of items whose difficulty parameters distribute widely and evenly, so that, for a wide interval of  $\theta$ , the negative and positive terms under the summation sign of the right hand side of (3.4) practically "cancel each other out".

Table 3-1 presents the estimated item discrimination parameter  $\hat{a}_g$  and item difficulty parameter  $\hat{b}_g$  for each of the forty-three dichotomous test items of the Level 11 Vocabulary Subtest of the Iowa Tests of Basic Skills, which were obtained by assuming the normal ogive model (Samejima, 1984a). The data were collected for 2,356 school children of approximately age eleven by the Iowa Testing Bureau, and were analyzed by using the tetrachoric correlation matrix and the principal factor solution of factor analysis. Thus we have for the item parameter estimates

$$(3.5) \quad \hat{a}_g = \rho_g(1 - \rho_g^2)^{-1/2}$$

and

$$(3.6) \quad \hat{b}_g = \gamma_g \rho_g^{-1} ,$$

where  $\rho_g$  is the factor loading of item  $g$  on the single, dominating common factor, which is operationally defined as  $\theta$ , and  $\gamma_g$  is the normal deviate corresponding to the proportion correct,  $p_g$ , of item  $g$ . Corresponding results for each of the fifty-five dichotomous test items of Test J1 of Shiba's Word/Phrase Comprehension Tests are also presented as Table 3-2 (cf. Shiba, 1978, Samejima, 1984b). Those data were based upon 2,259 junior high school students in Japan.

Figure 3-1 presents the square root of the test information function of each of these two tests by solid and dashed lines, respectively. We can see that these curves are fairly similar. The bias functions, which were obtained by (3.4) for the two tests, are shown in Figure 3-2. We can see that, over all, the bias is less conspicuous for Test J1 than for Iowa Subtest. In order to show the relationship between the amount of bias and that of test information, Figure 3-3 presents the square root of test information and the bias function together for each of the two tests. It is interesting to note that in both cases, if we tolerate biases of up to  $\pm 0.1$ , for example, then the range of  $\theta$  in which this is the case corresponds to the interval of  $\theta$  where the square root of test information is approximately 1.75 or greater, or where the amount of test information is approximately 3.0 or greater.

### 3.2 Logistic Model

Since the logistic model can be considered as a special case of the three-parameter logistic model in which we set the guessing parameter,  $c_g$ , equal to zero, Lord's bias function, which is written as formula (1.4) in the present paper, is also applicable. Note, however, that neither  $I_g(\theta)$  nor  $I(\theta)$  in the formula includes the guessing parameter  $c_g$ , when it is used for the logistic model.

A close examination of (1.4) reveals strong similarities of the logistic model with the normal ogive model, i.e., 1) for a fixed value of  $\theta$ , the sign of the term under the summation sign in (1.4) depends upon the difficulty parameter  $b_g$  of each item  $g$ ; 2) if  $b = b_1 = b_2 = \dots = b_n$ , the bias function is strictly increasing in  $\theta$  with positive and negative infinities as its two asymptotes, which equals zero only at  $\theta = b$ , and is point-symmetric with  $(b, 0.5)$  as the point of symmetry; and 3) in order to make the bias practically nil for a wide range of  $\theta$ , we must develop items whose difficulty parameters distribute widely and evenly.

Figures 3-4 and 3-5 present two sets of examples of the square root of the test information function and the bias function obtained by following the logistic model with  $D = 1.7$ , by solid and dashed lines, respectively. They are the results obtained by using the same set of estimated discrimination and difficulty parameters of the Iowa Level 11 Vocabulary Subtest and Shiba's Test J1, which are shown in Tables 3-1 and 3-2, respectively, as we did for the normal ogive model in the preceding section. As is expected, these results are close to those obtained by following the normal ogive model. As was done in the normal ogive model, the square root of test information and the bias function are put together for each of the two tests, and presented in Figure 3-6. The relationship between the square root of test information and the amount of bias appears to be almost the same as was recognized in the corresponding result obtained by following the normal ogive model, which we discussed in the preceding section.

TABLE 3-1

Estimated Item Discrimination Parameter  $\hat{a}_g$  and Item Difficulty Parameter  $\hat{b}_g$  for Each of the Forty-Three Dichotomous Test Items of the Level 11 Vocabulary Subtest of the Iowa Tests of Basic Skills, Based upon the Result Collected for 2,356 School Children of Approximately Age Eleven.

Item $g$	Discrimination Parameter $\hat{a}_g$	Difficulty Parameter $\hat{b}_g$
24	0.196	-4.257
25	0.829	-1.000
26	0.614	-0.821
27	0.594	-0.340
28	0.669	-0.900
29	0.867	-1.077
30	0.956	-0.557
31	0.938	-0.179
32	0.940	-0.803
33	0.434	-2.331
34	0.598	-1.210
35	0.489	-0.569
36	0.657	-0.987
37	0.351	0.577
38	0.665	-0.468
39	0.333	-0.676
40	0.683	0.402
41	0.531	-0.948
42	0.436	0.258
43	0.672	-0.867
44	0.143	4.175
45	0.898	-0.357

Item $g$	Discrimination Parameter $\hat{a}_g$	Difficulty Parameter $\hat{b}_g$
46	0.612	-0.318
47	0.494	-0.781
48	0.849	0.054
49	0.421	-0.626
50	0.346	-0.250
51	0.664	-0.420
52	0.640	0.217
53	0.402	0.526
54	0.573	0.126
55	0.667	-0.342
56	0.593	1.007
57	0.370	0.398
58	0.416	0.782
59	0.491	-0.731
60	0.678	-0.170
61	0.519	0.748
62	0.938	-0.485
63	0.637	-0.398
64	0.818	-0.042
65	0.606	0.595
66	0.604	-0.376



TABLE 3-2

Estimated Item Discrimination Parameter  $\hat{a}_g$  and Item Difficulty Parameter  $\hat{b}_g$  for Each of the Fifty-Five Dichotomous Test Items of Test J1 of Shiba's Word/Phrase Comprehension Tests Collected for 2,259 Junior High School Students.

Item $g$	Discrimination Parameter $\hat{a}_g$	Difficulty Parameter $\hat{b}_g$
J101	0.726	-0.238
J102	0.537	-0.956
J103	0.568	-1.263
J104	0.710	-0.809
J105	0.794	-0.097
J106	0.495	-0.741
J107	0.583	0.205
J108	0.771	-1.974
J109	0.386	-0.872
J110	0.572	-0.327
J111	0.950	-1.266
J112	0.437	-1.036
J113	0.508	-1.061
J114	0.472	0.486
J115	0.704	-0.224
J116	0.303	-1.671
J117	0.390	-0.626
J118	0.583	-1.573
J119	0.653	-0.972
J120	0.293	1.058
J121	0.470	-0.904
J122	0.451	-1.038
J123	0.456	0.151
J124	0.562	-1.313
J125	0.450	-1.691
J126	0.367	-0.424
J127	0.525	-1.299
J128	0.679	-1.094

Item $g$	Discrimination Parameter $\hat{a}_g$	Difficulty Parameter $\hat{b}_g$
J129	0.761	1.416
J130	0.351	-1.839
J131	0.798	-0.494
J132	0.322	0.162
J133	0.822	-1.377
J134	0.302	1.633
J135	0.850	-0.225
J136	0.368	0.264
J137	0.591	0.331
J138	—	—
J139	0.375	1.602
J140	0.422	0.216
J141	0.566	-0.689
J142	0.447	0.132
J143	0.586	-0.100
J144	0.384	-0.399
J145	0.630	-0.479
J146	0.880	0.057
J147	0.333	0.374
J148	0.521	-0.062
J149	0.509	-0.108
J150	0.512	-0.040
J151	0.462	0.907
J152	0.394	0.478
J153	0.384	2.029
J154	0.242	2.353
J155	0.738	1.258
J156	0.655	1.468

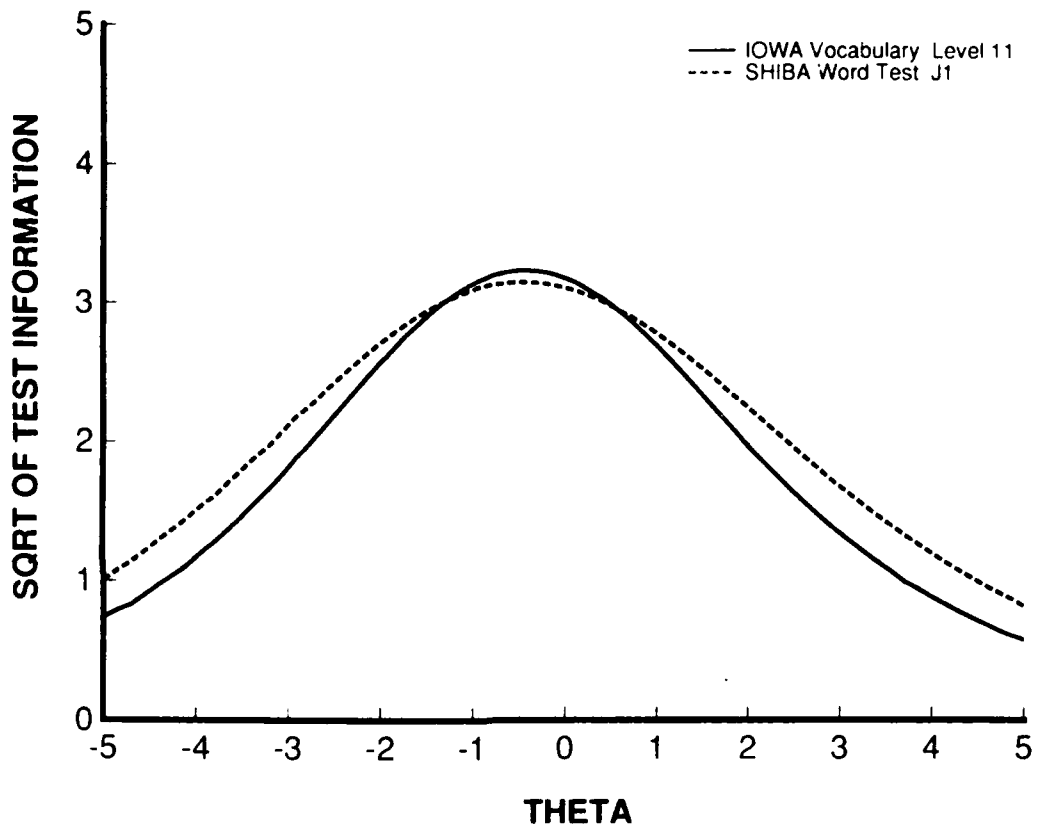
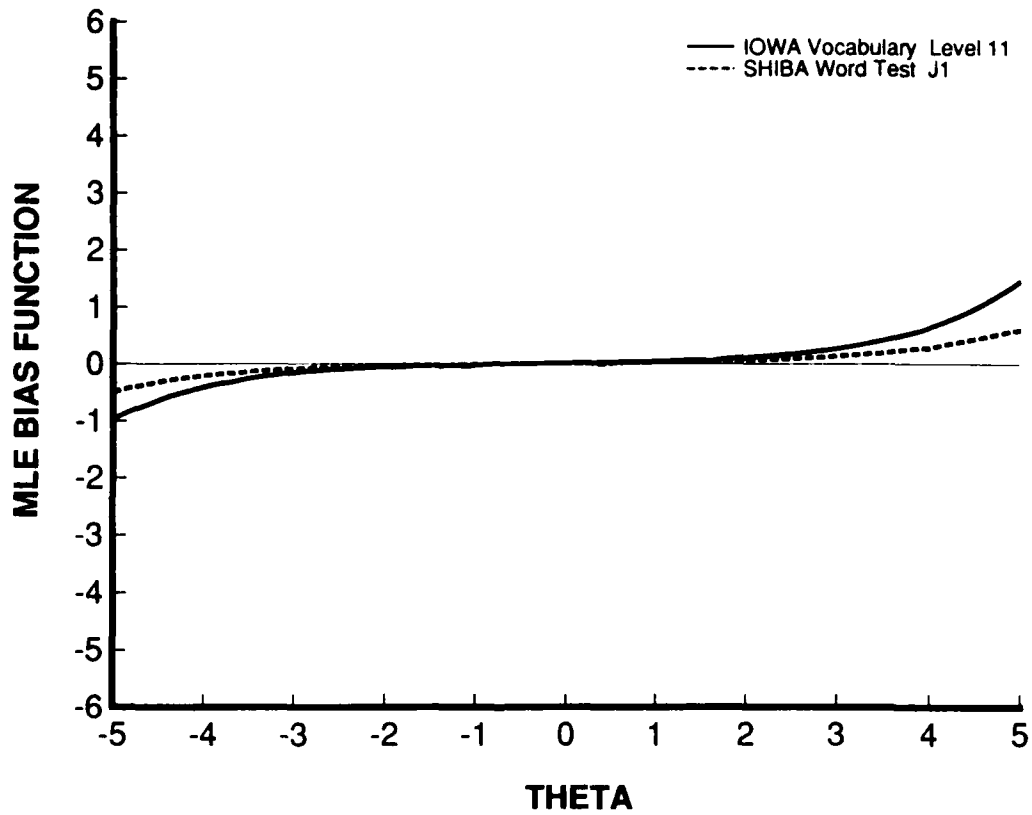


FIGURE 3-1

Square Roots of the Test Information Functions for the Iowa Level 11 Vocabulary Subtest (Solid Line) and for Shiba's Test J1 (Dashed Line), Following the Normal Ogive Model.



0 790 0 50 1 50 8 50 8 00

FIGURE 3-2

Bias Functions of the Iowa Level 11 Vocabulary Subtest (Solid Line) and of Shiba's Test J1 (Dashed Line), Following the Normal Ogive Model.

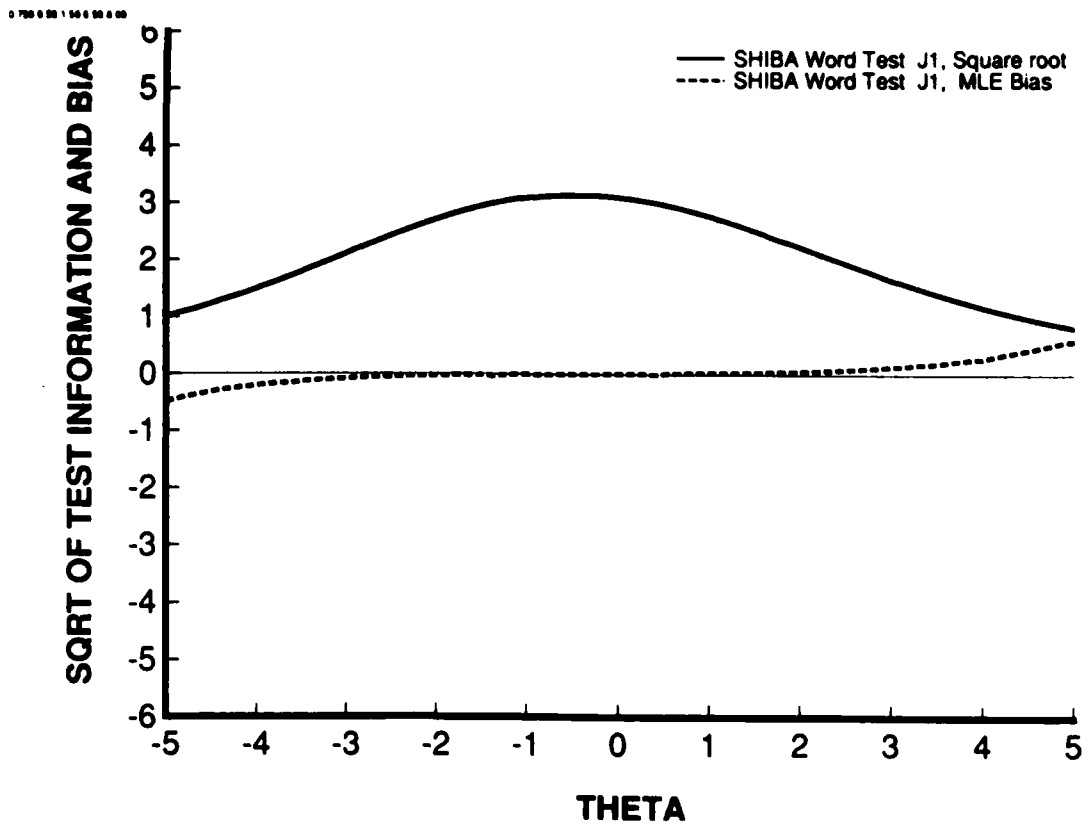
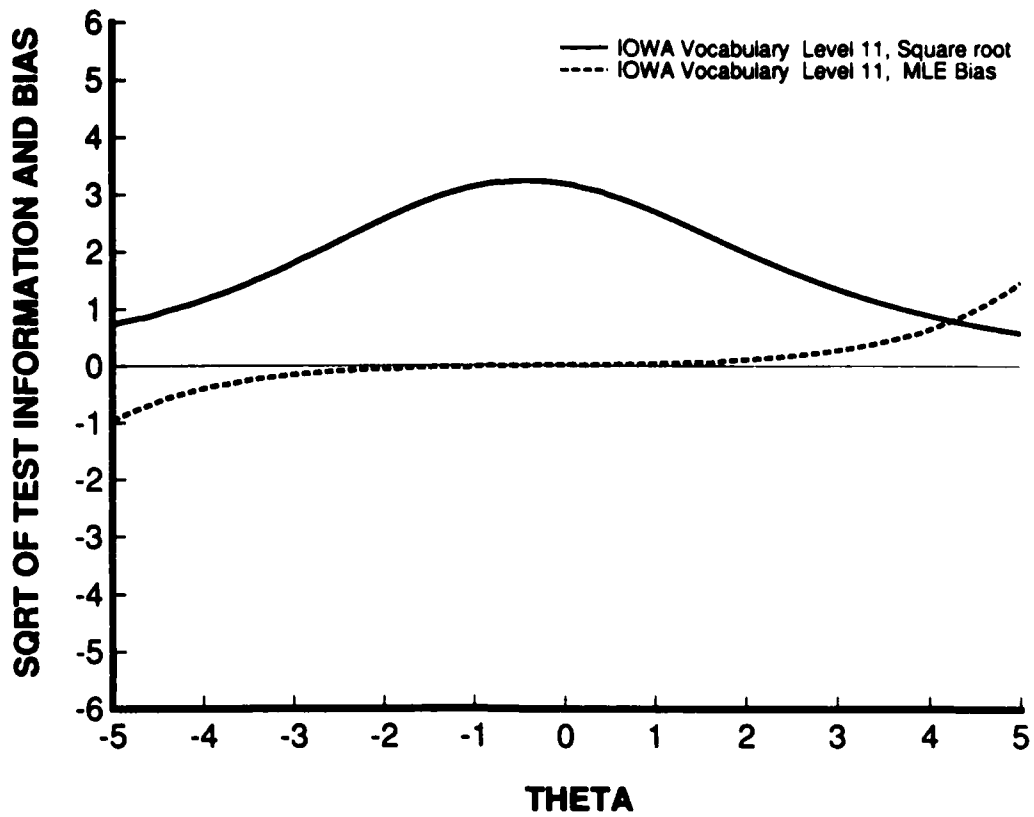
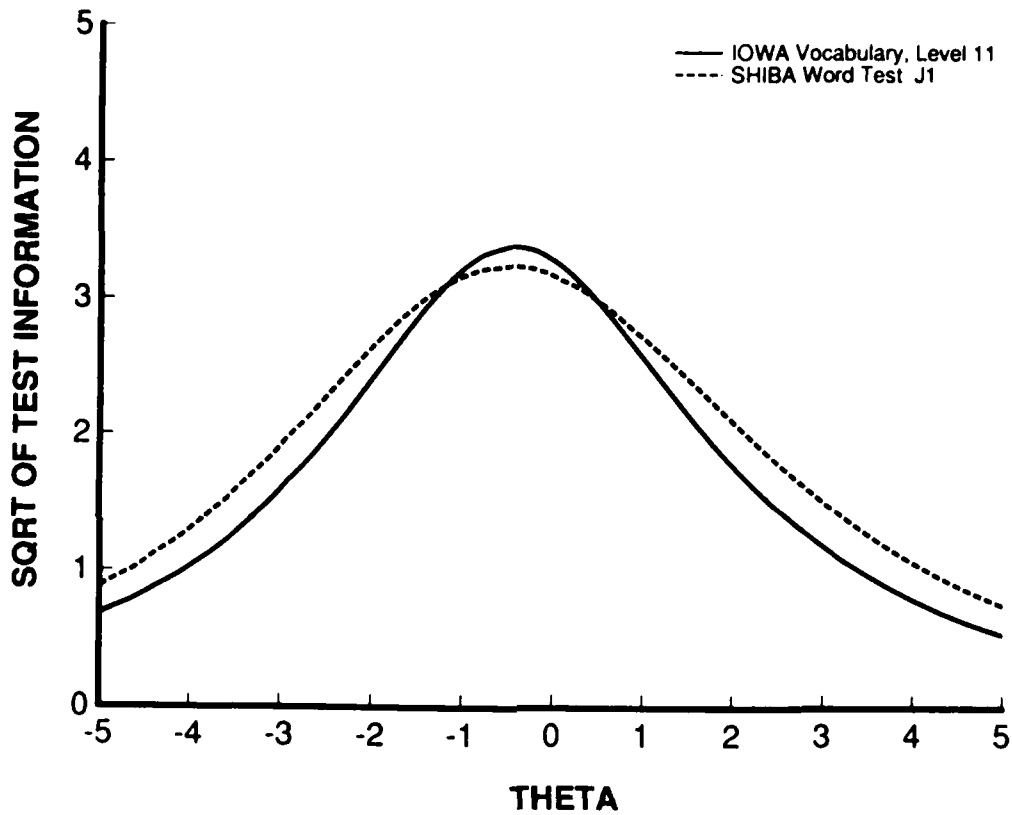


FIGURE 3-3

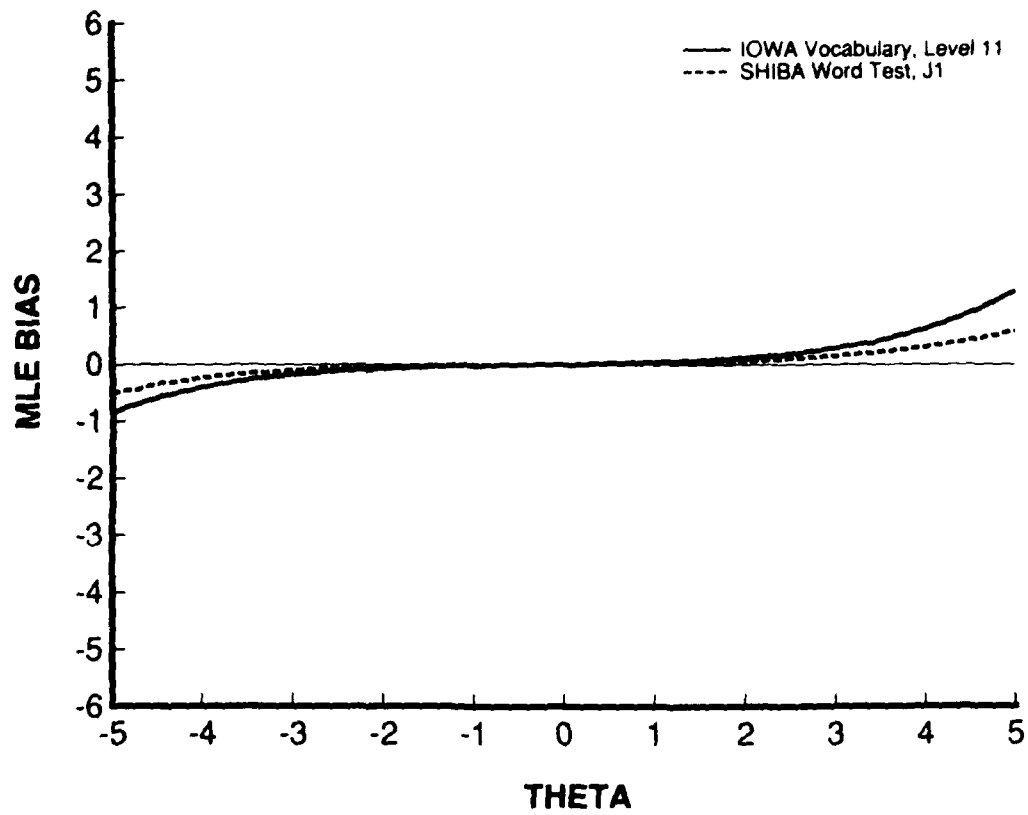
Comparison of the Square Root of the Test Information Function (Solid Line) and the Bias Function (Dashed Line) Following the Normal Ogive Model, of Each of the Two Tests, i.e., the Iowa Level 11 Vocabulary Subtest (Upper Graph) and Shiba's Test J1 (Lower Graph).



0 750 0 50 1 10 6 50 8 70

FIGURE 3-4

Square Roots of the Test Information Functions for the Iowa Level 11 Vocabulary Subtest (Solid Line) and for Shiba's Test J1 (Dashed Line), Following the Logistic Model.



0 750 0 50 1 50 5 50 8 00

FIGURE 3-5

Bias Functions of the Iowa Level 11 Vocabulary Subtest (Solid Line) and of Shiba's Test J1 (Dashed Line), Following the Logistic Model.

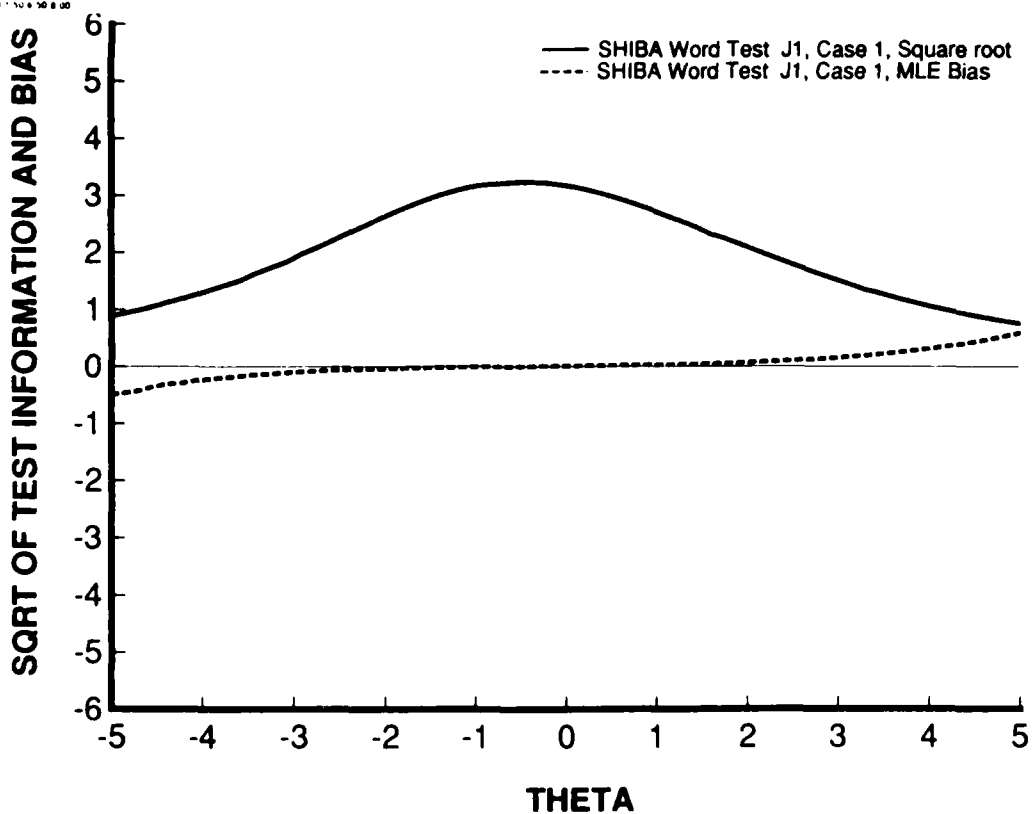
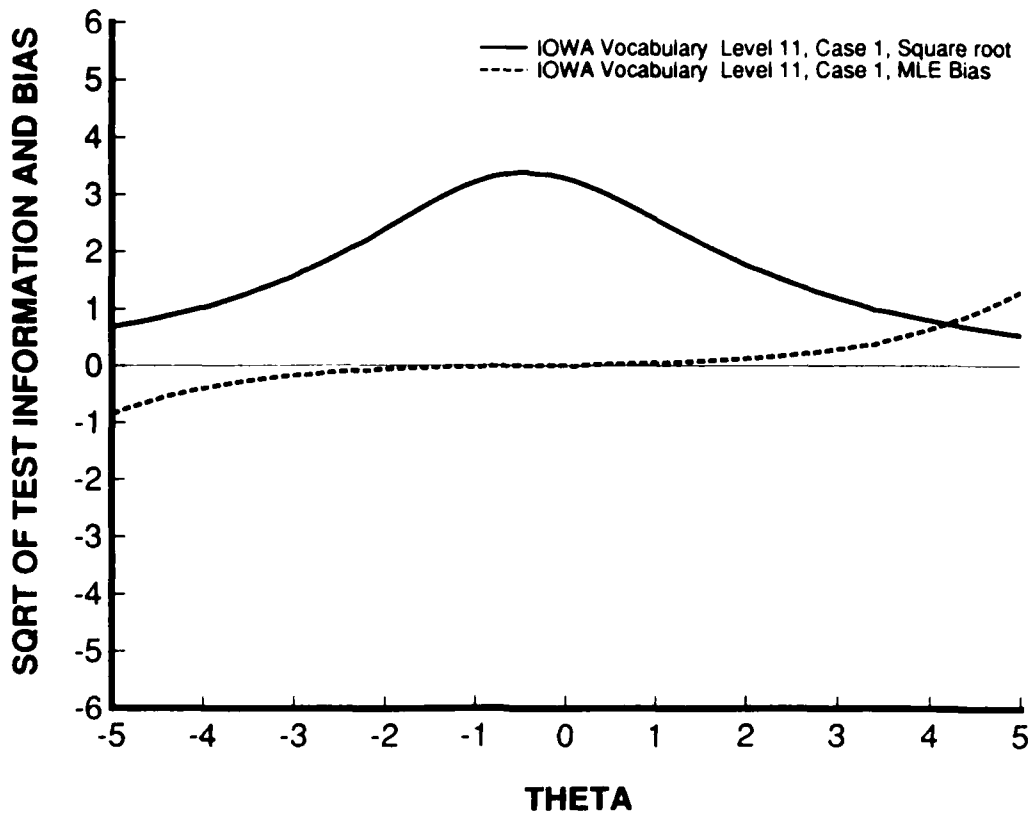


FIGURE 3-6

Comparison of the Square Root of the Test Information Function (Solid Line) and the Bias Function (Dashed Line) Following the Logistic Model, of Each of the Two Tests, i.e., the Iowa Level 11 Vocabulary Subtest (Upper Graph) and Shiba's Test J1 (Lower Graph).

### 3.3 Rasch Model

Since the Rasch model is a special case of the logistic model, in which the discrimination parameters of all the items are identical, i.e.,  $a = a_1 = a_2 = \dots = a_n$ , the bias function is obtained by removing  $a_y$  on the right hand side of formula (1.4), provided that the scale unit of  $\theta$  be adjusted to the common discrimination parameter. Thus all the observations made for the logistic model also apply for the Rasch model.

Figure 3-7 presents the item characteristic functions of 25 items, each of which follows the Rasch model with  $D = 1.7$ . The item difficulty parameters of these items are equally spaced, starting with  $b_y = -3.0$  and ending with  $b_y = +3.0$  with equal step widths of 0.25. The square root of the test information of this hypothetical test is shown by a solid line in Figure 3-8. In the same figure, also presented are the square root of test information of each of the two subtests, i.e. the subtest of 13 items which is constructed by taking every other curve in Figure 3-7, and the subtest of 7 items obtained by changing the equal steps from 0.25 to 1.0. They are drawn by dashed and dotted lines, respectively. The bias functions of these three tests are shown in Figure 3-9 using the same types of lines. In this figure, unlike the square root of test information, it looks as if substantial changes in the number of items did not affect the bias functions to a great extent, especially for the range of  $\theta$ ,  $-2.0$  through  $2.0$ . In order to show the relationship between the square root of test information and the bias function more clearly, Figure 3-10 presents both curves together for each of the three hypothetical tests. If we tolerate biases of  $\pm 0.1$ , as we did before, then for the 25 item test, the critical value of the square root of test information is approximately 2.0, and for the 13 and 7 item tests, these values are approximately 1.6 and 1.1, respectively.

This result indicates the importance of the configuration of the item difficulty parameters, i.e., even if the number of items is as small as seven, the approximate unbiasedness can be reached for a wide range of  $\theta$ , provided that the item difficulty parameters are distributed evenly for a wide range.

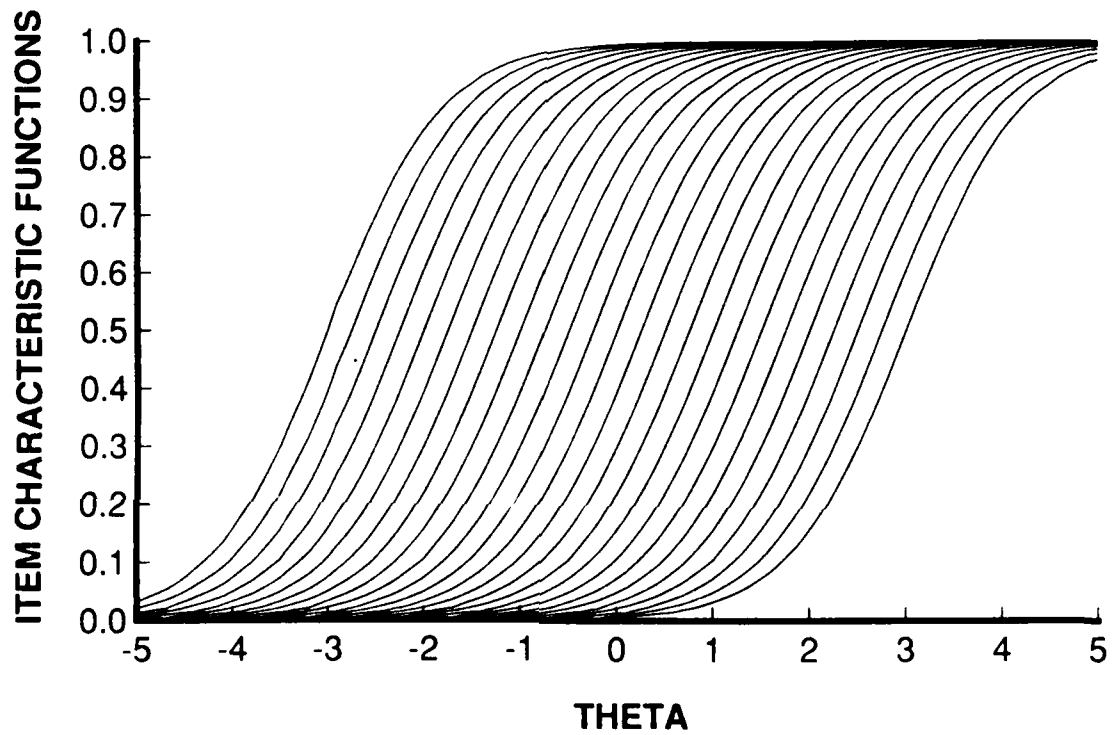
### 3.4 Three-Parameter Logistic Model

In the three-parameter logistic model, the item characteristic function is not point-symmetric, as is clear from the formula (1.3). For this reason, even if the difficulty parameters of all the  $n$  items are equal, the bias function, which is given by (1.4), is not point-symmetric either, unlike those in the normal ogive and logistic models. Since random guessing is nothing but noise, there is a certain amount of decrement in the accuracy of estimation, especially on the lower levels of ability or latent trait. Consequently, we must expect a larger amount of bias, especially on the lower levels.

For the purpose of illustration, the guessing parameters, 0.20 and 0.25, were added to the estimated discrimination and difficulty parameters of each of the 43 test items of the Iowa Level 11 Vocabulary Subtest, which are shown in Table 3-1, respectively, to create two more hypothetical tests. Figure 3-11 presents the square roots of the test information function of these two hypothetical tests by dashed and dotted lines, respectively, in comparison with the one following the (two-parameter) logistic model, which is shown by a solid line. We can see that, in each case, the decrement caused by the guessing parameters is substantial, especially on the lower levels of ability. The bias functions of these two hypothetical tests are shown in Figure 3-12 in comparison with the one for the logistic model, using the corresponding types of lines. It is obvious that random guessing causes a substantial amount of additional bias, especially on the lower levels of ability. Figure 3-13 compares the square root of test information with the amount of bias for each of the two hypothetical tests, as we did previously. It looks as if the same rule held in these two cases of the three-parameter logistic model, i.e., the amount of bias is within the range of  $\pm 0.1$  for the intervals of  $\theta$  for which the square root of test information is approximately 1.75 or greater, just as we observed in the examples of the (two-parameter) normal ogive and logistic models. Such intervals of  $\theta$  are substantially smaller, however.

In a similar manner, two additional hypothetical tests were created with  $c_y = 0.20$  and  $c_y = 0.25$  as the guessing parameters, respectively, added to the estimated discrimination and difficulty parameters of each item of Shiba's Test J1, which are shown in Table 3-2. These results are presented in Figures 3-14 through 3-16. We can see in these figures that the results are very similar to the corresponding





214001 501 50 6 10 8 30

FIGURE 3-7

Twenty-Five Item Characteristic Functions Following Rasch Model with  $D = 1.7$  and Equally Spaced Difficulty Parameters Ranging from  $-3.0$  to  $+3.0$ .

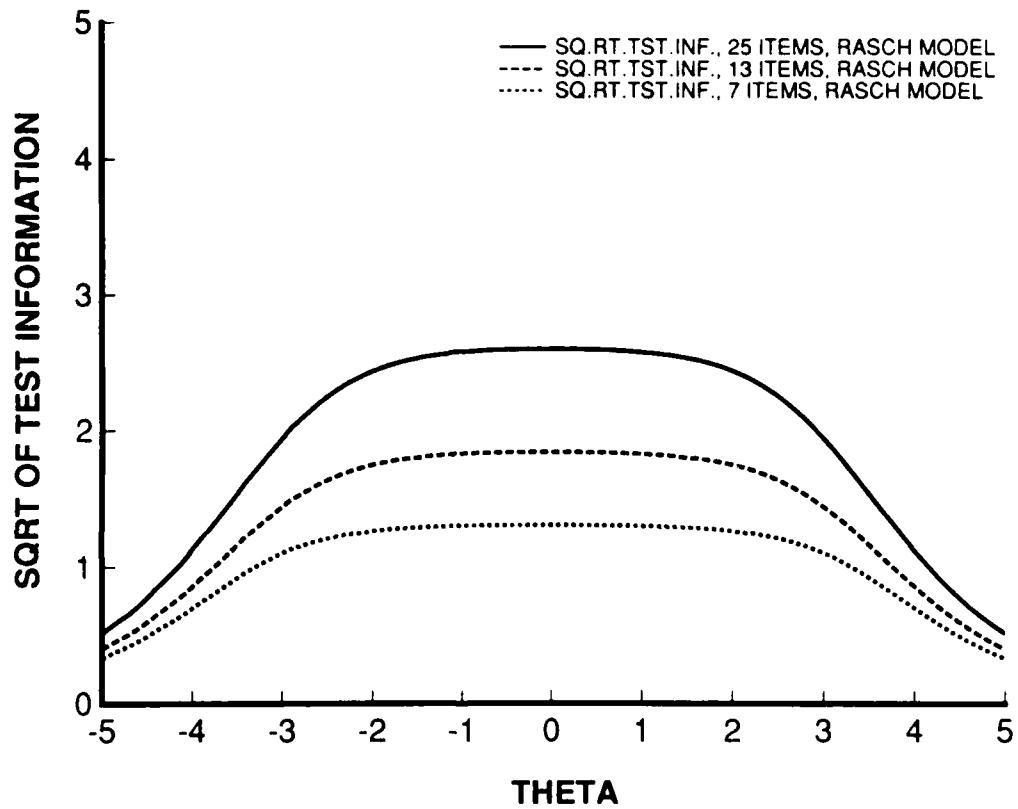
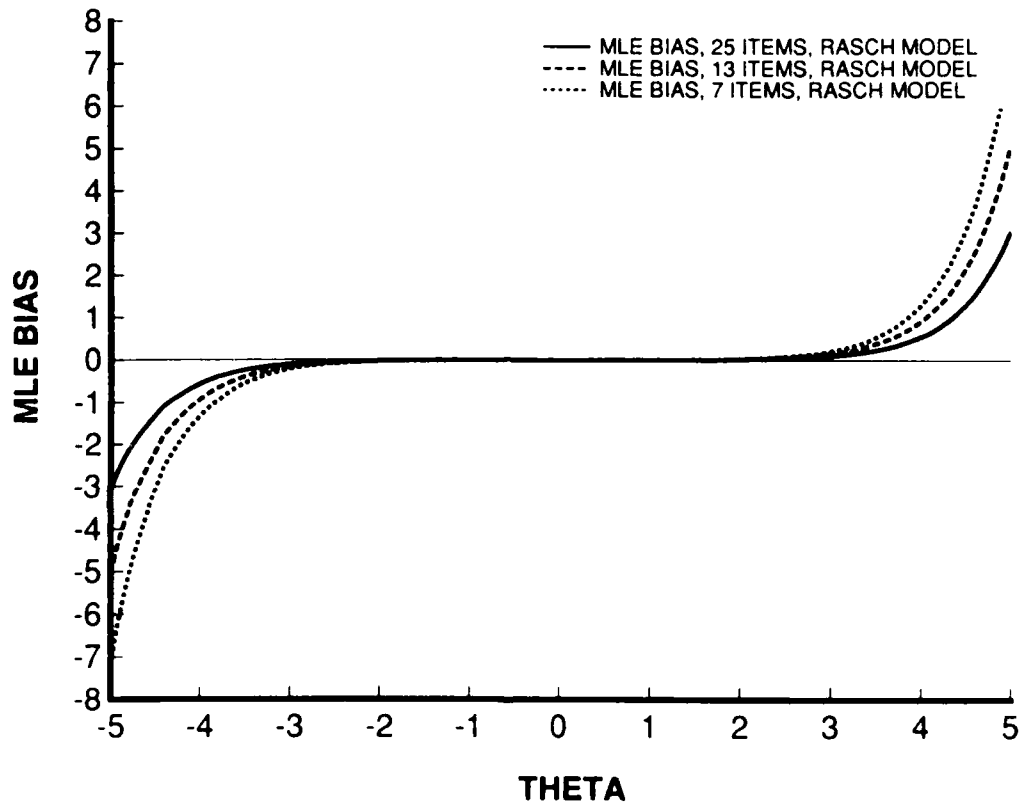


FIGURE 3-8

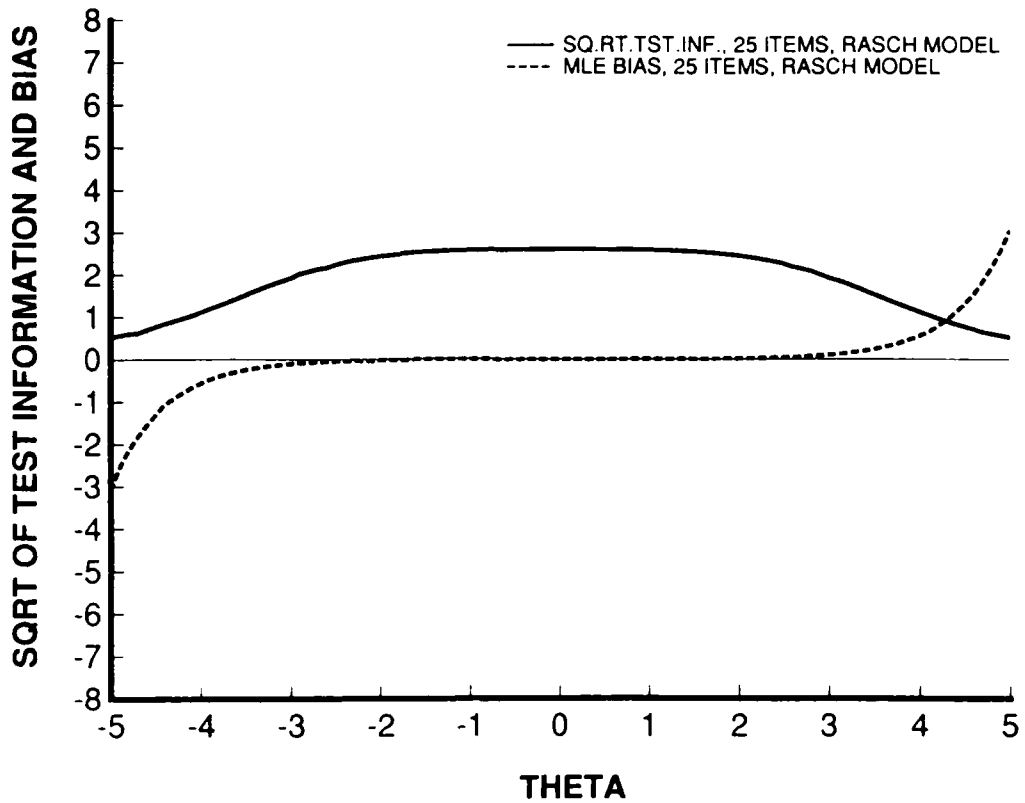
Square Root of Test Information of Each of Three Hypothetical Tests of 25 (Solid Line), 13 (Dashed Line) and 7 (Dotted Line) Items, Respectively, Following Rasch Model with  $D = 1.7$  and Equally Spaced Difficulty Parameters Ranging from  $-3.0$  to  $+3.0$ .



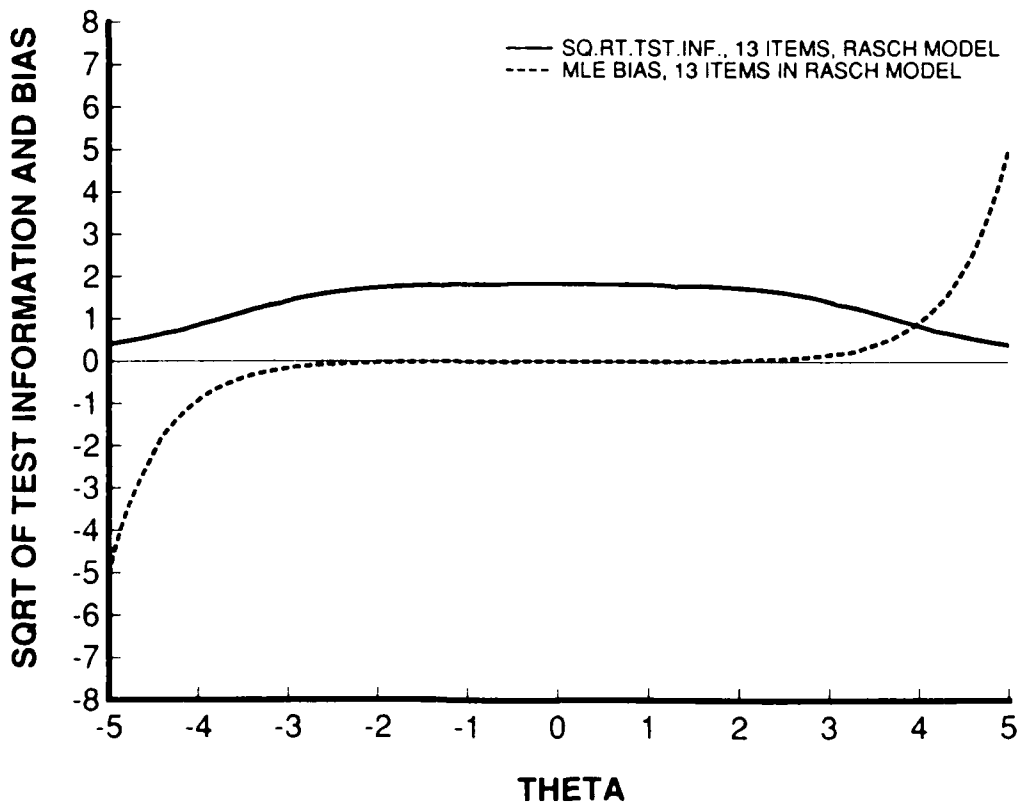
0 50 0 50 1 50 6 50 8 00

FIGURE 3-9

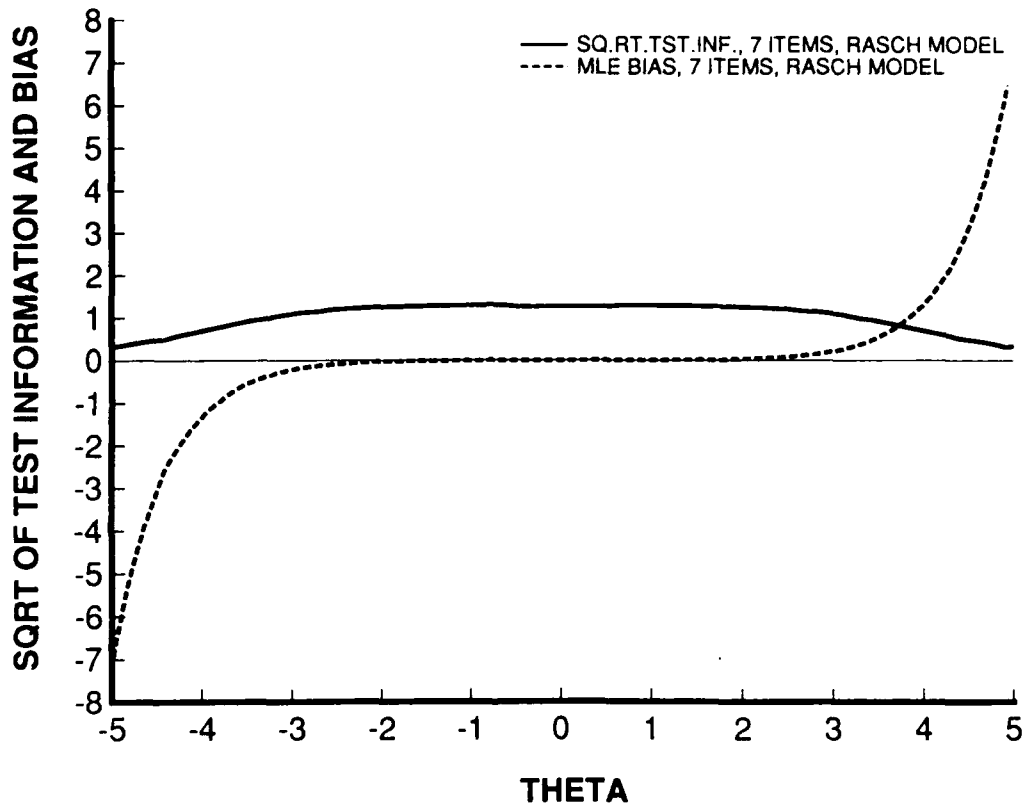
MLE Bias Function of Each of Three Hypothetical Tests of 25 (Solid Line), 13 (Dashed Line) and 7 (Dotted Line) Items, Respectively, Following Rasch Model with  $D = 1.7$  and Equally Spaced Difficulty Parameters Ranging from  $-3.0$  to  $+3.0$ .



0 50 0 50 1 50 8 50 8 30



0 750 0 50 1 50 8 50 8 30



0 750 0 50 1 50 6 50 8 00

FIGURE 3-10

MLE Bias Function (Dashed Line) and the Square Root of Test Information (Solid Line) of Each of Three Hypothetical Tests of 25 (Solid Line), 13 (Dashed Line) and 7 (Dotted Line) Items, Respectively, Following Rasch Model with  $D = 1.7$  and Equally Spaced Difficulty Parameters Ranging from  $-3.0$  to  $+3.0$ .

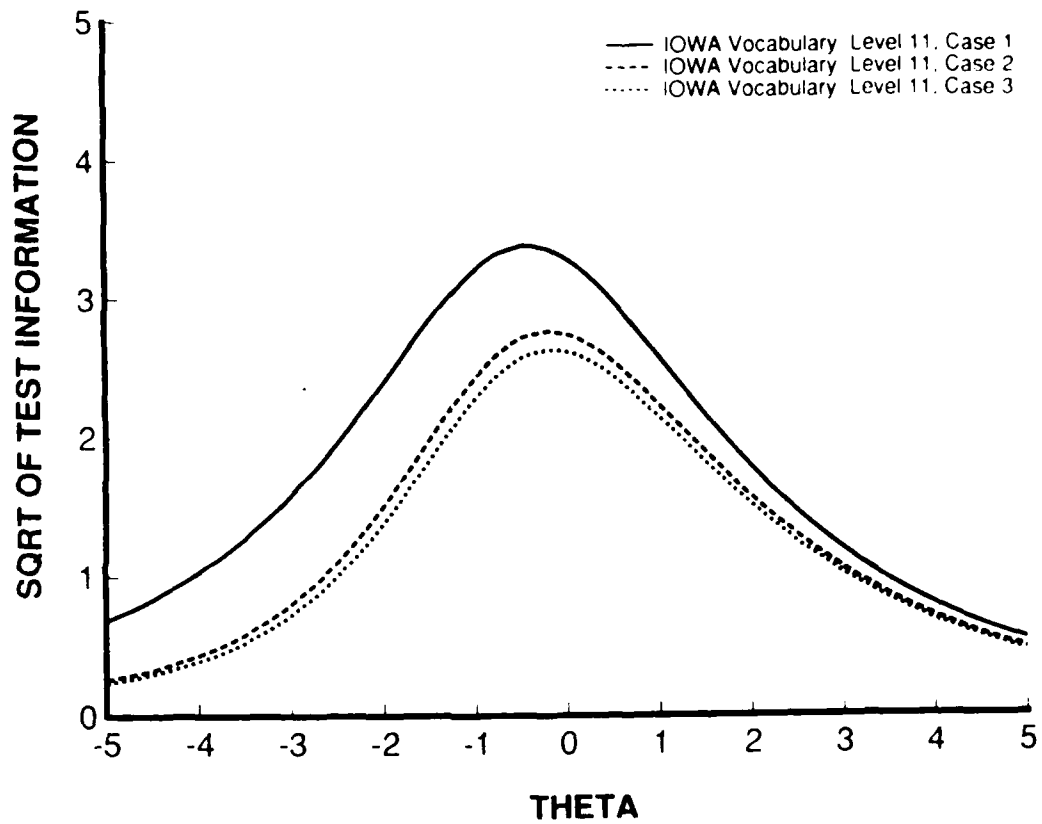
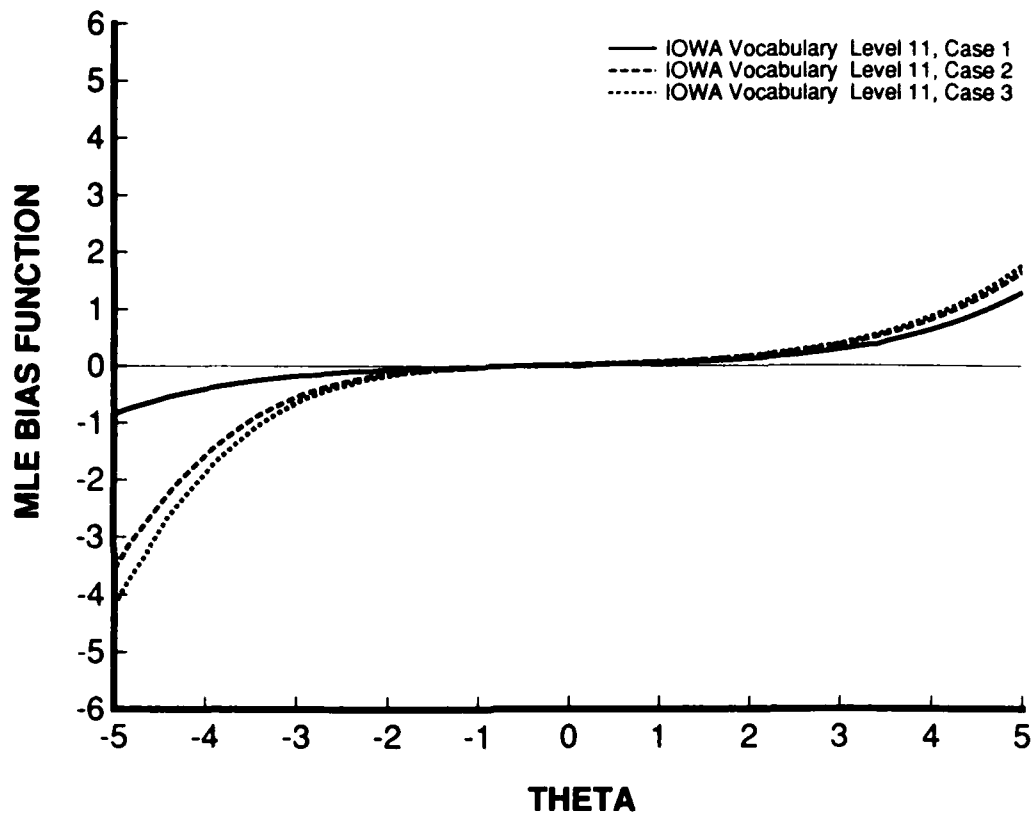


FIGURE 3-11

Square Roots of Test Information of the Iowa Level 11 Vocabulary Subtest (Solid Line) of 43 Items Following the Logistic Model, and of Two Hypothetical Tests of the Same Number of Items Each Following the Three-Parameter Logistic Model, which Share the Same Set of Item Discrimination and Difficulty Parameters as the Iowa Subtest and with the Guessing Parameters 0.20 (Dashed Line) and 0.25 (Dotted Line), Respectively.



0 750 0 50 1 50 8 50 8 00

FIGURE 3-12

MLE Bias Functions of the Iowa Level 11 Vocabulary Subtest of 43 Items Following the Logistic Model, and of Two Hypothetical Tests of the Same Number of Items Each Following the Three-Parameter Logistic Model, which Share the Same Set of Item Discrimination and Difficulty Parameters as the Iowa Subtest and with the Guessing Parameters 0.20 (Dashed Line) and 0.25 (Dotted Line), Respectively.

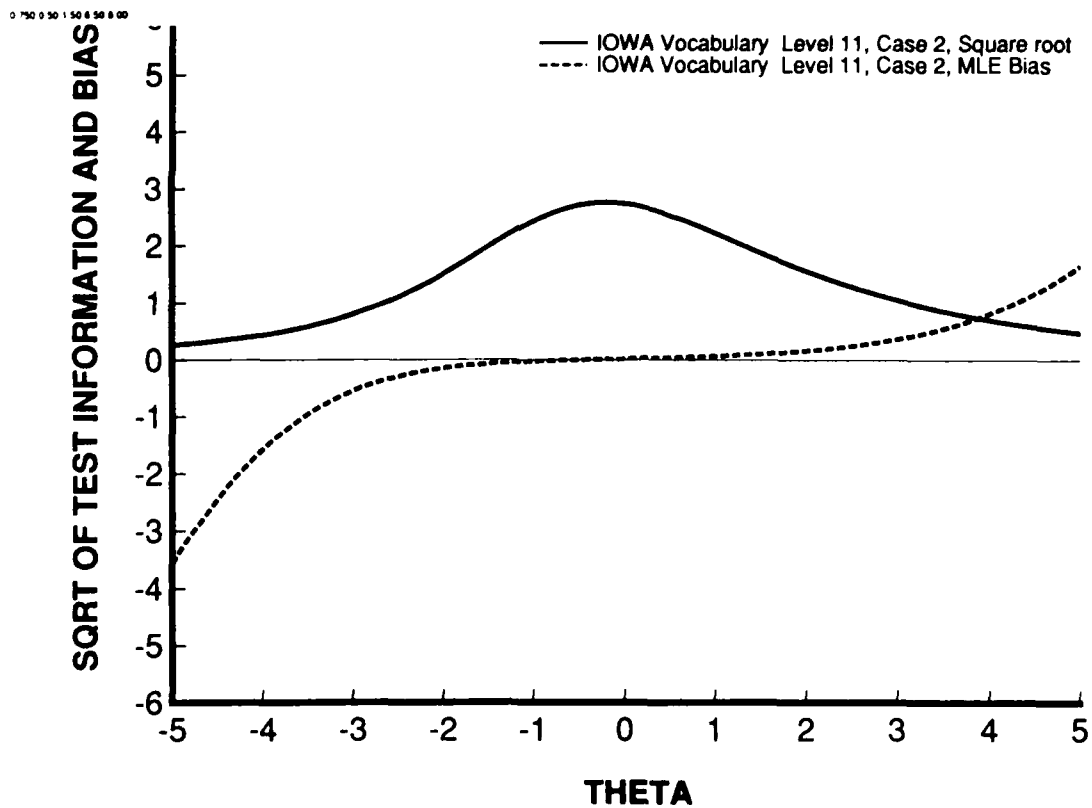
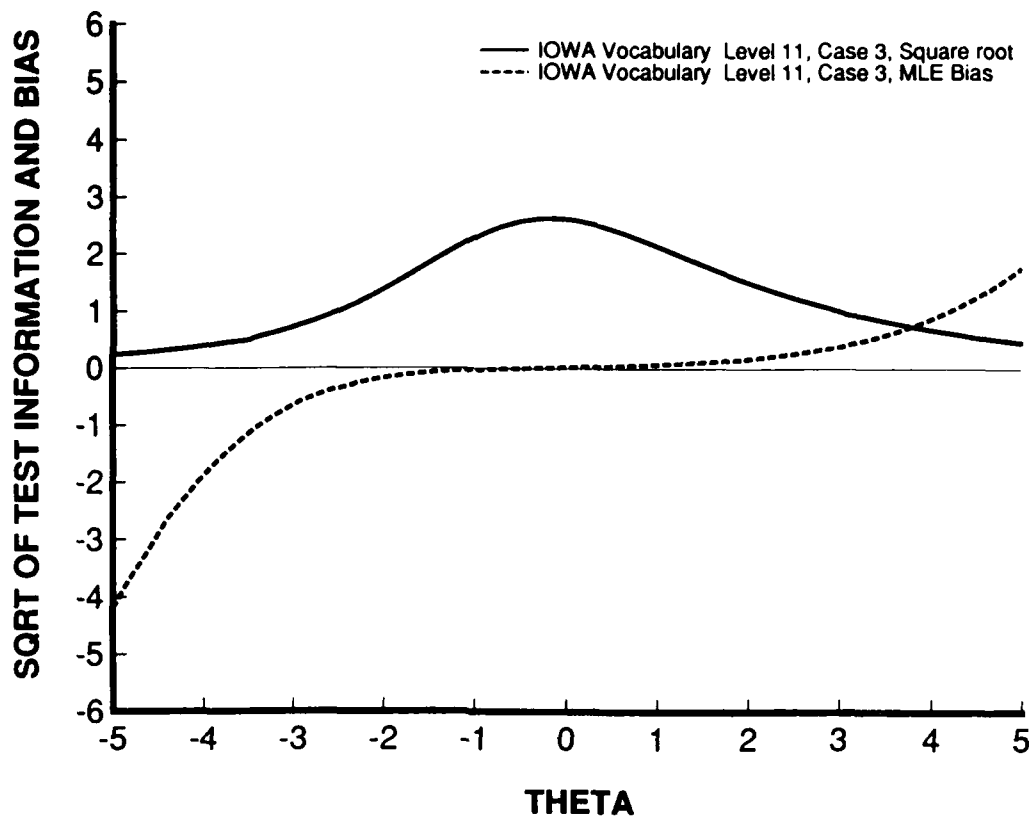
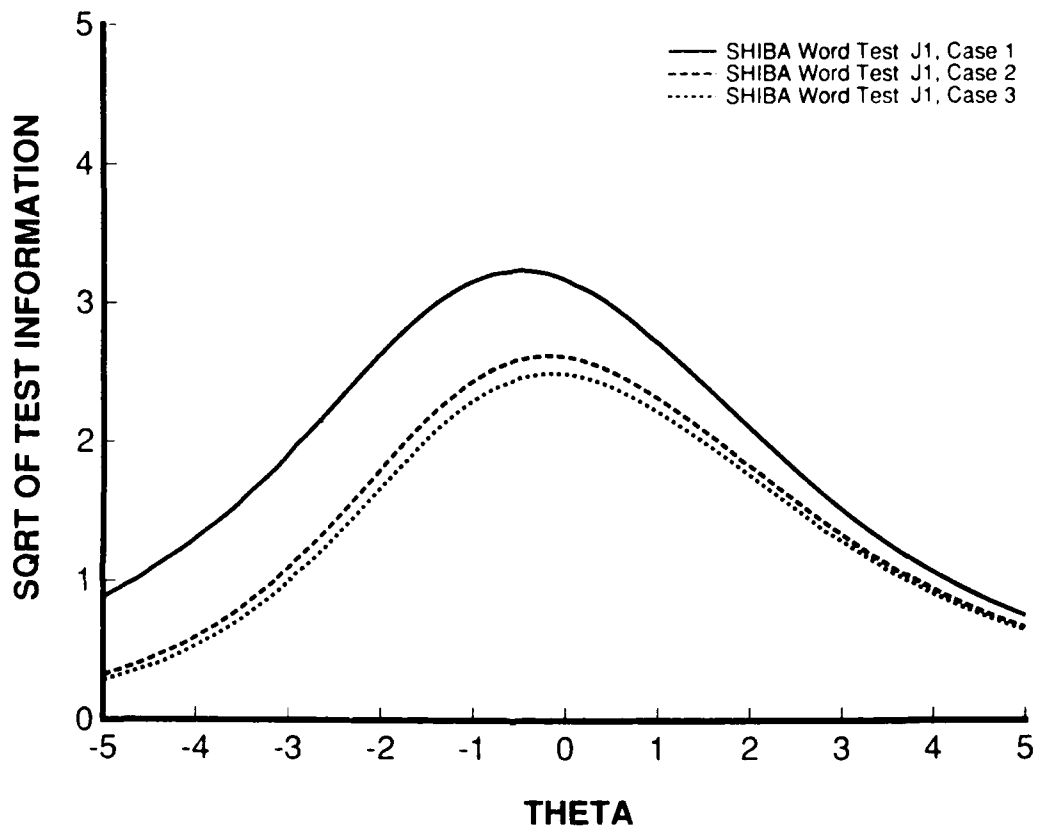


FIGURE 3-13

MLE Bias Function (Dashed Line) and the Square Root of Test Information (Solid Line) of Each of the Two Hypothetical Tests of 43 Items Each Following the Three-Parameter Logistic Model, which Share the Same Set of Item Discrimination and Difficulty Parameters as the Iowa Subtest and with the Guessing Parameters 0.20 (Dashed Line) and 0.25 (Dotted Line), Respectively.





0 750 0 50 1 51 6 51 4 30

FIGURE 3-14

Square Roots of Test Information of Shiba's Word/Phrase Comprehension Test J1 (Solid Line) of 55 Items Following the Logistic Model, and of Two Hypothetical Tests of the Same Number of Items Each Following the Three-Parameter Logistic Model, which Share the Same Set of Item Discrimination and Difficulty Parameters as Test J1 and with the Guessing Parameters 0.20 (Dashed Line) and 0.25 (Dotted Line), Respectively.

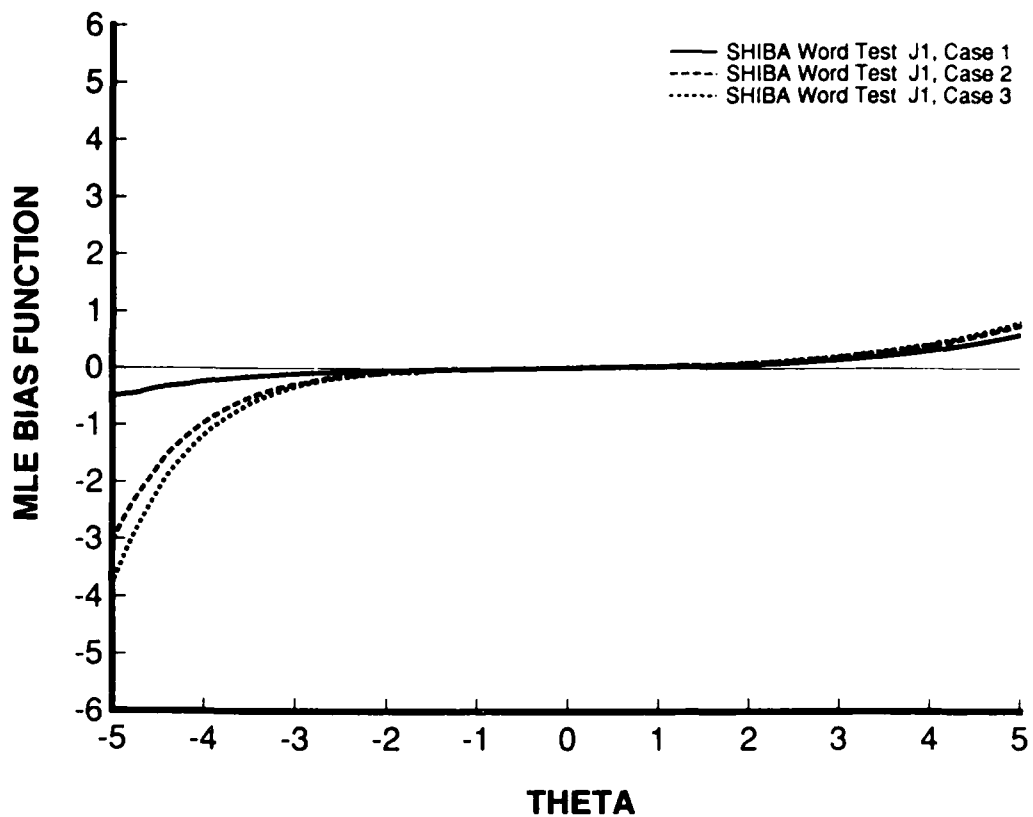


FIGURE 3-15

MLE Bias Functions of Shiba's Word/Phrase Comprehension Test J1 of 55 Items Following the Logistic Model, and of Two Hypothetical Tests of the Same Number of Items Each Following the Three-Parameter Logistic Model, which Share the Same Set of Item Discrimination and Difficulty Parameters as Test J1 and with the Guessing Parameters 0.20 (Dashed Line) and 0.25 (Dotted Line), Respectively.

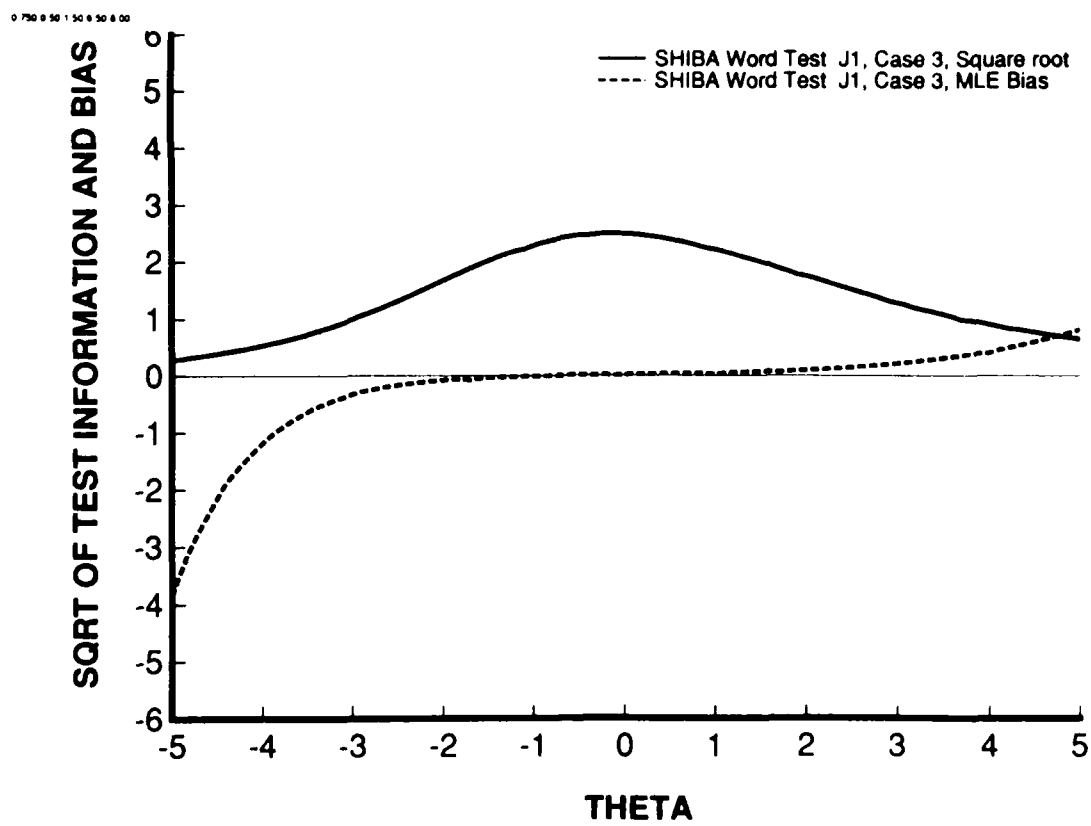
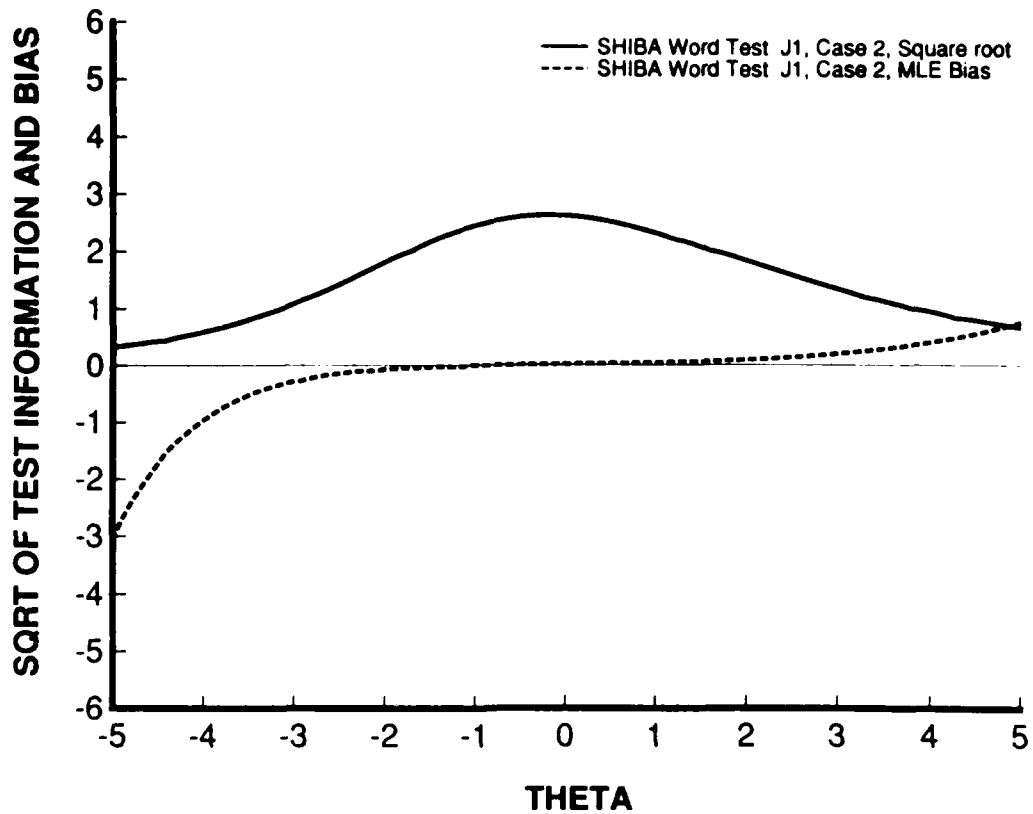


FIGURE 3-16

MLE Bias Function (Dashed Line) and the Square Root of Test Information (Solid Line) of Each of the Two Hypothetical Tests of 55 Items Each Following the Three-Parameter Logistic Model, which Share the Same Set of Item Discrimination and Difficulty Parameters as Test J1 and with the Guessing Parameters 0.20 (Dashed Line) and 0.25 (Dotted Line), Respectively.

results of the Iowa Level 11 Vocabulary Subtest, and similar conclusions can be reached.

## 4 Graded Response Level

On the graded response level, the bias function is directly given by (2.43) by replacing the general discrete response  $k_g$  to item  $g$  by the graded item score  $x_g (= 0, 1, \dots, m_g)$ .

In the homogenous case of the graded response level (Samejima, 1972), the general formula for the operating characteristic of the item score  $x_g$  is given by

$$(4.1) \quad P_{x_g}(\theta) = P_{x_g}^*(\theta) - P_{(x_g+1)}^*(\theta) ,$$

where

$$(4.2) \quad P_{x_g}^*(\theta) = \int_{-\infty}^{a_g(\theta - b_{x_g})} \psi_g(t) dt ,$$

$$(4.3) \quad -\infty = b_0 < b_1 < b_2 < \dots < b_{m_g} < b_{m_g+1} = \infty ,$$

and  $\psi_g(\theta)$  is some specified density function. When we replace the right hand side of (4.2) by that of (3.1) with  $b_g$  replaced by  $b_{x_g}$ , we have the operating characteristic of  $x_g$  in the normal ogive model on the graded response level; when we do a similar thing by using the right hand side of (1.5), we obtain the operating characteristic of  $x_g$  in the logistic model on the graded response level.

Since, in general, the graded item is more informative than the dichotomous item, we can expect smaller amounts of bias on the graded response level than on the dichotomous response level. Although the relationship between the configuration of the difficulty parameters of the  $n$  items and the amount of bias is more complicated on the graded response level, it will be easier in practice to develop a set of items which provides us with negligibly small amounts of bias for a wide range of  $\theta$ .

We shall see some examples here. In the past years, the author has been engaged in developing nonparametric approaches and methods of estimating the operating characteristics, or the conditional probabilities, given ability  $\theta$ , assigned to separate discrete item responses. In other words, these approaches and methods are based upon no assumptions concerning the mathematical forms of those operating characteristics. In so doing, the asymptotic normal property of the maximum likelihood estimate (MLE), i.e., the fact that, as the number of items increases, the conditional distribution of MLE, given  $\theta$ , approaches normality with  $\theta$  and the inverse of the square root of the test information function as the two parameters, is fully utilized. A set of simulated data has been used for testing these approaches and methods, in which 35 graded test items following the normal ogive model with three item score categories each are hypothesized as the Old Test (cf. Samejima, 1977, 1981). Table 4-1 presents the item discrimination parameter  $a_g$  and the two item response difficulty parameters, i.e.,  $b_{x_g}$  for  $x_g = 1, 2$ , for each of the 35 hypothesized items. The square root of the test information function of this Old Test is shown as the solid curve in Figure 4-1. The bias function, which was computed through (2.43), is shown in Figure 4-2 as the solid curve. We can see in this figure that for the interval of  $\theta$  covering  $(-4, 4)$  the bias of the maximum likelihood estimate is practically zero, i.e., the MLE of ability is practically unbiased for this range of  $\theta$ . Thus one of the necessary conditions to justify the use of the asymptotic normality as the approximation for the conditional distribution of MLE, given  $\theta$ , is satisfied.

We notice in Figure 4-1 that for the range of  $\theta$ ,  $(-3, 3)$ , the square root of the test information function of this Old Test assumes approximately a constant value of 4.65, and we have already seen that for the wider range of  $\theta$  the bias function assumes, practically, zero. It is interesting to note that

TABLE 4-1

Item Discrimination Parameter  $a_g$  and Two Item Difficulty Parameters  $b_{x_g}$ ,  $x_g = 1, 2$ ,  
for Each of the Thirty-Five Graded Test Items of the Old Test.

Item $g$	$a_g$	$b_1$	$b_2$
1	1.8	-4.75	-3.75
2	1.9	-4.50	-3.50
3	2.0	-4.25	-3.25
4	1.5	-4.00	-3.00
5	1.6	-3.75	-2.75
6	1.4	-3.50	-2.50
7	1.9	-3.00	-2.00
8	1.8	-3.00	-2.00
9	1.6	-2.75	-1.75
10	2.0	-2.50	-1.50
11	1.5	-2.25	-1.25
12	1.7	-2.00	-1.00
13	1.5	-1.75	-0.75
14	1.4	-1.50	-0.50
15	2.0	-1.25	-0.25
16	1.6	-1.00	0.00
17	1.8	-0.75	0.25
18	1.7	-0.50	0.50

Item $g$	$a_g$	$b_1$	$b_2$
19	1.9	-0.25	0.75
20	1.7	0.00	1.00
21	1.5	0.25	1.25
22	1.8	0.50	1.50
23	1.4	0.75	1.75
24	1.9	1.00	2.00
25	2.0	1.25	2.25
26	1.6	1.50	2.50
27	1.7	1.75	2.75
28	1.4	2.00	3.00
29	1.9	2.25	3.25
30	1.6	2.50	3.50
31	1.5	2.75	3.75
32	1.7	3.00	4.00
33	1.8	3.25	4.25
34	2.0	3.50	4.50
35	1.4	3.75	4.75

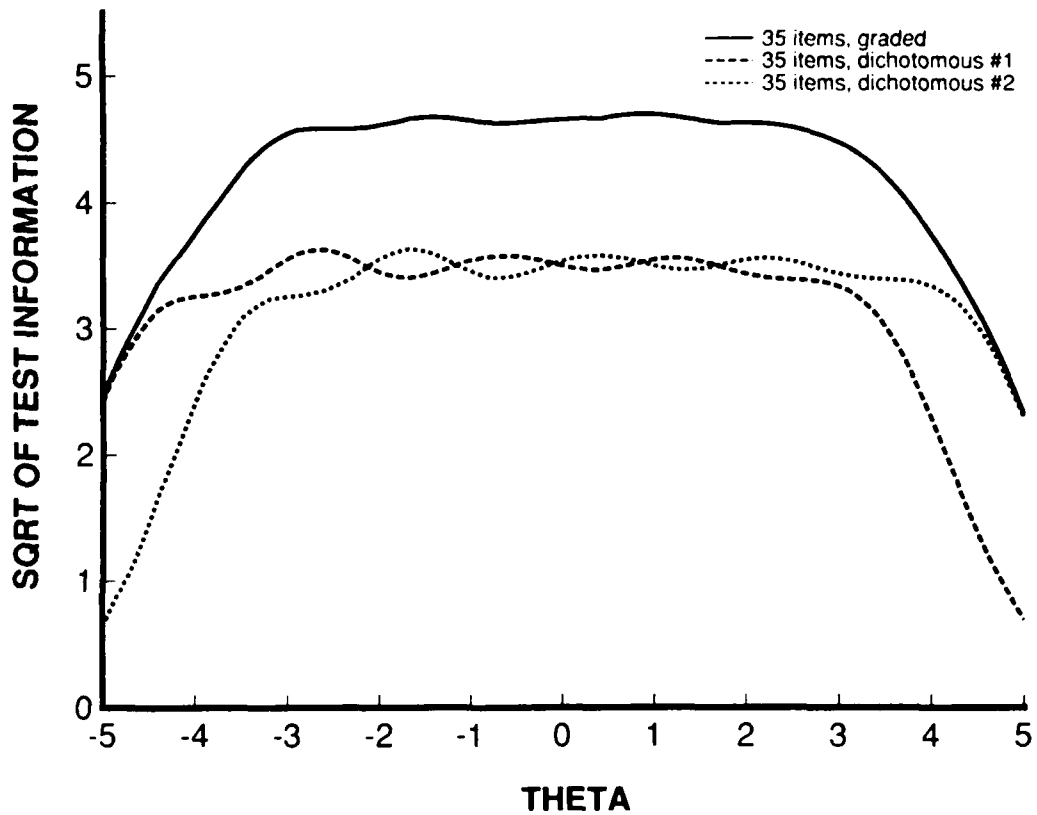


FIGURE 4-1

Square Roots of the Test Information Functions for the Old Test of 35 Graded Items (Solid Line), for the Two Sets of 35 Redichotomized Items Using the First Set (Dashed Line) and Second Set (Dotted Line) of the Difficulty Parameters of the Old Test, Respectively, Following the Normal Ogive Method.

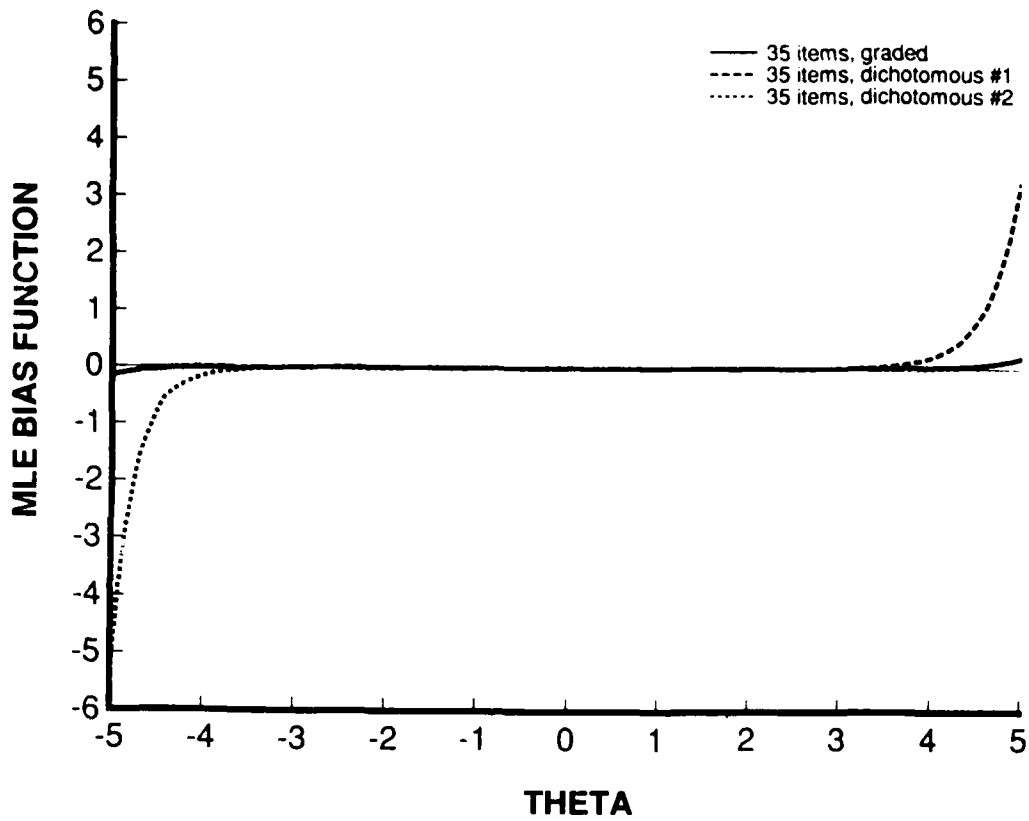
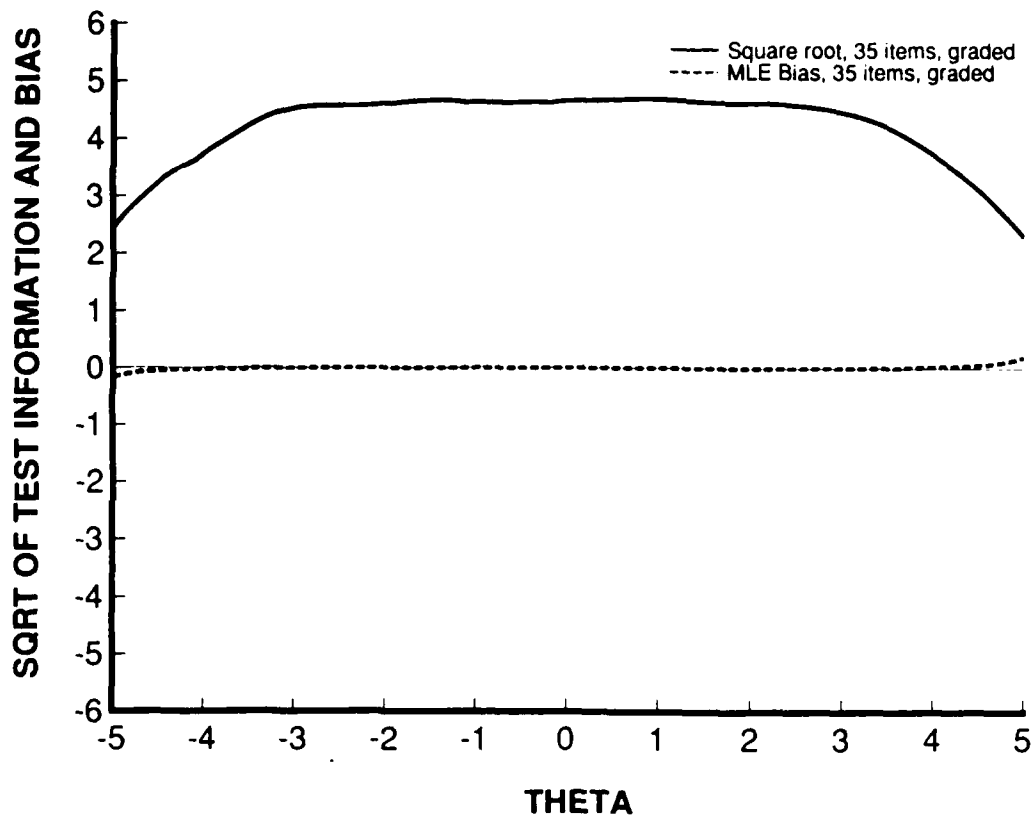
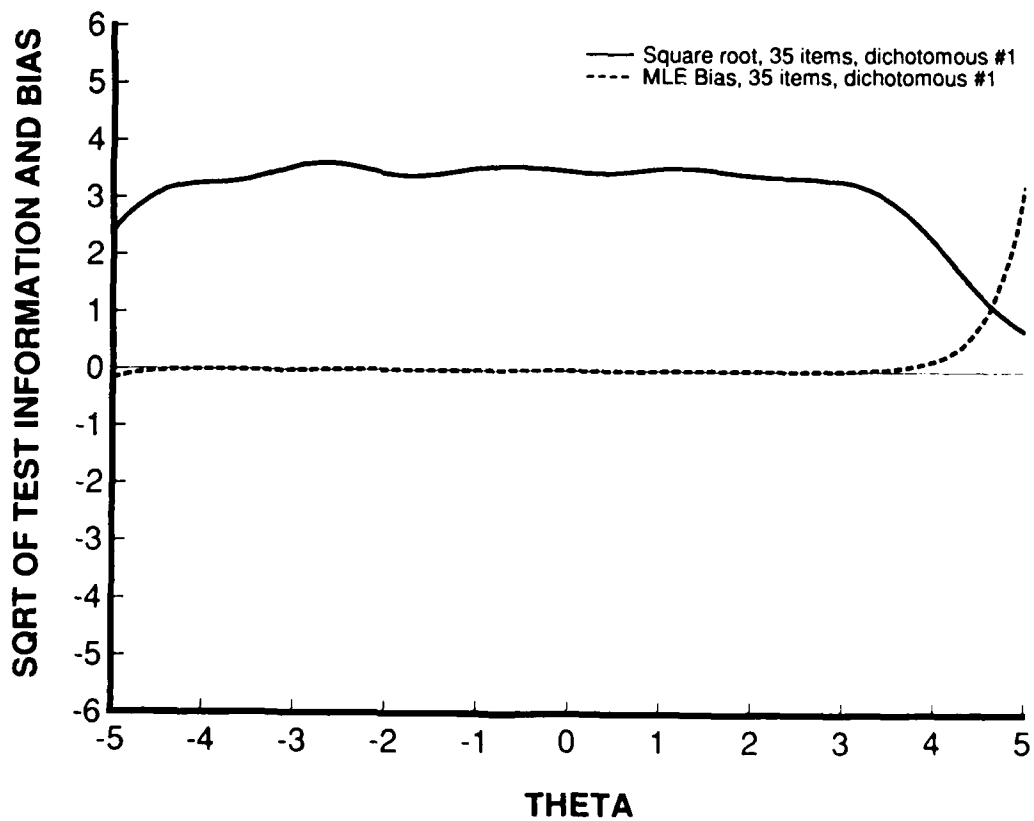


FIGURE 4-2

Bias Functions of the Old Test of 35 Graded Items (Solid Line) and of the Two Sets of 35 Redichotomized Items Using the First Set (Dashed Line) and the Second Set (Dotted Line) of the Difficulty Parameters of the Old Test, Respectively, Following the Normal Ogive Model.



0.750 0.50 1.50 6.50 8.00



0.750 0.50 1.50 6.50 8.00



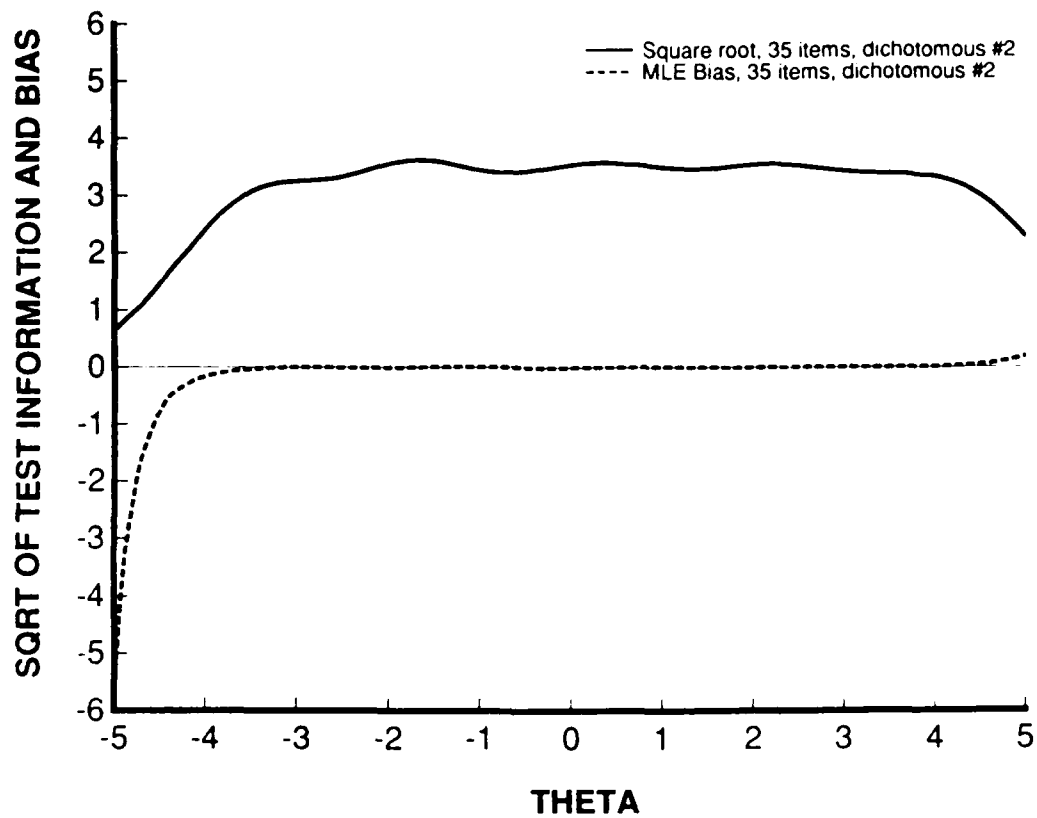


FIGURE 4-3

Comparison of the Square Root of the Test Information Function (Solid Line) and the Bias Function (Dashed Line) of Each of the Three Tests, i.e., the Old Test (First Graph) and the Two Sets of Redichotomized Items Using the First (Second Graph) and the Second (Third Graph) Sets of Difficulty Parameters, Respectively.

the bias starts showing up both positively and negatively when the square root of test information drops lower than a critical value, which is approximately 3.2, or the test information function drops lower than approximately 10. In order to pursue this relationship, two more sets of these two functions are also shown by dashed and dotted curves in Figures 4-1 and 4-2. These two sets were created by redichotomizing the graded items of the Old Test, using the first and second sets of the difficulty parameters in Table 4-1, respectively. We can see that for the wide range of  $\theta$  the square root of test information is substantially less than that of the original Old Test, which is the natural consequence of redichotomizing the items. It is noticed that for each of these two, the square root of the test information function is barely greater than 3.2 for a wide range of  $\theta$ , and the bias is still practically nil. Again the bias appears both positively and negatively when the square root of the test information function drops lower than approximately 3.2. In order to make this observation easier, both the square root of the test information function and the bias function are plotted together in Figure 4-3 for each of the three hypothetical tests, by solid and dashed lines, respectively. If we tolerate biases of up to  $\pm 0.1$ , as we did earlier, then the critical value of the square root of test information will approximately be 2.75, or that of the test information function approximately 7.5. When the square root of test information drops to less than 2.0, the bias turns out to be substantially large.

It is interesting to note that, in these examples, the amount of information required to make the bias negligibly small is larger than those observed in the previous examples, i.e., those of Iowa Level 11 Vocabulary Subtest and Shiba's Test J1. This has something to do with the fact that in the Old Test there are only 35 test items with the average discrimination parameter as high as 1.70, while there are as many as 43 and 55 items in the Iowa Level 11 Vocabulary Subtest and Shiba's J1 Test with the average values of discrimination parameters 0.601 and 0.538, respectively. We shall investigate the effect of discrimination parameters on the amount of bias from a somewhat different angle in the following section.

## 5 Effect of the Discrimination Parameter

In the previous sections, we have seen from several examples that, if the amount of test information is substantially large, the amount of bias is negligibly small, and it looks as if there were a critical value of the square root of test information to realize this approximate unbiasedness of the maximum likelihood estimate. This critical value differs from test to test, however, as we have observed in the examples of the Old Test, the Iowa Level 11 Vocabulary Subtest and Shiba's Test J1. On the dichotomous response level, the effect of the configuration of the difficulty parameters in a test can be seen fairly straightforwardly from the formulae of the bias function for different mathematical models, as we have observed earlier. In this section, we shall pursue the effect of the discrimination parameters on the bias function. In so doing, we choose tests of equivalent items, i.e., each of which consists of test items whose item characteristic functions are identical. We have seen earlier that in such a case, a substantial amount of bias starts showing up as  $\theta$  departs from the common value of difficulty parameters.

In order to simplify the notation, in this section, we use  $P$ ,  $Q$ ,  $I_\theta$ ,  $\psi$  and  $\phi$  to indicate  $P_\theta(\theta)$ ,  $Q_\theta(\theta)$ ,  $I_\theta(\theta)$ ,  $\Psi_\theta(\theta)$  and

$$(5.1) \quad \phi = \phi_\theta\{a(\theta - b)\} = (2\pi)^{-1/2} e^{-a^2(\theta - b)^2/2},$$

which are common for all the  $n$  items, where  $a = a_1 = \dots = a_n$  and  $b = b_1 = \dots = b_n$ . In the normal ogive model, (3.4) can be simplified for the set of  $n$  equivalent items so that we obtain

$$(5.2) \quad B(\theta) = (1/2)PQ(\theta - b)n^{-1}\phi^{-2},$$

because of (1.6), (1.7) and (3.2). We have for the partial derivative of  $B(\theta)$  with respect to  $a$

$$(5.3) \quad \frac{\partial}{\partial a} B(\theta) = (2n)^{-1}(\theta - b)\phi^{-4}[\phi^2 \frac{\partial}{\partial a}(PQ) - 2\phi PQ \frac{\partial}{\partial a}\phi] .$$

By virtue of the fact that

$$(5.4) \quad \frac{\partial}{\partial a}(PQ) = \phi(\theta - b)(Q - P)$$

and

$$(5.5) \quad \frac{\partial}{\partial a}\phi = \phi[-a(\theta - b)^2] ,$$

we can write for the last factor of the right hand side of (5.3)

$$(5.6) \quad \phi^2 \frac{\partial}{\partial a}(PQ) - 2\phi PQ \frac{\partial}{\partial a}\phi = 2\phi^2(\theta - b)PQ[a(\theta - b) - \frac{1}{2}\{-\frac{\phi}{P} + \frac{\phi}{Q}\}] .$$

We notice that the second term on the parenthesis on the right hand side of (5.6) equals  $\frac{1}{2}[E[u | u < a(\theta - b)] + E[u | u \geq a(\theta - b)]]$ , since we have

$$(5.7) \quad \begin{aligned} E[u | u < a(\theta - b)] &= \int_{-\infty}^{a(\theta - b)} u\phi(u) du / P \\ &= -\phi / P , \\ &= [-\phi(u) |_{-\infty}^{a(\theta - b)}] / P \end{aligned}$$

and

$$(5.8) \quad \begin{aligned} E[u | u \geq a(\theta - b)] &= \int_{a(\theta - b)}^{\infty} u\phi(u) du / Q \\ &= [-\phi(u) |_{a(\theta - b)}^{\infty}] / Q \\ &= \phi / Q . \end{aligned}$$

It is obvious that this average of the two expectations of  $u$  equals zero when  $\theta = b$  and assumes negative and positive values when  $\theta < b$  and  $\theta > b$ , respectively. In addition, we obtain

$$(5.9) \quad \frac{1}{2}\{-\frac{\phi}{P} + \frac{\phi}{Q}\} \begin{cases} > a(\theta - b) & \theta < b \\ = a(\theta - b) & \theta = b \\ < a(\theta - b) & \theta > b \end{cases} .$$

To prove this, since we can write for the item information function in the normal ogive model

$$(5.10) \quad I_y = a^2 \phi^2 (PQ)^{-1} ,$$

its first derivative  $I'_g$  with respect to  $\theta$  is given by

$$(5.11) \quad I'_g = 2aI_g \left[ \frac{1}{2} \left\{ -\frac{\phi}{P} + \frac{\phi}{Q} \right\} - a(\theta - b) \right] .$$

Setting (5.11) equal to zero, we obtain  $\theta = b$ . From (5.11), we can see that the second derivative  $I''_g$  of the item information function with respect to  $\theta$  assumes  $4a^2\pi^{-2}(2 - \pi)$  at  $\theta = b$ , which is negative. Thus  $\theta = b$  is the point of  $\theta$  at which  $I_g$  is maximal, and  $I'_g$  assumes positive values for  $\theta < b$  and negative values for  $\theta > b$ . Equation (5.9) is the direct consequence of this fact.

Figure 5-1 presents the square root of the test information function for each of the five examples following the normal ogive model by a solid line and dashed lines of various lengths. In these five examples, the numbers of equivalent items are uniformly 30, and the common values of the difficulty parameter are all 0.0. The common values of the discrimination parameter differ for different tests, however, i.e., they assume 0.4, 0.7, 1.0, 1.5 and 2.0, respectively. The five bias functions for these five hypothetical tests are shown in Figure 5-2, using the same set of solid and dashed lines.

We can see in these figures how rapidly the amount of bias increases when the common discrimination parameter is large, in both negative and positive directions as  $\theta$  departs from zero, at which the square root of test information is maximal, especially when  $a_g = 2.0$ . It is also noted that, taking the criterion of  $\pm 0.1$  again, in order to keep the practical unbiasedness the square root of test information must be as large as 4.0 when  $a_g = 2.0$ , while it can be as small as 1.3 when  $a_g = 0.4$ . For the intermediate values of the discrimination parameter, i.e., for  $a_g = 0.7, 1.0, 1.5$ , the corresponding criterion values of the square root of test information are approximately 2.0, 2.5 and 3.0, respectively.

In the logistic model, we can rewrite (1.4) for the set of  $n$  equivalent items to obtain

$$(5.12) \quad \begin{aligned} B(\theta) &= DnaI_g \left\{ \Psi - \frac{1}{2} \right\} [nI_g]^{-2} \\ &= \left\{ \Psi - \frac{1}{2} \right\} [nDa\Psi(1 - \Psi)] , \end{aligned}$$

by virtue of (1.3), (1.5), (1.6) and the fact that

$$(5.13) \quad \Psi'_g(\theta) = Da_g\Psi_g(\theta)[1 - \Psi_g(\theta)] ,$$

where  $\Psi'_g(\theta)$  indicates the first derivative of  $\Psi_g(\theta)$  with respect to  $\theta$ . Since we have

$$(5.14) \quad \frac{\partial}{\partial a_g} \Psi_g(\theta) = D(\theta - b_g)\Psi_g(\theta)[1 - \Psi_g(\theta)] ,$$

the numerator of the partial derivative of  $B(\theta)$  with respect to  $a$  can be written as

$$(5.15) \quad \begin{aligned} nD^2a\Psi^2[1 - \Psi]^2(\theta - b) - \left[ \Psi - \frac{1}{2} \right] nD^2a\Psi[1 - \Psi][1 - 2\Psi](\theta - b) \\ = nD^2a\Psi[1 - \Psi] \left[ \frac{1}{2} - \Psi(1 - \Psi) \right] (\theta - b) . \end{aligned}$$

Since all the factors on the right hand side of (5.15) are positive except for the last one, we can conclude that the amount of bias equals zero at  $\theta = b$  regardless of the value of  $a$ , and increases in the positive and negative directions for  $\theta > b$  and  $\theta < b$ , respectively.

Figures 5-3 and 5-4 present the square root of the test information function and the bias function for each of the five hypothetical tests following the logistic model, respectively, which share the same number of items and the parameter values as those five hypothetical tests following the normal ogive model. These results are very similar to those obtained for the normal ogive model, except for the fact that the intervals of practical unbiasedness are a little smaller.

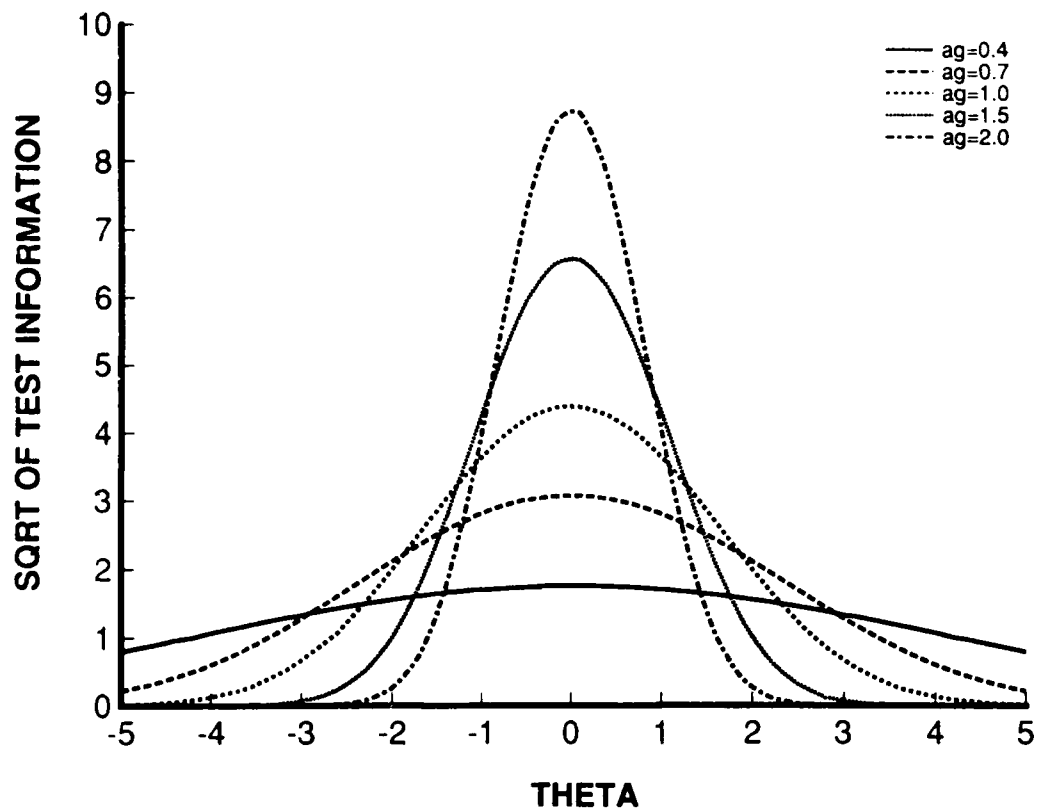
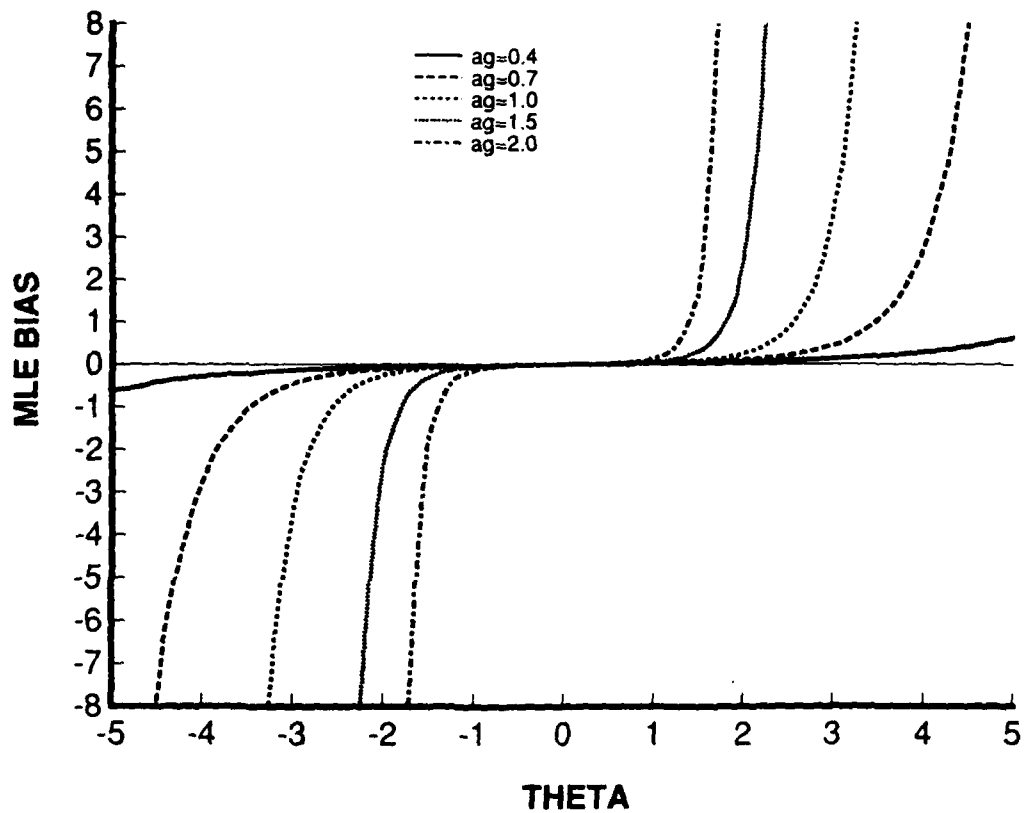


FIGURE 5-1

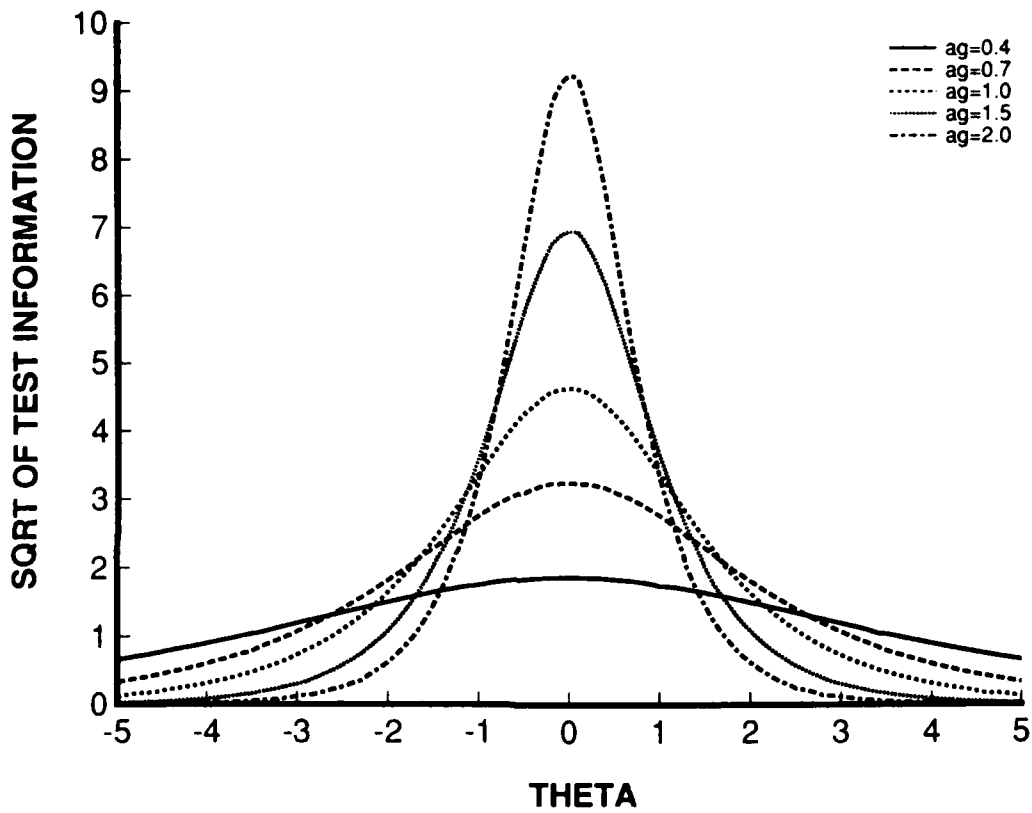
Square Roots of Test Information for the Five Hypothetical Tests of 30 Equivalent Items Following the Normal Ogive Model. The Common Values of the Difficulty Parameter Are Uniformly 0.0, and Those of the Discrimination Parameter are 0.4, 0.7, 1.0, 1.5, and 2.0, Respectively.



0 750 0 50 1 50 6 50 8 00

FIGURE 5-2

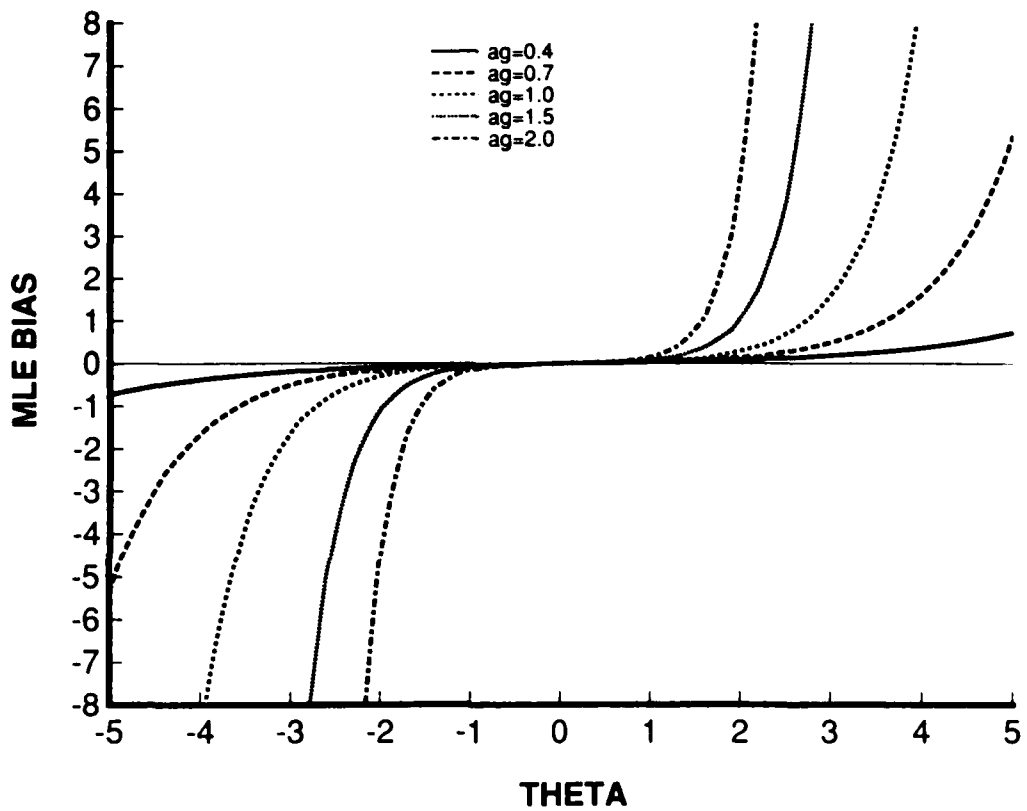
Bias Functions for the Five Hypothetical Tests of 30 Equivalent Items Following the Normal Ogive Model. The Common Values of the Difficulty Parameter Are Uniformly 0.0 , and Those of the Discrimination Parameter are 0.4 , 0.7 , 1.0 , 1.5 , and 2.0 , Respectively.



0 790 0 50 1 50 6 50 8 00

FIGURE 5-3

Square Roots of Test Information for the Five Hypothetical Tests of 30 Equivalent Items Following the Logistic Model. The Common Values of the Difficulty Parameter Are Uniformly 0.0 , and Those of the Discrimination Parameter are 0.4 , 0.7 , 1.0 , 1.5 , and 2.0 , Respectively.



0 50 0 50 1 50 8 50 8 00

FIGURE 5-4

Bias Functions for the Five Hypothetical Tests of 30 Equivalent Items Following the Logistic Model. The Common Values of the Difficulty Parameter Are Uniformly 0.0 and Those of the Discrimination Parameter are 0.4 , 0.7 , 1.0 , 1.5 and 2.0, Respectively.



## 6 Effect of the Number of Items

It is obvious from (1.7), (2.43) and (2.46) that the number of items in a test affects the amount of bias through the test information function, in the negative way such that its increase causes the decrease in the amount of bias. We have also observed from our examples that, even if the amounts of test information are the same for two different tests, they may not share the same amount of bias. In this regard, we have seen that the values of the item discrimination parameters affect the amount of bias in the positive direction. It has also been pointed out that the configuration of the difficulty parameters in a test affects the bias function.

In order to demonstrate these effects further, in this section, we shall observe the effect of the number of items using different numbers of equivalent items, each of which follows the constant information model (Samejima, 1979) on the dichotomous response level. The item characteristic function in the constant information model is defined by

$$(6.1) \quad P_g(\theta) = \sin^2\{\alpha_g(\theta - \beta_g) + (\pi/4)\} .$$

where  $\alpha_g$  and  $\beta_g$  are the item discrimination and difficulty parameters, respectively. From (6.1) we obtain

$$(6.2) \quad Q_g(\theta) = \cos^2\{\alpha_g(\theta - \beta_g) + (\pi/4)\} .$$

This model provides us with a constant amount of item information, i.e.,

$$(6.3) \quad I_g(\theta) = 4\alpha_g^2 ,$$

for the interval of  $\theta$  such that

$$(6.4) \quad -\pi(4\alpha_g)^{-1} + \beta_g < \theta < \pi(4\alpha_g)^{-1} + \beta_g .$$

Since we have

$$(6.5) \quad P'_g(\theta) = 2\alpha_g[P_g(\theta)Q_g(\theta)]^{1/2}$$

and

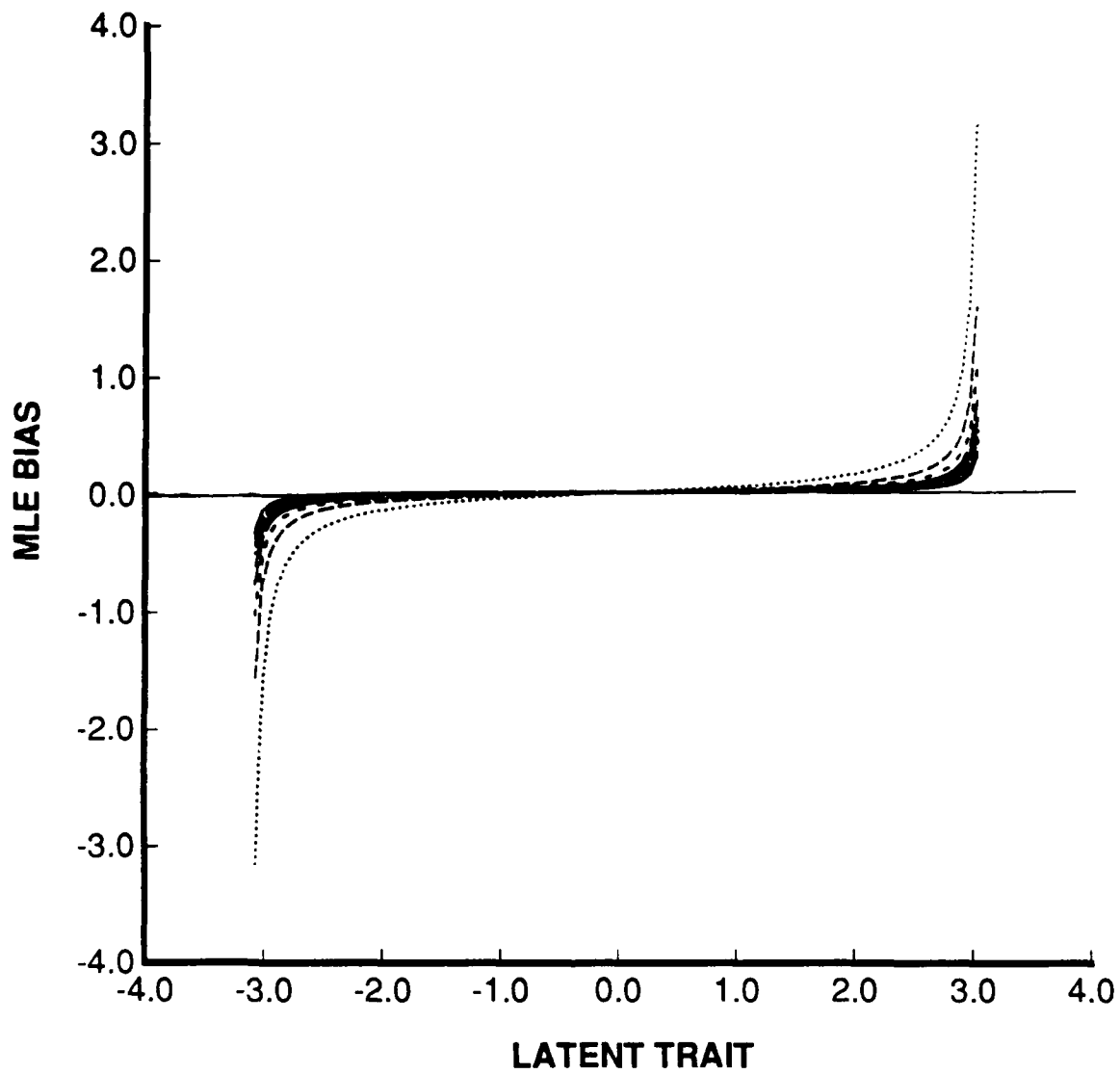
$$(6.6) \quad P''_g(\theta) = 2\alpha_g^2[Q_g(\theta) - P_g(\theta)] ,$$

substituting these and (6.3) into (2.46) and rearranging, we can write for the bias function in the constant information model

$$(6.7) \quad B(\theta) = 2\{I(\theta)\}^{-2} \sum_{g=1}^n \alpha_g^3 \{P_g(\theta) - Q_g(\theta)\} \{P_g(\theta)Q_g(\theta)\}^{-1/2} .$$

For the set on  $n$  equivalent items, we can simplify (6.7) into the form

$$(6.8) \quad B(\theta) = \{8n\alpha_g\}^{-1} \{P_g(\theta) - Q_g(\theta)\} \{P_g(\theta)Q_g(\theta)\}^{-1/2} .$$



1.000 0.50 1.00 10.00 6.055

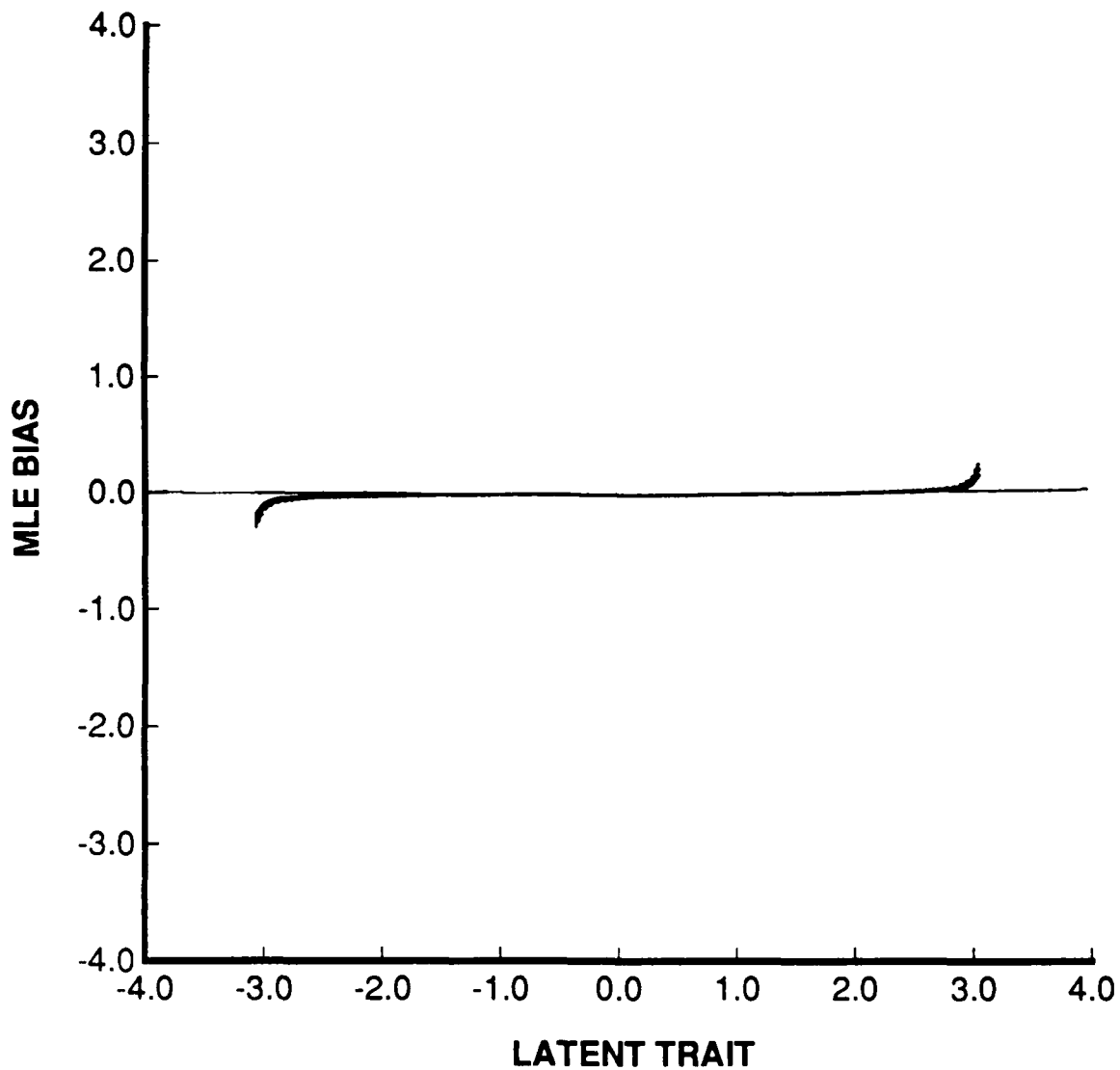


FIGURE 6-1

MLE Bias Function for Each of the Twenty Different Values of  $n$ , Which Starts with 10, Increases by 10 Successively, and Ends with 200. The Common Item Parameters for These Equivalent Items Are Given by  $\alpha_g = 0.25$  and  $\beta_g = 0.00$ , and, Therefore, in Each Set of Equivalent Items, the Test Information Function Assumes a Constant Value,  $0.25n$ , for the Range of  $\theta$ ,  $-\pi < \theta < \pi$ .

Figure 6-1 presents, in two graphs, this bias function for each of the twenty different values of  $n$ , which starts with 10, increases by 10 successively, and ends with 200. The common item parameters for these equivalent items are given by  $\alpha_g = 0.25$  and  $\beta_g = 0.00$ , and, therefore, in each set of equivalent items, the test information function assumes a constant value of  $0.25n$  for the range of  $\theta$ ,  $-\pi < \theta < \pi$ . We can see in this figure that the amount of bias is substantial as  $\theta$  departs from  $\beta_g = 0.00$  when  $n$  is relatively small, but it becomes negligibly small for larger values of  $n$ , especially when  $n$  exceeds 100. These results also illustrate the fact that the amount of test information alone does not control the amount of bias, since they are based upon the constant information model, and the amount of test information is a constant in each set of equivalent items for the range of  $\theta$ ,  $-\pi < \theta < \pi$ . A usefulness of the constant information model is that we can use it as the benchmark when we deal with a set of equivalent items following such frequently used mathematical models as the normal ogive model, the logistic model, Rasch model, etc. This will be shown with respect to the bias function in the following section.

## 7 Scale Transformation

It is obvious from (2.43) that the bias function belongs to a particular scale of the latent trait, and, if the scale is transformed, the function changes also. In this section, we shall see how the scale transformation affects the bias function of a test.

### 7.1 General Case of Discrete Responses

Let  $\tau$  be a strictly increasing transformation of  $\theta$ , so that we can write

$$(7.1) \quad \tau = \tau(\theta) .$$

We assume that  $\tau$  is twice differentiable with respect to  $\theta$ . Since the operating characteristic, or the conditional probability, of the discrete item response  $k_g$  is unchanged for the scale transformation, we can write

$$(7.2) \quad P_{k_g}^*(\tau) = P_{k_g}(\theta) ,$$

where  $P_{k_g}^*(\tau)$  denotes the operating characteristic of the discrete item response  $k_g$  as a function of the transformed latent trait  $\tau$ . From (7.2) we obtain for the first and second derivatives,  $P_{k_g}^{*'}(\tau)$  and  $P_{k_g}^{*''}(\tau)$ , of the operating characteristic of  $k_g$  with respect to  $\tau$

$$(7.3) \quad P_{k_g}^{*'}(\tau) = P_{k_g}'(\theta) \frac{d\theta}{d\tau}$$

and

$$(7.4) \quad P_{k_g}^{*''}(\tau) = P_{k_g}''(\theta) \left| \frac{d\theta}{d\tau} \right|^2 + P_{k_g}'(\theta) \frac{d^2\theta}{d\tau^2} .$$

From (7.3), (7.4) and the definitions of the item response information function and of the item information function, which are given by (1.1) and (1.2), respectively, we can write for the item information function  $I_g^*(\tau)$  of item  $g$  as a function of the transformed latent trait  $\tau$

$$(7.5) \quad I_g^*(\tau) = I_g(\theta) \left| \frac{d\theta}{d\tau} \right|^2 .$$

Hence we have for the test information function  $I^*(\tau)$  on the transformed scale of the latent trait

$$(7.6) \quad \begin{aligned} I^*(\tau) &= \sum_{g=1}^n I_g^*(\tau) = \sum_{g=1}^n I_g \left[ \frac{d\theta}{d\tau} \right]^2 \\ &= I(\theta) \left[ \frac{d\theta}{d\tau} \right]^2 . \end{aligned}$$

We can write from (2.43) for the bias function,  $B^*(\tau)$ , of a test after the scale transformation

$$(7.7) \quad B^*(\tau) = -\frac{1}{2} [I^*(\tau)]^{-2} \sum_{g=1}^n \sum_{k_g} P_{k_g}^{*'}(\tau) P_{k_g}^{*''}(\tau) [P_{k_g}^*(\tau)]^{-1} .$$

Substituting (7.2), (7.3), (7.4), and (7.6) into (7.7) and rearranging, we obtain

$$(7.8) \quad B^*(\tau) = B(\theta) \left[ \frac{d\theta}{d\tau} \right]^{-1} - \frac{1}{2} [I(\theta)]^{-1} \left[ \frac{d\theta}{d\tau} \right]^{-3} \frac{d^2\theta}{d\tau^2} .$$

The bias function of the transformed latent variable can be computed, therefore, from the original bias function, the original test information function and the first and second derivatives of  $\theta$  with respect to  $\tau$ .

## 7.2 Scale Transformation to Generate a Constant Test Information

In a nonparametric approach for estimating the operating characteristics of discrete item responses, the latent trait  $\theta$  is transformed to  $\tau$  in such a way that the resultant test information function  $I^*(\tau)$  assumes a constant value for the interval of  $\tau$  of interest, in which the ability levels of most subjects are included (cf. Samejima, 1981). In this way, the approximation of the conditional distribution of the maximum likelihood estimate  $\hat{\tau}$ , given  $\tau$ , by a normal distribution with the parameters  $\tau$  and  $\sigma$ , the latter of which does not depend upon  $\tau$ , becomes more justifiable. Thus we can write

$$(7.9) \quad I^*(\tau) = C^2 , \quad C > 0 .$$

Substituting this into (7.6) and rearranging, we obtain for the first derivative of  $\theta$

$$(7.10) \quad \frac{d\theta}{d\tau} = C [I(\theta)]^{-1/2} .$$

The transformation of the latent trait  $\theta$  to  $\tau$  is given by

$$(7.11) \quad \tau = \tau(\theta) = C^{-1} \int_{-\infty}^{\theta} [I(t)]^{1/2} dt + \delta ,$$

where  $\delta$  indicates a constant which determines the origin of  $\tau$ . From (7.10) we have for the second derivative of  $\theta$  with respect to  $\tau$

$$(7.12) \quad \frac{d^2\theta}{d\tau^2} = -\frac{1}{2} C^2 [I(\theta)]^{-2} I'(\theta) ,$$

where  $I'(\theta)$  indicates the first derivative of the test information function  $I(\theta)$  with respect to  $\theta$ , for which we can write

$$(7.13) \quad I'(\theta) = \sum_{g=1}^n \sum_{k_g} P_{k_g}'(\theta) [2P_{k_g}(\theta) P_{k_g}''(\theta) - \{P_{k_g}'(\theta)\}^2] [P_{k_g}(\theta)]^{-2} .$$

Substituting (7.12) and (7.13) into (7.8) and rearranging, we obtain

$$(7.14) \quad B^*(\tau) = B(\theta)C^{-1}[I(\theta)]^{1/2} + \frac{1}{4}C^{-1}[I(\theta)]^{-3/2}I'(\theta) .$$

Thus in this specific situation (7.8) is simplified to include only the original bias function, the original test information function and its derivative, and the constant square root of test information after the latent variable has been transformed.

Let  $K(\theta)$  denote the square root of the test information function  $I(\theta)$ . Thus we can write

$$(7.15) \quad K(\theta) = I(\theta)^{1/2} .$$

Since we have

$$(7.16) \quad I'(\theta) = 2K(\theta)K'(\theta) ,$$

where  $K'(\theta)$  is the derivative of  $K(\theta)$  with respect to  $\theta$ , we can rewrite (7.14) in the form

$$(7.17) \quad B^*(\tau) = B(\theta)C^{-1}K(\theta) + \frac{1}{2}C^{-1}[K(\theta)]^{-2}K'(\theta) .$$

We notice from (7.17) that, in a typical situation where the square root of test information is unimodal, as is exemplified by those functions obtained for the Iowa Level 1. Vocabulary Subtest and Shiba's Word/Phrase Comprehension Test J1, which are shown in Figure 3-1, the amount of bias is decreased by the transformation of  $\theta$  to  $\tau$  for extreme values of the transformed latent variable where there used to be substantial amounts of bias either in negative or positive direction (cf. Figure 3-2). If we set the value of  $C$  in such a way that, for a meaningful interval of  $\theta$ , the values of the two endpoints are practically unchanged over the transformation from  $\theta$  to  $\tau$  in order to avoid overall radical changes of scale values, then the factor  $C^{-1}K(\theta)$  by which the original bias function  $B(\theta)$  is multiplied, takes on values less than unity for extreme values of  $\tau$ , causing reduction in the amount of bias. In addition to this fact, the second term of the right hand side of (7.17) further reduces negative biases as we go toward extremely lower levels of the transformed scale and decreases positive biases as we go toward the other extreme, for  $K'(\theta)$  is positive on lower levels of  $\tau$  and negative on higher levels, respectively, and  $K(\theta)$  assumes small positive values on both.

Figure 7-1 presents the constant square root  $C$  of test information  $I^*(\tau)$  by dashed lines, in comparison with the square root of the original test information function  $I(\theta)$  which is drawn by a solid line, of the Iowa Level 11 Vocabulary Subtest. The original square root of the test information function is based upon the normal ogive model and has already been shown in Figure 3-1. The values of  $C$  and  $\delta$  in (7.12) were chosen in such a way that  $\tau$  is set equal to  $\theta$  at  $\theta = \pm 4.0$ . In this way, we can avoid radical changes between the two sets of scale values. As the result, the constant square root of test information  $C$  turned out to be approximately 2.22617674. We can see in Figure 7-1 that for the interval,  $(-4.0, 4.0)$ , the areas under the two square roots of test information are equal. The resulting bias function  $B^*(\tau)$  is shown in Figure 7-2 by a dashed line, in comparison with the original bias function  $B(\theta)$ . We can see a substantial decrease in the amount of bias caused by the scale transformation.

Figures 7-3 and 7-4 present the corresponding results for Shiba's Word/Phrase Comprehension Test J1. The transformation of  $\theta$  to  $\tau$  was made following the same strategy that was used for the Iowa Subtest. The resultant constant square root of test information  $C$  is approximately 2.39633860. We can see in this result that, after the scale transformation, the maximum likelihood estimate of  $\tau$  is practically unbiased, if we accept the criterion of  $\pm 0.1$  as we did before.

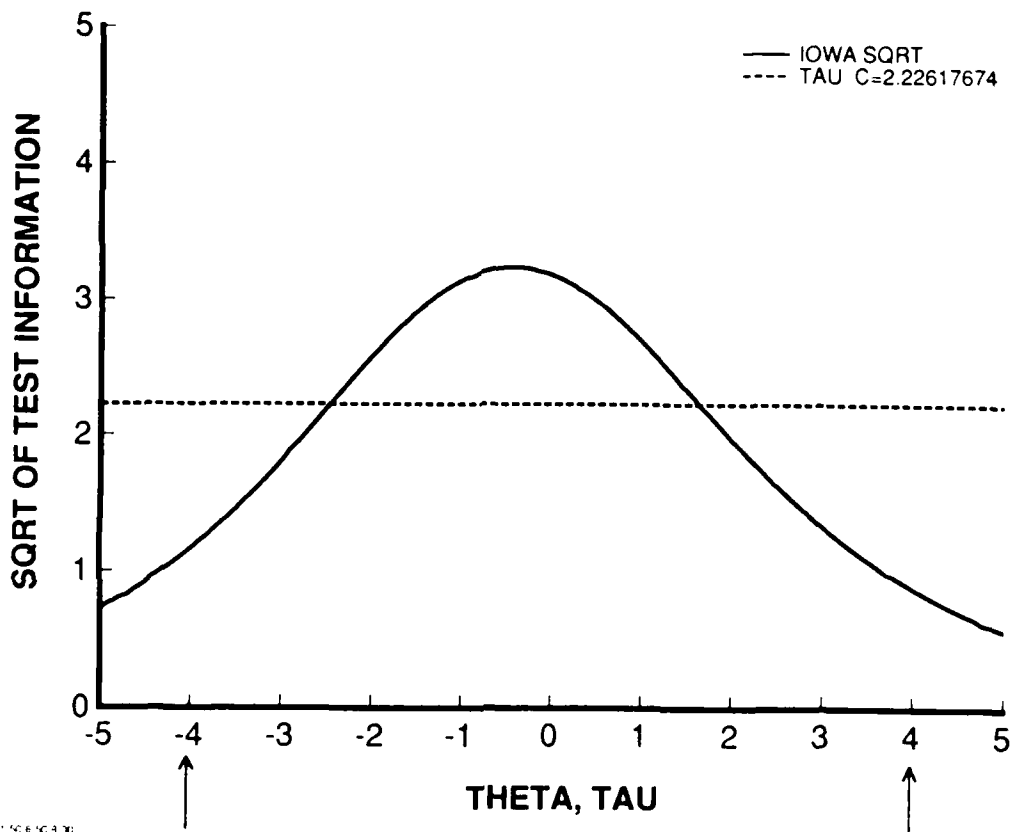
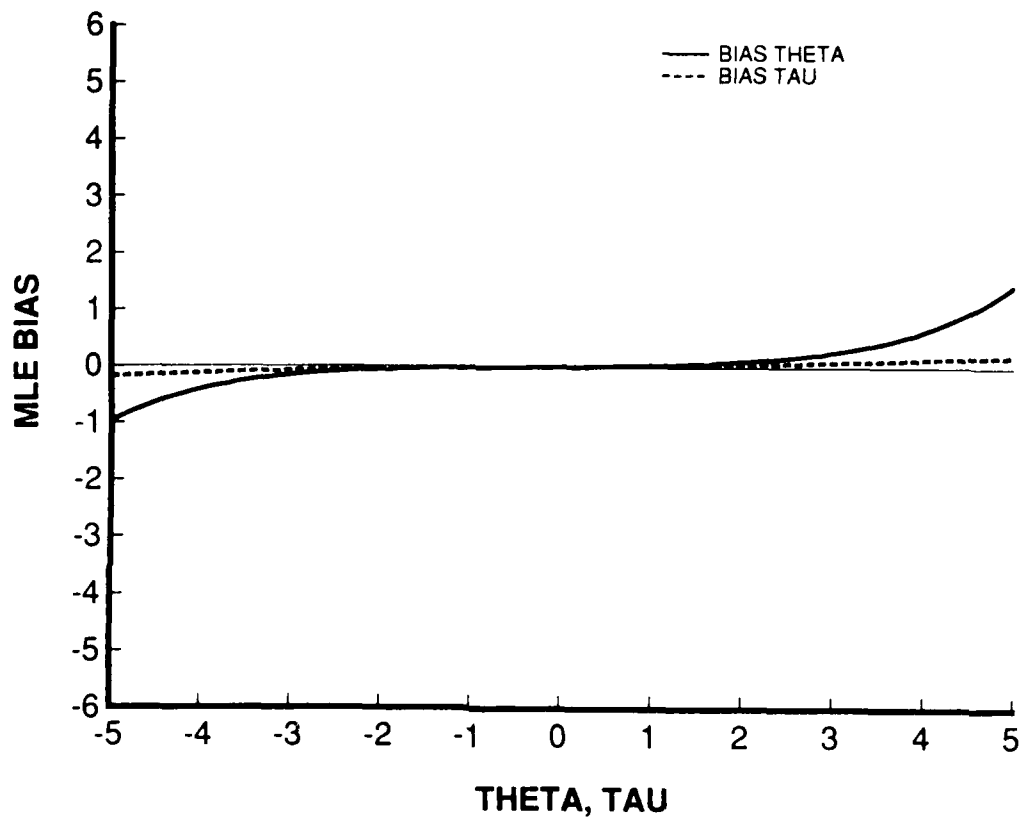


FIGURE 7-1

Square Roots of Test Information of the Iowa Level 11 Vocabulary Subtest Before (Solid Line) and After (Dashed Line) the Scale Transformation. Transformation is made in such a way that the two scales match at  $-4.0$  and  $+4.0$ .



0 7 5 0 0 5 0 1 5 0 8 5 0 8 0 0

FIGURE 7-2

MLE Bias Function of the Iowa Level 11 Vocabulary Subtest as a Function of the Transformed Latent Variable  $\tau$  (Dashed Line) in Comparison to the Original MLE Bias Function (Solid Line) of  $\theta$ .



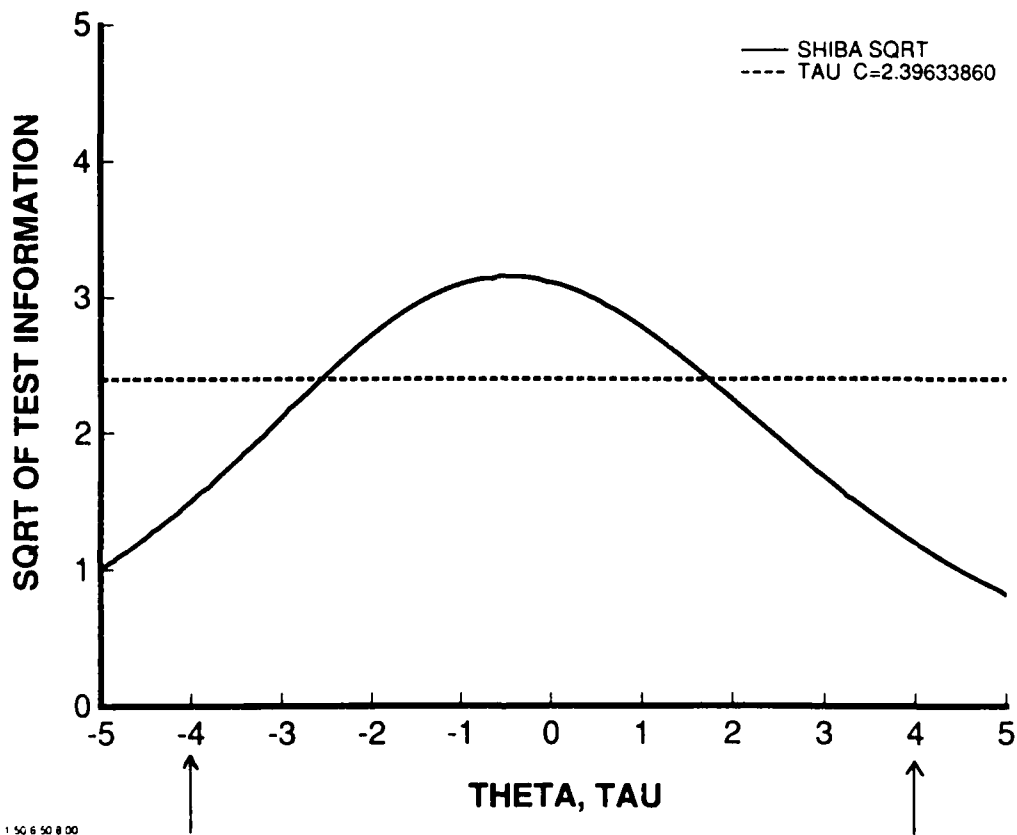
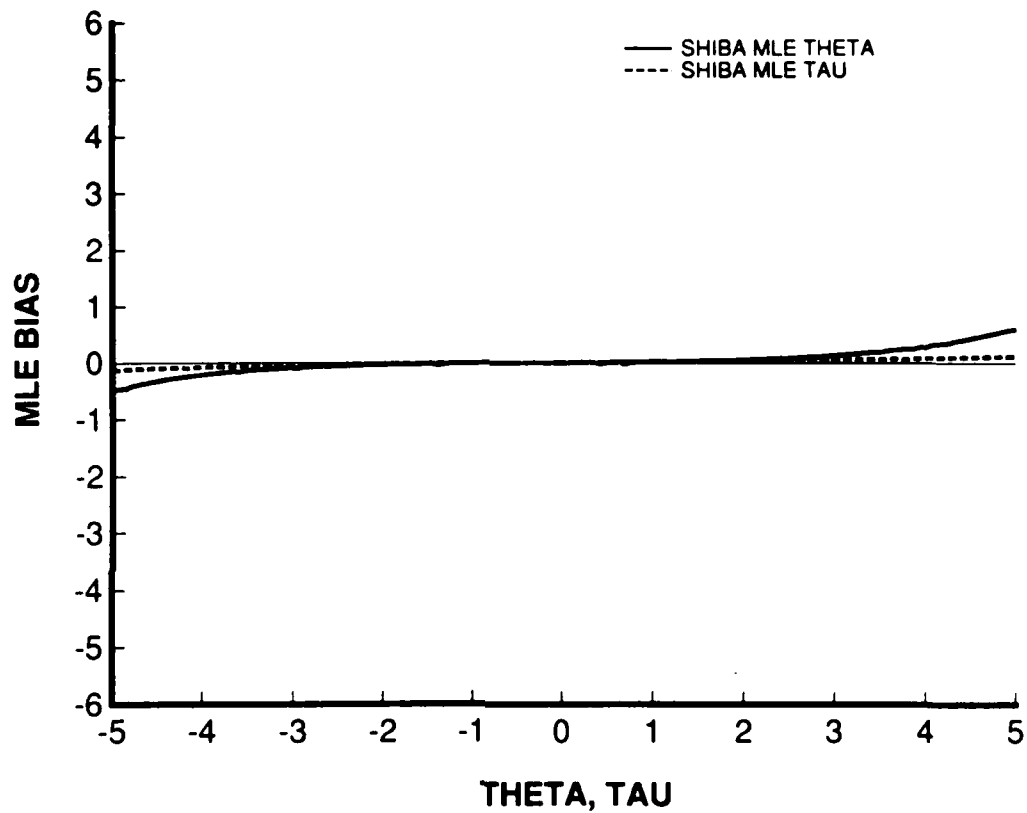


FIGURE 7-3

Square Roots of Test Information of Shiba's Word/Phrase Comprehension Test J1 Before (Solid Line) and After (Dashed Line) the Scale Transformation. Transformation is made in such a way that the two scales match at  $-4.0$  and  $+4.0$ .



0 150 0 50 1 50 6 50 8 00

FIGURE 7-4

MLE Bias Function of Shiba's Word/Phrase Comprehension Test J1 as a Function of the Transformed Latent Variable  $\tau$  (Dashed Line) in Comparison to the Original MLE Bias Function (Solid Line) of  $\theta$ . Transformation of  $\theta$  to  $\tau$  Is Made by a Polynomial Approximation.

It has been demonstrated that the square root of the test information can be approximated by a polynomial of a suitable degree obtained by the method of moments, which proves to be also the least square solution (cf. Samejima and Livingston, 1979). It has also been shown that with many different sets of data such approximations have worked well (e.g. Samejima, 1981, 1984a). For the purpose of illustration, the polynomial approximation was used with the Iowa Level 11 Vocabulary Subtest, and the resulting scale transformation is given by

$$(7.18) \quad \tau = 0.3777014 + 1.4301120\theta - 0.0528854\theta^2 - 0.0408096\theta^3 + 0.0029404\theta^4 \\ + 0.0011037\theta^5 - 0.0000858\theta^6 - 0.0000146\theta^7 + 0.0000010\theta^8$$

In this scale transformation, the same strategy was taken as before, so that  $\tau(\theta) = \theta$  at  $\theta = \pm 4.0$ . The constant square root of the test information function of  $\tau$  turned out to be 2.231709, which is very close to the corresponding value of 2.22617674, which was obtained without the polynomial approximation. The bias function  $B^*(\tau)$  thus obtained is shown in Figure 7-5 by a dashed line, in comparison with the original  $B(\theta)$  which is drawn by a solid line. We can see that this result is practically identical with the one obtained without the polynomial approximation, which is shown in Figure 7-2.

Figure 7-6 presents the three separate scale transformations of the Iowa Level 11 Vocabulary Subtest, of Shiba's Test J1 and of the Iowa Subtest with the polynomial approximation, by solid, dashed and dotted lines, respectively. Actually, we can only see two curves, for the dotted curve practically coincides with the solid curve.

### 7.3 Equivalent Items on the Dichotomous Response Level

We have seen in a previous section how the amount of bias decreases as the number of items increases, using the example of equivalent items on the dichotomous response level, which follow the constant information model. It should be noted that the corresponding set of bias functions for equivalent items following any mathematical model, which provides us with a strictly increasing item characteristic function with zero and unity as its two asymptotes, can be produced from these results by an appropriate strictly increasing scale transformation. Let  $\tau = \tau(\theta)$  be such a transformation of  $\theta$ , and  $P_g^*(\tau)$  denote the item characteristic function following one of such models. Setting

$$(7.19) \quad P_g^*(\tau) = P_g(\theta) ,$$

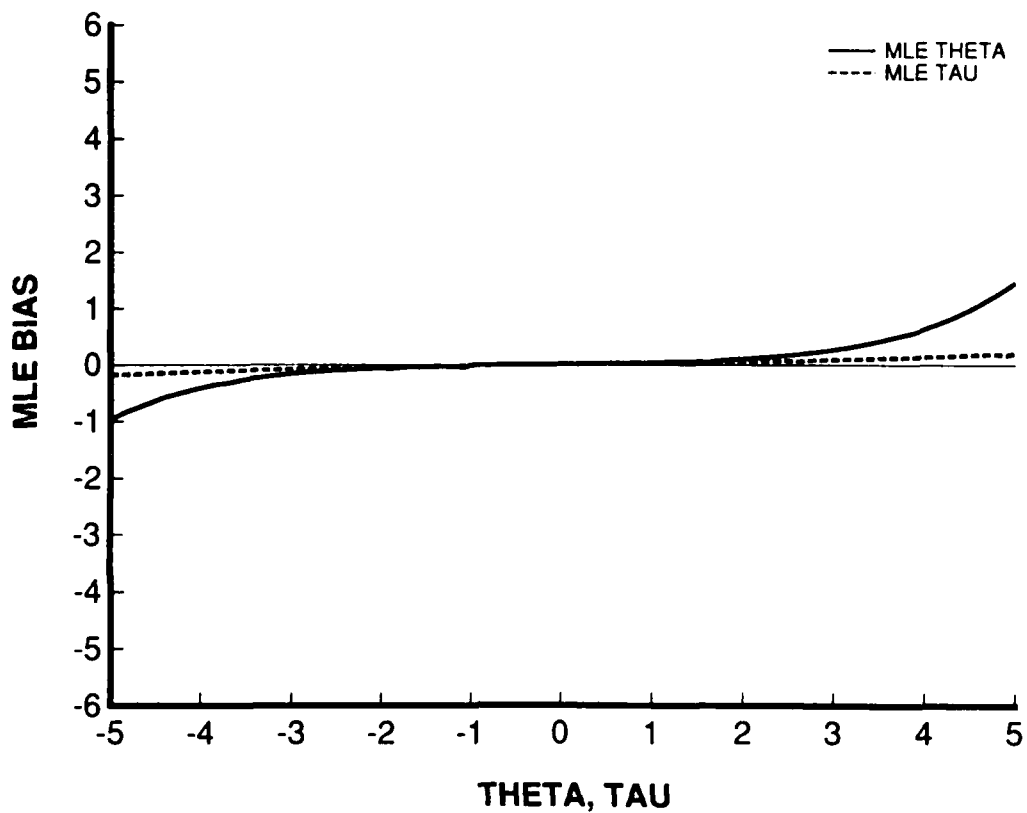
we obtain

$$(7.20) \quad P_g^{*'}(\tau) = P_g'(\theta) \frac{d\theta}{d\tau} = 2\alpha_g \{P_g(\theta)Q_g(\theta)\}^{1/2} \frac{d\theta}{d\tau} ,$$

$$(7.21) \quad P_g^{*''}(\tau) = P_g''(\theta) \left\{ \frac{d\theta}{d\tau} \right\}^2 + P_g'(\theta) \frac{d^2\theta}{d\tau^2} \\ = 2\alpha_g [\alpha_g \{Q_g(\theta) - P_g(\theta)\} \left\{ \frac{d\theta}{d\tau} \right\}^2 + \{P_g(\theta)Q_g(\theta)\}^{1/2} \frac{d^2\theta}{d\tau^2}]$$

and

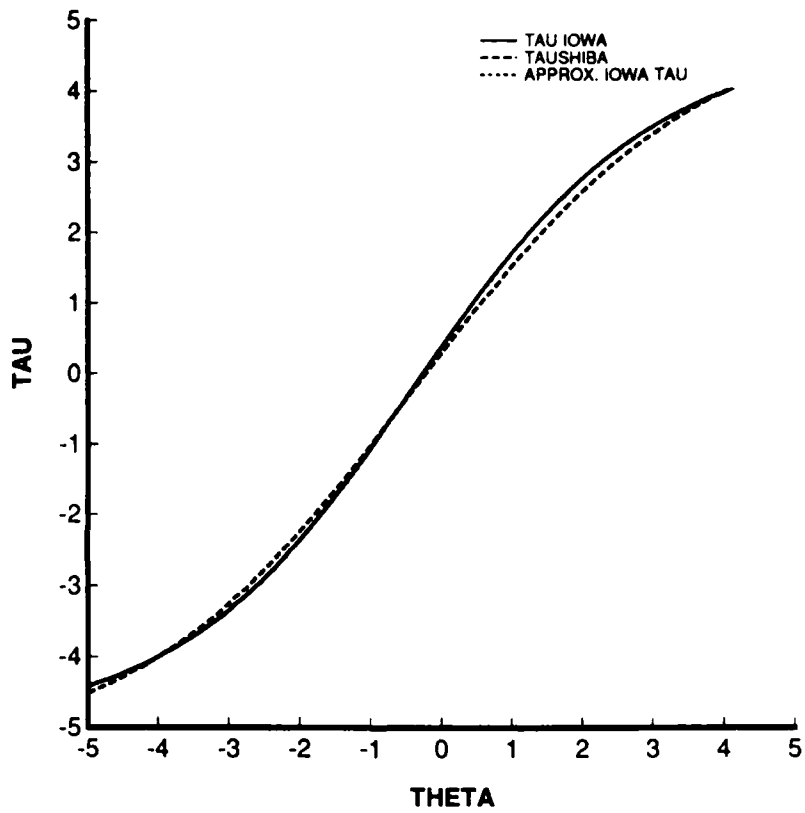
$$(7.22) \quad I_g^*(\tau) = I_g(\theta) \left\{ \frac{d\theta}{d\tau} \right\}^2 = 4\alpha_g^2 \left\{ \frac{d\theta}{d\tau} \right\}^2 ,$$



50 3 50 1 50 4 50 8 00

FIGURE 7-5

MLE Bias Function of the Iowa Level 11 Vocabulary Subtest as a Function of the Transformed Latent Variable  $\tau$  (Dashed Line) in Comparison to the Original MLE Bias Function (Solid Line) of  $\theta$ . Transformation of  $\theta$  to  $\tau$  Is Made by a Polynomial Approximation.



1 000 0 50 1 50 6 50 8 00

FIGURE 7-6

Transformation of  $\theta$  to  $\tau$  Based Upon The Iowa Level 11 Vocabulary Subtest (Solid Line), Upon Shiba's Word/Phrase Comprehension Test J1 (Dashed Line), and Upon the Iowa Level 11 Vocabulary Subtest Using the Polynomial Approximation.

where  $P_g^*(\tau)$  and  $I_g^*(\tau)$  indicate the item characteristic function and the item information function, respectively, after the scale transformation, and  $P_g^{*'}(\tau)$  and  $P_g^{*''}(\tau)$  denote the first and second derivatives of  $P_g^*(\tau)$  with respect to  $\tau$ . Since we can write for a set of  $n$  equivalent items

$$(7.23) \quad B^*(\tau) = (-2n)^{-1} P_g^*(\tau) Q_g^*(\tau) P_g^{*''}(\tau) \{P_g^{*''}(\tau)\}^{-3} ,$$

where  $B^*(\tau)$  is the bias function after the scale transformation, we obtain

$$(7.24) \quad B^*(\tau) = B(\theta) \left\{ \frac{d\theta}{d\tau} \right\}^{-1} - (8n\alpha_g^2)^{-1} \left\{ \frac{d\theta}{d\tau} \right\}^{-3} \frac{d^2\theta}{d\tau^2} .$$

For the purpose of illustration, let us consider the scale transformation which changes the constant information model to the logistic model. Thus we have

$$(7.25) \quad P_g^*(\tau) = \{1 + e^{-D\alpha_g(\tau - b_g)}\}^{-1} .$$

The functional relationship between  $\theta$  and  $\tau$  is given by

$$(7.26) \quad \tau = (D\alpha_g)^{-1} \log[\tan^2\{\alpha_g(\theta - \beta_g) + (\pi/4)\}] + b_g ,$$

or

$$(7.27) \quad \theta = \alpha_g^{-1} [\tan^{-1}\{e^{(1/2)D\alpha_g(\tau - b_g)}\} - (\pi/4)] + \beta_g .$$

The first and second derivatives of  $\theta$  with respect to  $\tau$  are given by

$$(7.28) \quad \frac{d\theta}{d\tau} = D\alpha_g(2\alpha_g)^{-1} \{P_g(\theta)Q_g(\theta)\}^{1/2}$$

and

$$(7.29) \quad \frac{d^2\theta}{d\tau^2} = D^2\alpha_g^2(4\alpha_g)^{-1} \{P_g(\theta)Q_g(\theta)\}^{1/2} \{Q_g(\theta) - P_g(\theta)\} ,$$

respectively. Thus we can write from (7.19), (7.20), (7.21), (7.22), (7.28), and (7.29)

$$(7.30) \quad P_g^{*'}(\tau) = D\alpha_g P_g(\theta) Q_g(\theta) = D\alpha_g P_g^*(\tau) Q_g^*(\tau) ,$$

$$(7.31) \quad \begin{aligned} P_g^{*''}(\tau) &= D^2\alpha_g^2 P_g(\theta) Q_g(\theta) \{Q_g(\theta) - P_g(\theta)\} \\ &= D^2\alpha_g^2 P_g^*(\tau) Q_g^*(\tau) \{Q_g^*(\tau) - P_g^*(\tau)\} \end{aligned}$$

and

$$(7.32) \quad I_g^*(\tau) = D^2\alpha_g^2 P_g(\theta) Q_g(\theta) = D^2\alpha_g^2 P_g^*(\tau) Q_g^*(\tau) ,$$

where

$$(7.33) \quad Q_g^*(\tau) = 1 - P_g^*(\tau) .$$

We can easily see that these results are agreeable with those obtained directly from (7.25). For the bias function  $B^*(\tau)$ , we have from (7.24), (7.28), and (7.29)

$$(7.34) \quad \begin{aligned} B^*(\tau) &= (2nDa_g)^{-1} \{P_g(\theta) - Q_g(\theta)\} \{P_g(\theta)Q_g(\theta)\}^{-1} \\ &= (nDa_g)^{-1} \{P_g^*(\tau)Q_g^*(\tau)\}^{-1} \{P_g^*(\tau) - (1/2)\} . \end{aligned}$$

We can see that (7.34) is a special case of (1.4) when all the  $n$  items are equivalent, by replacing  $\theta$  by  $\tau$  and  $\Psi_g(\theta)$  by  $P_g^*(\tau)$ .

## 8 Adaptive Testing

Observations made in previous sections provide us with ideas how things go in adaptive testing. First of all, in order to reach the practical unbiasedness in estimating the individual subject's ability in adaptive testing, we need to make sure that a sufficient amount of test information has been reached for each individual subject, before terminating the presentation of new items. We can control it easily, if we use the amount of test information as the criterion for the termination of presenting new items, or the stopping rule. If the items follow the normal ogive or logistic model in the adaptive testing situation, for subjects of intermediate ability levels it is likely that on the initial stage the item difficulty parameters fluctuate both negatively and positively around the subject's true ability level, and consequently, the biases of negative and positive directions are cancelled out, since an item pool usually has plenty of items of intermediate difficulties. In such a case, we do not have to worry too much about the influence of initial items on the eventual bias of the ability estimate. When the maximum likelihood estimate has started being more or less stabilized, chances are slim that the additional item causes a substantial bias, provided that the program is written in such a way that an item of a large amount of information at the current estimated ability level will be presented next, and that the item pool has a sufficient number of items whose difficulty levels are around the subject's true ability level. There is a greater possibility that the examinee obtains a biased ability estimate if his ability level is close to either end of the configuration of difficulty parameters, since biases caused by the initially presented items are not likely to cancel themselves out, and, moreover, there may not be a sufficient number of items whose difficulty levels are close to his true ability level.

If the item pool consists of items following the three-parameter normal ogive or logistic model, the effect of random guessing on the amount of bias can be substantial, especially on the lower levels of ability. In such a case, it is imperative to include many easy items in the item pool.

In any case, the bias function can be a good indicator in evaluating the item pool, if we use it wisely and effectively. Those results that were described in previous sections will give us information and suggestions as to how to improve an existing item pool.

## 9 Discussion and Conclusions

The bias function of the maximum likelihood estimate has been proposed for the general discrete response level, which includes Lord's bias function in the three-parameter logistic model as a special case. The function has also been observed both on the dichotomous and graded response levels, with respect to various mathematical models. Effects of the item discrimination parameters, of the item difficulty parameters, and of the number of items have also been observed. Local changes in the amount of bias caused by the scale transformation have also been observed from various different angles, and it has also been discussed in the context of adaptive testing.

Since the local unbiasedness is important, the proposed function will find its usefulness directly in the estimation of the subject's latent trait. An even greater usefulness of the function can be seen in the context of more elaborated methodologies, as is exemplified in the nonparametric approach to the estimation of the operating characteristics of discrete responses.

## References

- [1] Lord, F. M. Unbiased estimators of ability parameters, of their variance, and of their parallel-forms reliability. *Psychometrika*, 48, 1983, 233-245.
- [2] Samejima, F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*, No. 17, 1969.
- [3] Samejima, F. A general model for free-response data. *Psychometrika Monograph*, No. 18, 1972.
- [4] Samejima, F. A method of estimating item characteristic functions using the maximum likelihood estimate of ability. *Psychometrika*, 42, 1977, 163-191.
- [5] Samejima, F. Constant Information Model: A New, Promising Item Characteristic Function. *ONR Research Report*, 79-1, 1979.
- [6] Samejima, F. and Livingston, P.S. Method of moments as the least squares solution for fitting a polynomial. *ONR Research Report*, 79-2, 1979.
- [7] Samejima, F. Final Report: Efficient methods of estimating the operating characteristics of item response categories and challenge to a new model for the multiple-choice item. *Final Report of N00014-77-C-0960*, Office of Naval Research, 1981.
- [8] Samejima, F. Plausibility Functions of Iowa Vocabulary Test Items Estimated by the Simple Sum Procedure of the Conditional P.D.F. Approach. *ONR Research Report*, 84-1, 1984.
- [9] Samejima, F. Comparison of the Estimated Item Parameters of Shiba's Word/Phrase Comprehension Tests Obtained by Logist 5 and Those by the Tetrachoric Method. *ONR Research Report*, 84-2, 1984.
- [10] Shiba, S. Construction of a Scale for Acquisition of Word Meanings. *Bulletin of Faculty of Education, University of Tokyo*, 17, 1978. (in Japanese)

ONRR8701.87  
June 16, 1987



1987/05/13

University of Tennessee/Samejias

Dr. Terry Ackerman  
American College Testing Programs  
P.O. Box 168  
Iowa City, IA 52243

Dr. Robert Ahlers  
Code N711  
Human Factors Laboratory  
Naval Training Systems Center  
Orlando, FL 32813

Dr. James Aigins  
University of Florida  
Gainesville, FL 32605

Dr. Erling B. Andersen  
Department of Statistics  
Stuejstraede 6  
1455 Copenhagen  
DENMARK

Dr. Eva L. Baker  
UCLA Center for the Study  
of Evaluation  
145 Moore Hall  
University of California  
Los Angeles, CA 90024

Dr. Isaac Bejar  
Educational Testing Service  
Princeton, NJ 08450

Dr. Menucha Birnbaum  
School of Education  
Tel Aviv University  
Tel Aviv, Ramat Aviv 69978  
ISRAEL

Dr. Arthur S. Blaisies  
Code N711  
Naval Training Systems Center  
Orlando, FL 32813

Dr. Bruce Blaxom  
Defense Manpower Data Center  
550 Camino El Estero,  
Suite 200  
Monterey, CA 93943-3231

Dr. R. Darrell Bock  
University of Chicago  
NORC  
5030 South Ellis  
Chicago, IL 60637

Cdt. Arnold Bohrer  
Sectie Psychologisch Onderzoek  
Rekruterings-En Selectiecentrum  
Kwartier Koningen Astrid  
Bruijnstreat  
1120 Brussels, BELGIUM

Dr. Robert Breaux  
Code N-056R  
Naval Training Systems Center  
Orlando, FL 32813

Dr. Robert Brennan  
American College Testing  
Programs  
P. O. Box 168  
Iowa City, IA 52243

Dr. Lyle D. Broeseel  
Code 1111SP  
800 North Quincy Street  
Arlington, VA 22217

Mr. James M. Carey  
Commandant (G-PTE)  
U.S. Coast Guard  
2100 Second Street, S.W.  
Washington, DC 20593

Dr. James Carlson  
American College Testing  
Program  
P.O. Box 168  
Iowa City, IA 52243

Dr. John B. Carroll  
409 Elliott Rd.  
Chapel Hill, NC 27514

Dr. Robert Carroll  
OP 01B7  
Washington, DC 20370

Mr. Raymond E. Christal  
AFMRL/MOE  
Brooks AFB, TX 78235

Dr. Norman Cliff  
Department of Psychology  
Univ. of So. California  
University Park  
Los Angeles, CA 90007

Director,  
Manpower Support and  
Readiness Program  
Center for Naval Analysis  
2000 North Beauregard Street  
Alexandria, VA 22311

Dr. Stanley Collyer  
Office of Naval Technology  
Code 222  
800 M. Quinoy Street  
Arlington, VA 22217-5000

Dr. Hans Crombag  
University of Leyden  
Education Research Center  
Boerhaavelaan 2  
2334 EN Leyden  
The NETHERLANDS

Mr. Timothy Davey  
University of Illinois  
Educational Psychology  
Urbana, IL 61801

Dr. Doug Davis  
Chief of Naval Education  
and Training  
Naval Air Station  
Pensacola, FL 32508

Dr. Dattprasad Divgi  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Dr. Hei-Ki Dong  
Ball Foundation  
800 Roosevelt Road  
Building C, Suite 206  
Glen Ellyn, IL 60137

Dr. Fritz Dreagow  
University of Illinois  
Department of Psychology  
603 E. Daniel St.  
Champaign, IL 61820

Defense Technical  
Information Center  
Cameron Station, Bldg 5  
Alexandria, VA 22314  
Attn: TC  
(12 Copies)

Dr. Stephen Dunbar  
Lindquist Center  
for Measurement  
University of Iowa  
Iowa City, IA 52242

Dr. James A. Earles  
Air Force Human Resources Lab  
Brooks AFB, TX 78235

Dr. Kent Esten  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Dr. John M. Eddins  
University of Illinois  
252 Engineering Research  
Laboratory  
103 South Mathews Street  
Urbana, IL 61801

Dr. Susan Embretson  
University of Kansas  
Psychology Department  
425 Fraser  
Lawrence, KS 66045

Dr. Benjamin A. Fairbank  
Performance Metrics, Inc.  
5825 Callaghan  
Suite 225  
San Antonio, TX 78228

Dr. Pat Federico  
Code 511  
NPRDC  
San Diego, CA 92152-6800

1987/05/13

University of Tennessee/Samejias

Dr. Leonard Feldt  
Lindquist Center  
for Measurement  
University of Iowa  
Iowa City, IA 52242

Dr. Richard L. Ferguson  
American College Testing  
Program  
P.O. Box 168  
Iowa City, IA 52240

Dr. Gerhard Fischer  
Liebigasse 5/3  
A 1010 Vienna  
AUSTRIA

Dr. Myron Fischl  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Prof. Donald Fitzgerald  
University of New England  
Department of Psychology  
Armidale, New South Wales 2351  
AUSTRALIA

Mr. Paul Foley  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Alfred R. Fogly  
AFGR/ML  
Bolling AFB, DC 20332

Dr. Robert D. Gibbons  
Illinois State Psychiatric Inst.  
Rm 523M  
1601 W. Taylor Street  
Chicago, IL 60612

Dr. Janice Gifford  
University of Massachusetts  
School of Education  
Amherst, MA 01003

Dr. Robert Glaser  
Learning Research  
& Development Center  
University of Pittsburgh  
3939 O'Hara Street  
Pittsburgh, PA 15260

Dr. Bert Green  
Johns Hopkins University  
Department of Psychology  
Charles & 34th Street  
Baltimore, MD 21218

Dipl. Ped. Michael M. Habon  
Universität Düsseldorf  
Erziehungswissenschaftliches  
Universitätsstr. 1  
D-4000 Düsseldorf 1  
WEST GERMANY

Dr. Ronald K. Hambleton  
Prof. of Education & Psychology  
University of Massachusetts  
at Amherst  
Hille House  
Amherst, MA 01003

Dr. Delwyn Harnisch  
University of Illinois  
51 Serty Drive  
Champaign, IL 61820

Ms. Rebecca Hetter  
Navy Personnel R&D Center  
Cede 62  
San Diego, CA 92152-6800

Dr. Paul M. Holland  
Educational Testing Service  
Rosedale Road  
Princeton, NJ 08541

Prof. Lutz F. Hornke  
Institut für Psychologie  
RMTH Aachen  
Jaegerstrasse 17/19  
D-5100 Aachen  
WEST GERMANY

Dr. Paul Horst  
677 G Street, #184  
Chula Vista, CA 90010

Mr. Dick Hoshaw  
OP-135  
Arlington Annex  
Room 2834  
Washington, DC 20350

Dr. Lloyd Humphreys  
University of Illinois  
Department of Psychology  
Charles & 34th Street  
Champaign, IL 61820

Dr. Steven Hunke  
Department of Education  
University of Alberta  
Edmonton, Alberta  
CANADA

Dr. Huynh Huynh  
College of Education  
Univ. of South Carolina  
Columbia, SC 29208

Dr. Robert Jennerone  
Department of Psychology  
University of South Carolina  
Columbia, SC 29208

Dr. Dennis E. Jennings  
Department of Statistics  
University of Illinois  
1409 West Green Street  
Urbana, IL 61801

Dr. Douglas H. Jones  
Thatcher Jones Associates  
P.O. Box 6640  
10 Trafalgar Court  
Lawrenceville, NJ 08648

Dr. Milton S. Katz  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Prof. John A. Keats  
Department of Psychology  
University of Newcastle  
N.S.W. 2308  
AUSTRALIA

Dr. G. Gage Kingsbury  
Portland Public Schools  
Research and Evaluation Department  
501 North Dixon Street  
P. O. Box 3107  
Portland, OR 97209-3107

Dr. William Koch  
University of Texas-Austin  
Measurement and Evaluation  
Center  
Austin, TX 78703

Dr. James Krantz  
Computer-based Education  
Research Laboratory  
University of Illinois  
Urbana, IL 61801

Dr. Leonard Kroeker  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Daryll Lang  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Thomas Leonard  
University of Wisconsin  
Department of Statistics  
1210 West Dayton Street  
Madison, WI 53705

Dr. Michael Levine  
Educational Psychology  
210 Education Bldg.  
University of Illinois  
Champaign, IL 61801

Dr. Charles Lewis  
Educational Testing Service  
Princeton, NJ 08541

Dr. Robert Linn  
College of Education  
University of Illinois  
Urbana, IL 61801

Dr. Robert Lochman  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Dr. Frederic M. Lord  
Educational Testing Service  
Princeton, NJ 08541

Dr. James Lueden  
Department of Psychology  
University of Western Australia  
Nedlands M.A. 6009  
AUSTRALIA

Dr. Milton Maier  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Dr. William L. Meloy  
Chief of Naval Education  
and Training  
Naval Air Station  
Pensacola, FL 32508

Dr. Gary Marco  
Step 31-E  
Educational Testing Service  
Princeton, NJ 08541

Dr. Clessen Martin  
Army Research Institute  
5001 Eisenhower Blvd.  
Alexandria, VA 22333

Dr. James McBride  
Psychological Corporation  
c/o Harcourt, Brace,  
Jovanovich Inc.  
1250 West 6th Street  
San Diego, CA 92101

Dr. Clarence McCormick  
HQ, MEPCOM  
MEPT-P  
2500 Green Bay Road  
North Chicago, IL 60064

Dr. Robert McKinley  
Educational Testing Service  
Princeton, NJ 08541

Dr. James McMichael  
Technical Director  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Barbara Means  
Human Resources  
Research Organization  
1100 South Washington  
Alexandria, VA 22314

Dr. Robert Miallevy  
Educational Testing Service  
Princeton, NJ 08541

Dr. William Montague  
NPRDC Code 13  
San Diego, CA 92152-6800

Ms. Kathleen Moreno  
Navy Personnel R&D Center  
Code 62  
San Diego, CA 92152-6800

Headquarters, Marine Corps  
Code MP1-20  
Washington, DC 20380

Dr. M. Alan Nicewander  
University of Oklahoma  
Department of Psychology  
Oklahoma City, OK 73069

Deputy Technical Director  
NPRDC Code 01A  
San Diego, CA 92152-6800

Director, Training Laboratory,  
NPRDC (Code 05)  
San Diego, CA 92152-6800

Director, Manpower and Personnel  
Laboratory,  
NPRDC (Code 06)  
San Diego, CA 92152-6800

Director, Human Factors  
& Organizational Systems Lab,  
NPRDC (Code 07)  
San Diego, CA 92152-6800

Fleet Support Office,  
NPRDC (Code 301)  
San Diego, CA 92152-6800

Library, NPRDC  
Code P201L  
San Diego, CA 92152-6800

Commanding Officer,  
Naval Research Laboratory  
Code 2627  
Washington, DC 20390

Dr. Harold F. O'Neil, Jr.  
School of Education - MPH 801  
Department of Educational  
Psychology & Technology  
University of Southern California  
Los Angeles, CA 90089-0031

Dr. James Olson  
WICAT, Inc.  
1875 South State Street  
Drem, UT 84057

Office of Naval Research,  
Code 1142PT  
800 N. Quincy Street  
Arlington, VA 22217-5000  
(6 Copies)

Office of Naval Research,  
Code 125  
800 N. Quincy Street  
Arlington, VA 22217-5000

Assistant for MPT Research,  
Development and Studies  
OP 01B7  
Washington, DC 20370

Dr. Judith Orasenu  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Dr. Jesse Orlensky  
Institute for Defense Analyses  
1801 N. Beauregard St.  
Alexandria, VA 22311

Dr. Randolph Park  
Army Research Institute  
5001 Eisenhower Blvd.  
Alexandria, VA 22333

Wayne M. Patience  
American Council on Education  
GED Testing Service, Suite 20  
One Dupont Circle, NW  
Washington, DC 20036

Dr. James Paulson  
Department of Psychology  
Portland State University  
P.O. Box 751  
Portland, OR 97207

Administrative Sciences Department,  
Naval Postgraduate School  
Monterey, CA 93940

Department of Operations Research,  
Naval Postgraduate School  
Monterey, CA 93940

Dr. Mark D. Reckase  
ACT  
P. O. Box 168  
Iowa City, IA 52243

Dr. Malcolm Ree  
AFHRL/MP  
Brooks AFB, TX 78235

Dr. Barry Riegelhaupt  
HuePDR  
1100 South Washington Street  
Alexandria, VA 22314

Dr. Carl Ross  
CNET-PDCC  
Building 90  
Great Lakes NTC, IL 60088

Dr. J. Ryan  
Department of Education  
University of South Carolina  
Columbia, SC 29208

Dr. Fumiko Samejima  
Department of Psychology  
University of Tennessee  
Knoxville, TN 37916

Mr. Drew Sands  
NPRDC Code 62  
San Diego, CA 92152-6800

Lowell Schoer  
Psychological & Quantitative  
Foundations  
College of Education  
University of Iowa  
Iowa City, IA 52242

## University of Tennessee/Samejima

## University of Tennessee/Samejima

Dr. Mery Schratz  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Dan Segall  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. M. Steve Sallman  
OASD(MR&A)  
28269 The Pentagon  
Washington, DC 20301

Dr. Kazuo Shigemasu  
7-9-24 Kugunab-Kaigan  
Fujisawa 251  
JAPAN

Dr. William Sims  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0768

Dr. H. Wallace Simeko  
Manpower Research  
and Advisory Services  
Smithsonian Institution  
801 North Pitt Street  
Alexandria, VA 22314

Dr. Richard E. Snow  
Department of Psychology  
Stanford University  
Stanford, CA 94306

Dr. Richard Sorensen  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Paul Spackman  
University of Missouri  
Department of Statistics  
Columbia, MO 65201

Dr. Judy Spray  
ACT  
P.O. Box 168  
Iowa City, IA 52243

Dr. Martha Stocking  
Educational Testing Service  
Princeton, NJ 08541

Dr. Peter Stolf  
Center for Naval Analysis  
200 North Beauregard Street  
Alexandria, VA 22311

Dr. William Stout  
University of Illinois  
Department of Mathematics  
Urbana, IL 61801

Maj. Bill Strickland  
AF/MPXDA  
4E168 Pentagon  
Washington, DC 20330

Dr. Hariharan Swaminathan  
Laboratory of Psychometric and  
Evaluation Research  
School of Education  
University of Massachusetts  
Amherst, MA 01003

Mr. Brad Symson  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. John Tangney  
AFOSR/ML  
Bolling AFB, DC 20332

Dr. Kikumai Tatsuoka  
CERL  
252 Engineering Research  
Laboratory  
Urbana, IL 61801

Dr. Maurice Tatsuoka  
220 Education Bldg  
1310 S. Sixth St.  
Champaign, IL 61820

Dr. David Thissen  
Department of Psychology  
University of Kansas  
Lawrence, KS 66044

Mr. Gary Thomason  
University of Illinois  
Educational Psychology  
Champaign, IL 61820

Dr. Robert Tautkous  
University of Missouri  
Department of Statistics  
222 Math. Sciences Bldg.  
Columbia, MO 65211

Dr. Ledyard Tucker  
University of Illinois  
Department of Psychology  
603 E. Daniel Street  
Champaign, IL 61820

Dr. Vern M. Urry  
Personnel R&D Center  
Office of Personnel Management  
1900 E. Street, NW  
Washington, DC 20415

Dr. David Valle  
Assessment Systems Corp.  
2233 University Avenue  
Suite 310  
St. Paul, MN 55114

Dr. Frank Vicino  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Howard Weiner  
Division of Psychological Studies  
Educational Testing Service  
Princeton, NJ 08541

Dr. Ming-Mei Wang  
Lindquist Center  
for Measurement  
University of Iowa  
Iowa City, IA 52242

Dr. Thomas A. Warm  
Coast Guard Institute  
P. O. Substation 1B  
Oklahoma City, OK 73169

Dr. Brian Waters  
Program Manager  
Manpower Analysis Program  
HumRRD  
1100 S. Washington St.  
Alexandria, VA 22314

Dr. David J. Weis  
N660 Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455

Dr. Ronald A. Weitzman  
NPS, Code 54Mz  
Monterey, CA 92152-6800

Major John Weisb  
AFHRL/MDAM  
Brooks AFB, TX 78223

Dr. Douglas Metzler  
Code 12  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Rand R. Wilcox  
University of Southern  
California  
Department of Psychology  
Los Angeles, CA 90007

German Military Representative  
ATTN: Wolfgang Wildgrube  
Streitkräfteeinsatz  
D-5300 Bonn 2  
4000 Brandywine Street, NW  
Washington, DC 20016

Dr. Bruce Williams  
Department of Educational  
Psychology  
University of Illinois  
Urbana, IL 61801

Dr. Hilda Ming  
Psychological Corporation  
c/o Harcourt, Brace,  
Jovanovich Inc.  
1250 West 6th Street  
San Diego, CA 92101

Dr. Martin F. Wiskoff  
Navy Personnel R & D Center  
San Diego, CA 92152-6800

Mr. John H. Wolfe  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

1987/05/13

University of Tennessee/Seemjms

Dr. George Mong  
Biostatistics Laboratory  
Memorial Sloan-Kettering  
Cancer Center  
1275 York Avenue  
New York, NY 10021

Dr. Wallace Mulleck, III  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Kentaro Yamamoto  
Computer-based Education  
Research Laboratory  
University of Illinois  
Urbana, IL 61801

Dr. Mandy Yen  
CTB/McGraw Hill  
Del Monte Research Park  
Monterey, CA 93940

Dr. Joseph L. Young  
Memory & Cognitive  
Processes  
National Science Foundation  
Washington, DC 20550

NON GOVT

- 1 Dr. Robert Guton  
Department of Psychology  
Bowling Green State University  
Bowling Green, OH 43403
- 1 Dr. William W. Turnbuil  
Educational Testing Service  
Princeton, NJ 08542
- 1 Dr. P. Mengal  
Faculte de Psychologie  
et des Sciences de l'Education  
Universite de Geneve  
3 fl. de l'Universite  
1201 Geneva SWITZERLAND
- 1 Dr. Wim J. van der Linden  
Vakgroep Onderwijskunde  
Postbus 217  
7500 EA Endschede  
THE NETHERLANDS
- 1 Dr. Lutz Hornke  
University Duesseeldorf  
Erz. Wiss.  
D-4000 Dueseldorf  
WEST GERMANY
- 1 Dr. Wolfgang Buchtala  
8346 Simbach Inn  
Postfach 1306  
Industriestrasse 1  
WEST GERMANY
- 1 Dr. Albert Beaton  
Educational Testing Service  
Princeton, NJ 08542
- 1 Dr. Sukeyori Shiba  
Faculty of Education  
University of Tokyo  
Hongo, Bunkyo-ku  
Tokyo, JAPAN 113

NON GOVT

- 1 Dr. Takahiro Sato  
Nippon Electric Co., Ltd.  
C & C Systems Research Laboratories  
4-1-1 Miyazaki  
Miyamaeku, Kawasaki  
Kanagawaken 213, JAPAN
- 1 Dr. J. Uhlener  
Perceptronics, Inc.  
6271 Variel Ave.  
Woodland Hills, CA 91634
- 1 Mr. Susumu Fujimori  
1-21-36 Hiyoshi  
Kohokuku, Yokohama 223  
JAPAN
- 1 Mr. Kenji Goto  
Ina Nyuhaim #208  
7-2-12 Honcho, Tanashi-shi  
Tokyo 188, JAPAN
- 1 Dr. G. Gage Kingsbury  
Portland Public Schools  
Evaluation Department  
501 North Dixon Street  
Portland, Oregon 97227
- 1 Mr. Tadaashi Shibayama  
Ono-kohpo #201  
901 Hirasasaki  
Nagareyama-shi  
Chiba-ken, 270-01  
JAPAN
- 1 Dr. Mark Wilson  
School of Education  
University of California  
Berkeley, CA 94720
- 1 Dr. Chang-I Ronnie Chen  
Graduate School of Psychology  
National Chengchi University  
Taipei, Taiwan  
R.O.C.

NON GOVT

- 1 Dr. Ivo W. Molenaar  
F.S.W.-R.U.G.  
Oude Boteringestraat 23  
9712 GC Groningen  
THE NETHERLANDS
- NAVY
- 1 Mr. Thomas Bryant  
Office of Naval Research  
206 O'Keefe Building  
Atlanta, GA 30332
- ARMY
- 1 Dr. Randall M. Chambers  
U.S. Army Research Institute for  
the Behavioral & Social Sciences  
Fort Sill Field Unit  
P.O. Box 3066  
Fort Sill, OK 73503

END

9-87

DTIC