

AD-A063 918

TENNESSEE UNIV KNOXVILLE DEPT OF PSYCHOLOGY

F/G 12/1

ESTIMATION OF THE OPERATING CHARACTERISTICS OF ITEM RESPONSE CA--ETC(U)

DEC 78 F SAMEJIMA

N00014-77-C-0360

UNCLASSIFIED

RR-78-5

NL

1 OF 2  
AD  
A063918



2

**LEVEL**

III

12

RESEARCH REPORT 78-5

AD A063918

ESTIMATION OF THE OPERATING CHARACTERISTICS OF ITEM  
RESPONSE CATEGORIES VI: PROPORTIONED SUM PROCEDURE  
IN THE CONDITIONAL P.D.F. APPROACH

FUMIKO SAMEJIMA

DEPARTMENT OF PSYCHOLOGY  
UNIVERSITY OF TENNESSEE  
KNOXVILLE, TENN. 37916

DECEMBER, 1978

DDC FILE COPY

DDC  
JAN 30 1979  
A

Prepared under the contract number N00014-77-C-360,  
NR 150-402 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.

79 01 29 098

Unclassified

14 RR-78-5 *moves*

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

<p>9. REPORT NUMBER Research Report, 78-5</p>	<p>2. GOVT ACCESSION NO.</p>	<p>3. RECIPIENT'S CATALOG NUMBER</p>
<p>6. TITLE (and Subtitle) Estimation of the Operating Characteristics of Item Response Categories, VI. Proportioned Sum Procedure in the Conditional P.D.F. Approach</p>		<p>5. TYPE OF REPORT &amp; PERIOD COVERED Technical Report</p>
<p>7. AUTHOR(s) 10. Dr. Fumiko Samejima</p>	<p>15.</p>	<p>8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0360</p>
<p>9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology University of Tennessee Knoxville, TN 37916</p>	<p>16. 17. PE: 61153N; PROJ: RR 042 04 TA: RR 042 04 01 WU: RR 042 04 01</p>	<p>12. REPORT DATE 15 December 78</p>
<p>11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, VA 22217</p>	<p>18. NUMBER OF PAGES 119</p> <p>15. SECURITY CLASS. (of this report) Unclassified</p> <p>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE</p>	<p>18. MONITORING AGENCY NAME &amp; ADDRESS (if different from Controlling Office) 12. 119p.</p>
<p>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.</p>		
<p>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</p>		
<p>18. SUPPLEMENTARY NOTES</p>		
<p>19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Operating Characteristic Estimation Tailored Testing Latent Trait Theory</p>		
<p>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  (Please see reverse side)</p> <p style="text-align: right;">404 225 <i>Gu</i></p>		

Following Simple Sum Procedure and Weighted Sum Procedure, another method, Proportioned Sum Procedure, is introduced in the context of the Conditional P.D.F. Approach. The new method is somewhat different from the previous two, however, in the sense that the set of conditional density functions is not exclusively recategorized into the item score groups, but they are proportioned into each item score category. The same hypothetical data, i.e., the maximum likelihood estimates of the five hundred hypothetical subjects and their responses to the ten binary items, each of which follows the normal ogive model, are used to try the method. The criterion item characteristic function for each binary item is obtained and compared with those obtained by the other two procedures. The Pearson System Method and the Two-Parameter Beta Method are used for both Degree 3 and 4 Cases, and the results are compared with the previous ones. The mean square errors are adopted in evaluating both the resultant estimated item characteristic functions and probability density functions of ability. The two item parameters in the normal ogive model are also estimated for each item, and the results are compared with the previous ones.

ACCESSION BY	
NTS	White Section <input checked="" type="checkbox"/>
DD	Ref Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION.....	
BY.....	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. nos./or SPECIAL
A	

ESTIMATION OF THE OPERATING CHARACTERISTICS OF ITEM RESPONSE  
CATEGORIES VI: PROPORTIONED SUM PROCEDURE IN THE CONDITIONAL  
P.D.F. APPROACH

ABSTRACT

Following Simple Sum Procedure and Weighted Sum Procedure, another method, Proportioned Sum Procedure, is introduced in the context of the Conditional P.D.F. Approach. The new method is somewhat different from the previous two, however, in the sense that the set of conditional density functions is not exclusively recategorized into the item score groups, but they are proportioned into each item score category. The same hypothetical data, i.e., the maximum likelihood estimates of the five hundred hypothetical subjects and their responses to the ten binary items, each of which follows the normal ogive model, are used to try the method. The criterion item characteristic function for each binary item is obtained and compared with those obtained by the other two procedures. The Pearson System Method and the Two-Parameter Beta Method are used for both Degree 3 and 4 Cases, and the results are compared with the previous ones. The mean square errors are adopted in evaluating both the resultant estimated item characteristic functions and probability density functions of ability. The two item parameters in the normal ogive model are also estimated for each item, and the results are compared with the previous ones.

---

The research was conducted at the principal investigator's laboratory, 409 Austin Peay Hall, Department of Psychology, University of Tennessee, Knoxville, Tennessee. Those who worked for her as assistants at various times include Paul S. Chagas, Robert L. Trestman and Philip S. Livingston.

## TABLE OF CONTENTS

	Page
I Introduction	1
II Conditional P.D.F. Approach	3
III Proportioned Sum Procedure	5
IV Results I: Proportioned Weights	7
V Results II: Estimated Density Functions of Ability	19
VI Results III: Estimated Item Characteristic Functions	56
VII Discussion and Conclusion	99
References	101
<u>Appendices</u>	
Appendix I	103
Appendix II	107

## I Introduction

Following previous studies (Samejima, 1977b, 1977d, 1978a, 1978b, 1978c, 1978d), the present paper proposes another method for estimating the operating characteristics of the graded items (Samejima, 1969, 1972), in which:

- (1) no mathematical forms are assumed for the operating characteristics;
- (2) a relatively small number of examinees are used in the whole process of estimation;
- (3) the maximum likelihood estimates of the examinees are available because of a previous testing; and
- (4) the test which was used for (3) has a constant test information function for the range of ability of our interest, and the amount of information is substantially large.

The process of the maximum likelihood estimation using the response pattern of a specified examinee has been described (Samejima, 1969). The asymptotic property of the maximum likelihood estimate (Samejima, 1975), i.e., the normality of its conditional distribution, given ability, with the given ability itself and the inverse of the test information function as its two parameters, is fully used. Since the constancy of the test information has many theoretical advantages (Samejima, 1977a), the estimation of the conditional moments of ability, given its maximum likelihood estimate, is relatively easy\*, provided that the density function of the maximum likelihood estimate is given. By virtue of the concept of weakly parallel tests (Samejima, 1977c), the results are directly comparable with those of similar studies using

---

\* Refer to Appendix II.

different sets of test items but following the same principles, (1) through (4).

The actual data used in the present study are simulated data of five hundred hypothetical examinees, and our target is to estimate the item characteristic function, i.e., the operating characteristic of the item score 1 when the item is binary, of each of the ten binary items. A brief description of these data is given as Appendix I. The ten binary items follow the normal ogive model, and the two parameters, discrimination and difficulty indices, are given in Tables 6-3, 6-4, etc., for each item.

Throughout the study, the method of moments (Elderton and Johnson, 1969, Johnson and Kotz, 1970) is frequently used. Among others, the density function of the maximum likelihood estimate is approximated by a polynomial of degree 3 or 4, applying the method of moments for the set of 500 maximum likelihood estimates.



## II Conditional P.D.F. Approach

Conditional P.D.F. has been presented and developed (Samejima, 1978a, 1978b, 1978d). Let  $f(\theta)$  be the density function of ability, or latent trait,  $\theta$ ,  $\psi(\hat{\theta}|\theta)$  be the conditional density of  $\hat{\theta}$ , given  $\theta$ , which is approximated by  $n(\theta, \sigma^2)$  where  $\sigma^2$  is the inverse of the test information function,  $I(\theta) = 21.63$ , and  $g(\hat{\theta})$  is the density function of the maximum likelihood estimate,  $\hat{\theta}$ . We can write for the conditional density function of  $\hat{\theta}$ , given  $\theta$ ,

$$(2.1) \quad \begin{aligned} \phi(\theta|\hat{\theta}) &= \psi(\hat{\theta}|\theta) f(\theta) \left[ \int_{-\infty}^{\infty} \psi(\hat{\theta}|\theta) f(\theta) d\theta \right]^{-1} \\ &= \psi(\hat{\theta}|\theta) f(\theta) [g(\hat{\theta})]^{-1} . \end{aligned}$$

Let  $P_{x_g}(\theta)$  be the operating characteristic of the graded item score category  $x_g$ . The estimate of the operating characteristic,  $\hat{P}_{x_g}(\theta)$ , in the Conditional P.D.F. Approach can be written in the general form

$$(2.2) \quad \hat{P}_{x_g}(\theta) = \sum_{s \in G_{x_g}} w(\hat{\theta}_s) \hat{\phi}(\theta|\hat{\theta}_s) \left[ \sum_{s=1}^N w(\hat{\theta}_s) \hat{\phi}(\theta|\hat{\theta}_s) \right]^{-1} ,$$

where  $G_{x_g}$  indicates the group of examinees having the item score  $x_g$ ,  $N (=500)$  is the number of examinees, and  $w(\hat{\theta}_s)$  is the weight assigned to the maximum likelihood estimate of the examinee  $s$ . When this weight is constant for all the examinees, (2.2) is called Simple Sum Procedure, to distinguish itself from the Weighted Sum Procedure (Samejima, 1978d).

To use (2.2), it is necessary to obtain the estimated conditional density of  $\theta$ , given  $\hat{\theta}$ , and so far three methods, i.e., Pearson System Method, Normal Approach Method and Two-Parameter Beta Method, have been used. In the first method,  $\phi(\theta|\hat{\theta}_g)$  is approximated by

one of the Pearson System density functions, depending upon the value of Pearson's criterion  $\kappa$ , which is obtained from the first four conditional moments of  $\theta$ , given  $\hat{\theta}$ . In the second method,  $\phi(\theta|\hat{\theta}_g)$  is approximated by a normal density function, using the first two conditional moments of  $\theta$ , given  $\hat{\theta}$ , as the parameters. In the third method, we use a Beta density function for  $\hat{\phi}(\theta|\hat{\theta}_g)$ , with a priori set two parameters, which are the two endpoints of the interval for which the Beta density function assumes positive values, and the other two parameters estimated from the first two conditional moments of  $\theta$ , given  $\hat{\theta}$ . (For further detail, cf. Samejima, 1977d, 1978a, 1978b, 1978d).

In the present study, the concept of Conditional P.D.F. Method is expanded and generalized to mean that the estimated conditional density  $\hat{\phi}(\theta|\hat{\theta})$  is fully utilized, without using any estimated joint density of  $\theta$  and  $\hat{\theta}$ . In the following section, therefore, a new procedure is proposed which is not based on (2.2).

III Proportioned Sum Procedure

Let  $p(s \in G_{x_g})$  be the probability with which the examinee  $s$  belongs to the group whose item score for item  $g$  is  $x_g$ , and  $\hat{p}(s \in G_{x_g})$  be its estimate. In the present method, the estimated operating characteristic,  $\hat{P}_{x_g}(\theta)$ , is defined by

$$(3.1) \quad \hat{P}_{x_g}(\theta) = \frac{\sum_{s=1}^N \hat{p}(s \in G_{x_g}) \hat{\phi}(\theta | \hat{\theta}_s)}{\sum_{s=1}^N \hat{\phi}(\theta | \hat{\theta}_s)}^{-1},$$

where

$$(3.2) \quad \sum_{x_g=0}^m p(s \in G_{x_g}) = \sum_{x_g=0}^m \hat{p}(s \in G_{x_g}) = 1.$$

The main difference between (3.1) and (2.2) is that in the numerator of (3.1) the estimated conditional densities for all the  $\hat{\theta}_s$ 's participate, while that of (2.2) includes  $\hat{\phi}(\theta | \hat{\theta}_g)$  for the examinees who belong to the group  $G_{x_g}$  only. The procedure represented by (3.1) will be called Proportioned Sum Procedure of the Conditional P.D.F. Approach.

Since our data used in the present study are for estimating the operating characteristics of binary items, it is sufficient to estimate the item characteristic function, i.e., the operating characteristic of  $x_g = 1$ , of each item. Let  $P_g(\theta)$  be the item characteristic function of item  $g$ . We can rewrite (3.1) such that

$$(3.3) \quad \hat{P}_g(\theta) = \frac{\sum_{s=1}^N \hat{p}(s \in G) \hat{\phi}(\theta | \hat{\theta}_s)}{\sum_{s=1}^N \hat{\phi}(\theta | \hat{\theta}_s)}^{-1},$$

where  $G$  is the group of examinees whose item score for item  $g$  is 1.

The estimated probability,  $\hat{p}(s \in G)$ , or the proportioned weight, in the present study is given by the proportion of the examinees who belong to the group  $G$  within a specified interval whose midpoint is  $\hat{\theta}_s$ . This interval is more or less arbitrary, and we have taken  $\hat{\theta}_s \pm \sigma^2$  and  $\hat{\theta}_s \pm 2\sigma^2$ , where  $\sigma^2$  is the inverse of the constant test information (=21.63, cf. Appendix I) and also the second parameter of the conditional distribution of  $\hat{\theta}$ , given  $\theta$ . Thus these intervals are  $\hat{\theta}_s \pm 0.215$  and  $\hat{\theta}_s \pm 0.430$ , respectively.

If we replace  $\hat{\phi}(\theta|\hat{\theta}_s)$  in (3.3) by  $\phi(\theta|\hat{\theta}_s)$ , then we will obtain the estimated item characteristic function which indicates the maximal possible attainment for the present procedure and data. This function is not obtainable from empirical data, and only the simulation study can provide us with it. As usual, we shall call it the criterion item characteristic function. Since we adopted two different sets of proportioned weights, as was mentioned in the preceding paragraph, we shall obtain two criterion item characteristic functions for each item.

For the same reason described in the preceding study (Samejima, 1978d), we use the two methods, i.e., Pearson System Method and Two-Parameter Beta Method, for obtaining  $\hat{\phi}(\theta|\hat{\theta}_s)$ , and exclude Normal Approach Method.

#### IV Results I: Proportioned Weights

Figure 4-1 presents the estimated probability, or proportioned weight,  $\hat{p}(s \in G)$ , using each of the two intervals,  $\hat{\theta}_s \pm 0.215$  and  $\hat{\theta}_s \pm 0.430$ , for each of the 500 maximum likelihood estimates  $\hat{\theta}_s$ . We can see that these two sets of results, plotted by solid triangles and crosses respectively, are for the most part overlapping, except for some extreme values of  $\hat{\theta}_s$ . This finding suggests that we are likely to obtain similar estimated item characteristic functions as the results of using these two sets of proportioned weights.

It should be noted that these proportioned weights adopted in the Proportioned Sum Procedure in the present study themselves can be regarded as estimates of the item characteristic function of each binary item. Since we have the relationship (Samejima, 1977b)

$$(4.1) \quad \text{Var.}(\hat{\theta}) \doteq \text{Var.}(\theta) + \sigma^2 ,$$

and we obtain

$$(4.2) \quad E(\theta) = 0 ,$$

$$(4.3) \quad \begin{aligned} \text{Var.}(\theta) &= \int_{-\infty}^{\infty} \theta^2 f(\theta) d\theta = 0.2 \left[ \frac{1}{3} \theta^3 \right]_{-2.5}^{2.5} \\ &= \frac{25}{12} \doteq 2.083333 , \end{aligned}$$

$$(4.4) \quad \text{Var.}(\hat{\theta})^{1/2} \doteq [2.083333 + 0.046225]^{1/2} \doteq 1.459301$$

and

$$(4.5) \quad \text{Var.}(\theta)^{1/2} \doteq 1.443376 ,$$

it is expected that the abscissa be stretched by approximately 1.1 %

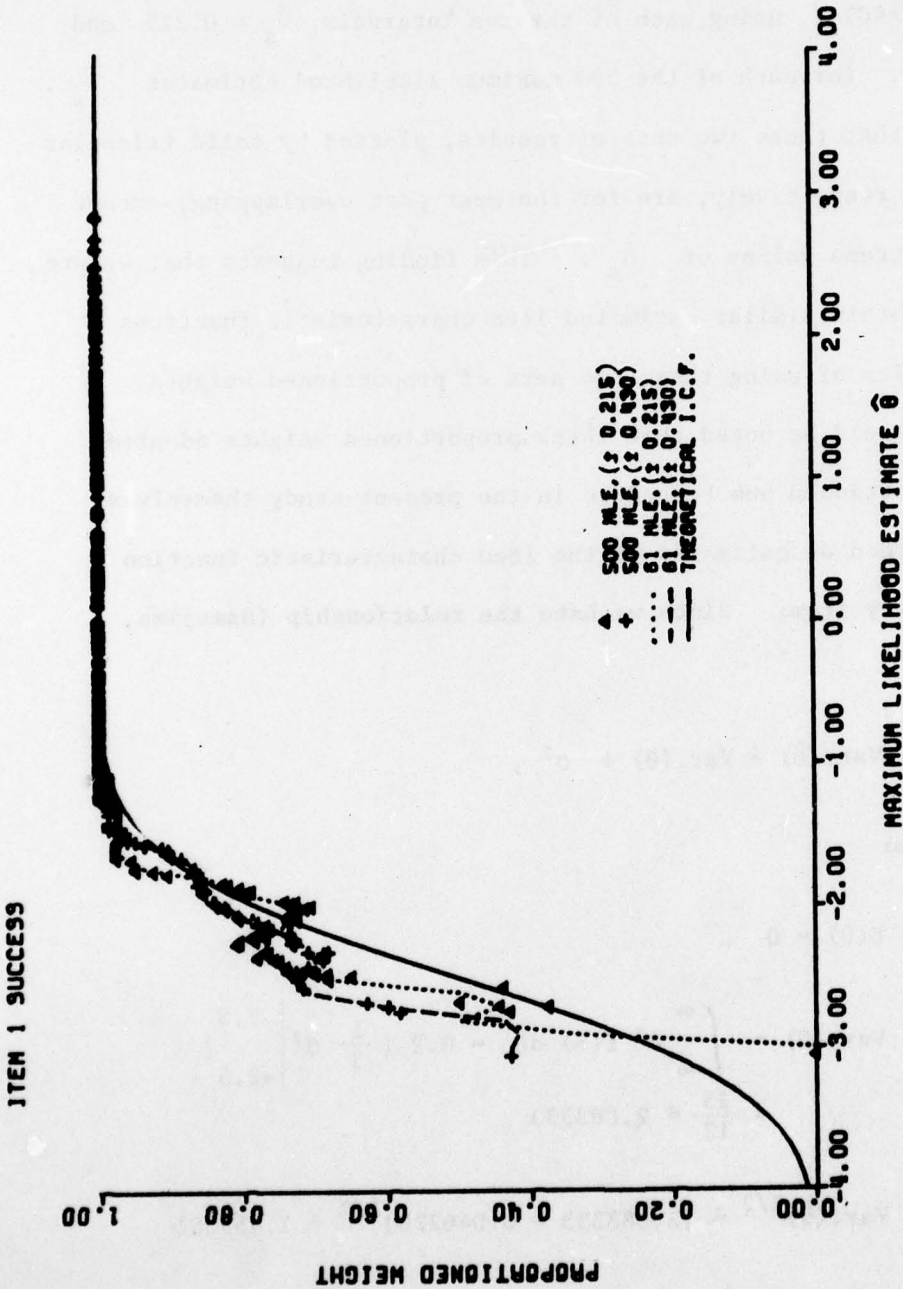


FIGURE 4-1

Four Sets of Proportioned Weights for the Success Group of the Item, Used in the Proportioned Sum Procedure, Presented with the Item Characteristic Function. Two of Them Are Based on the 500 Maximum Likelihood Estimates (MLE), with the Intervals  $\pm 0.215$  and  $\pm 0.430$ , Respectively, and the Other Two Are Based on the 61 Equally Intervalled Points of  $\theta$  with the Respective Intervals.

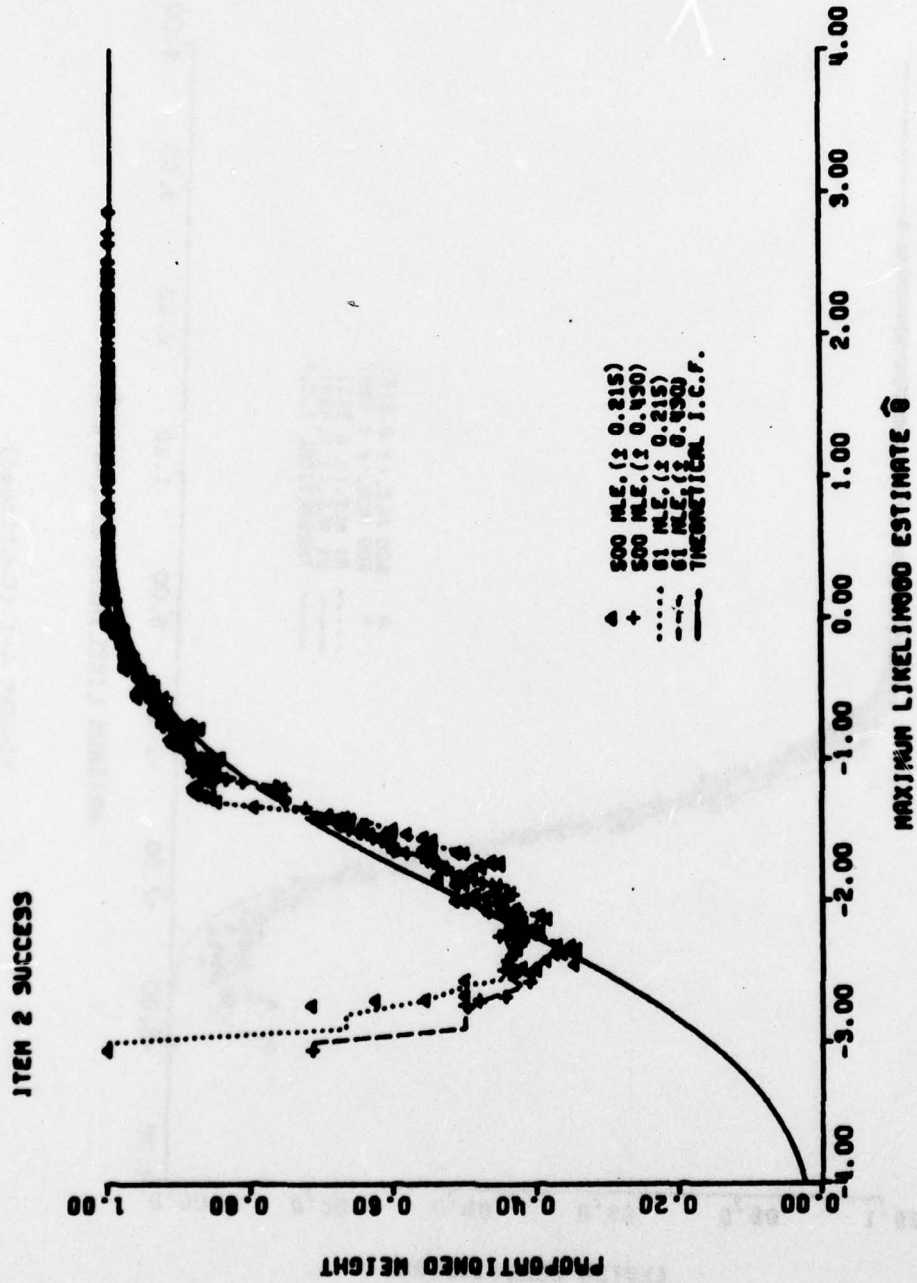


FIGURE 4-1 (Continued)

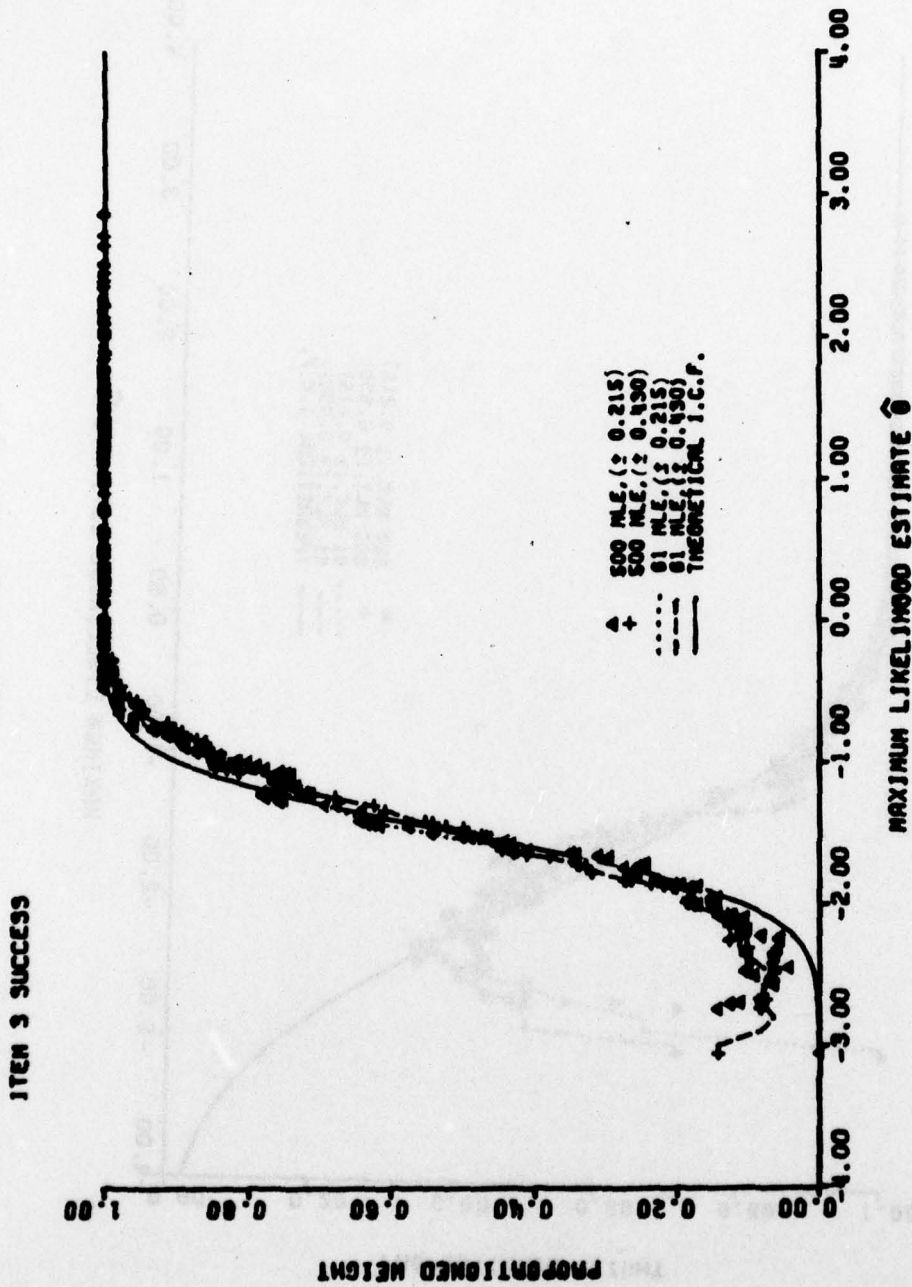


FIGURE 4-1 (Continued)



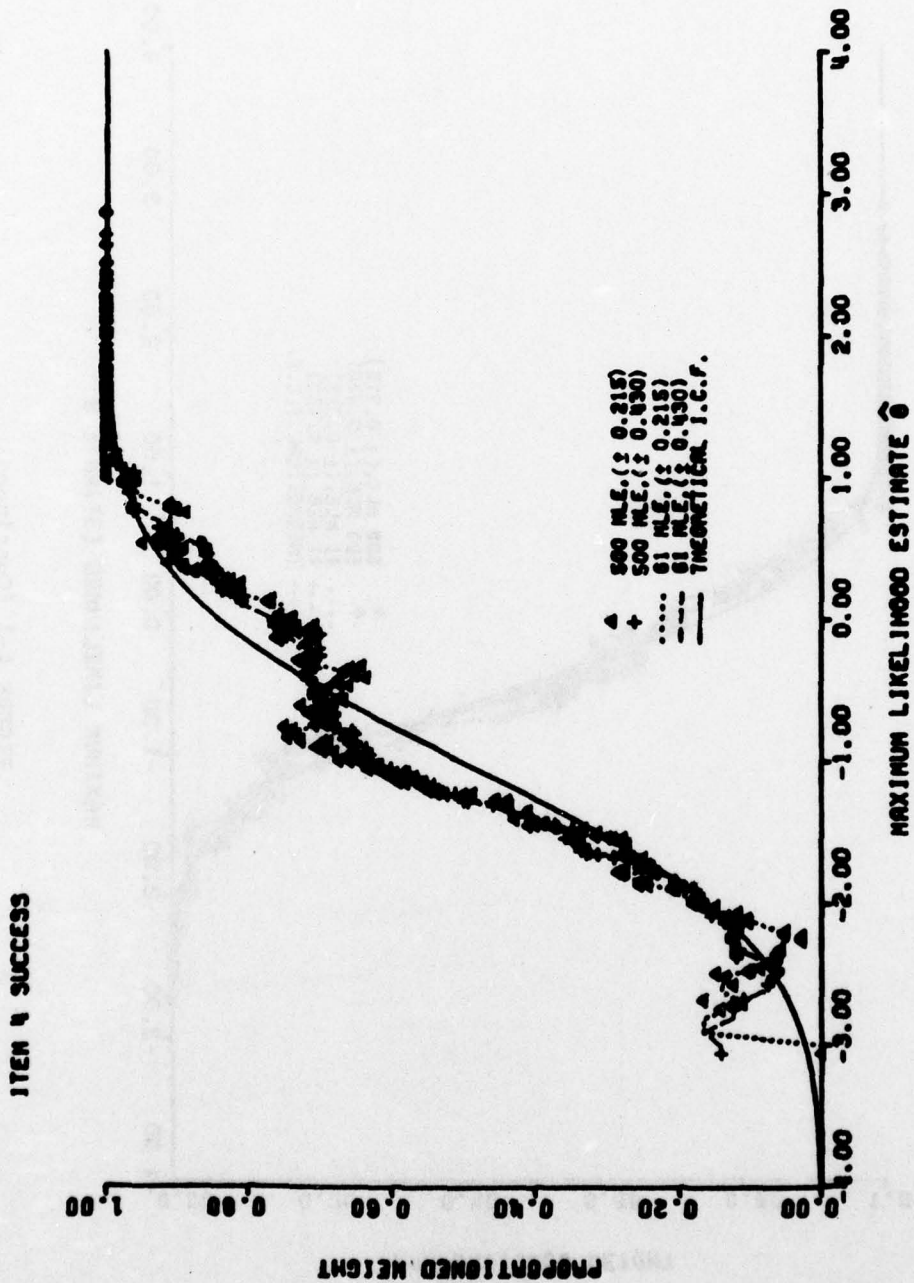


FIGURE 4-1 (Continued)

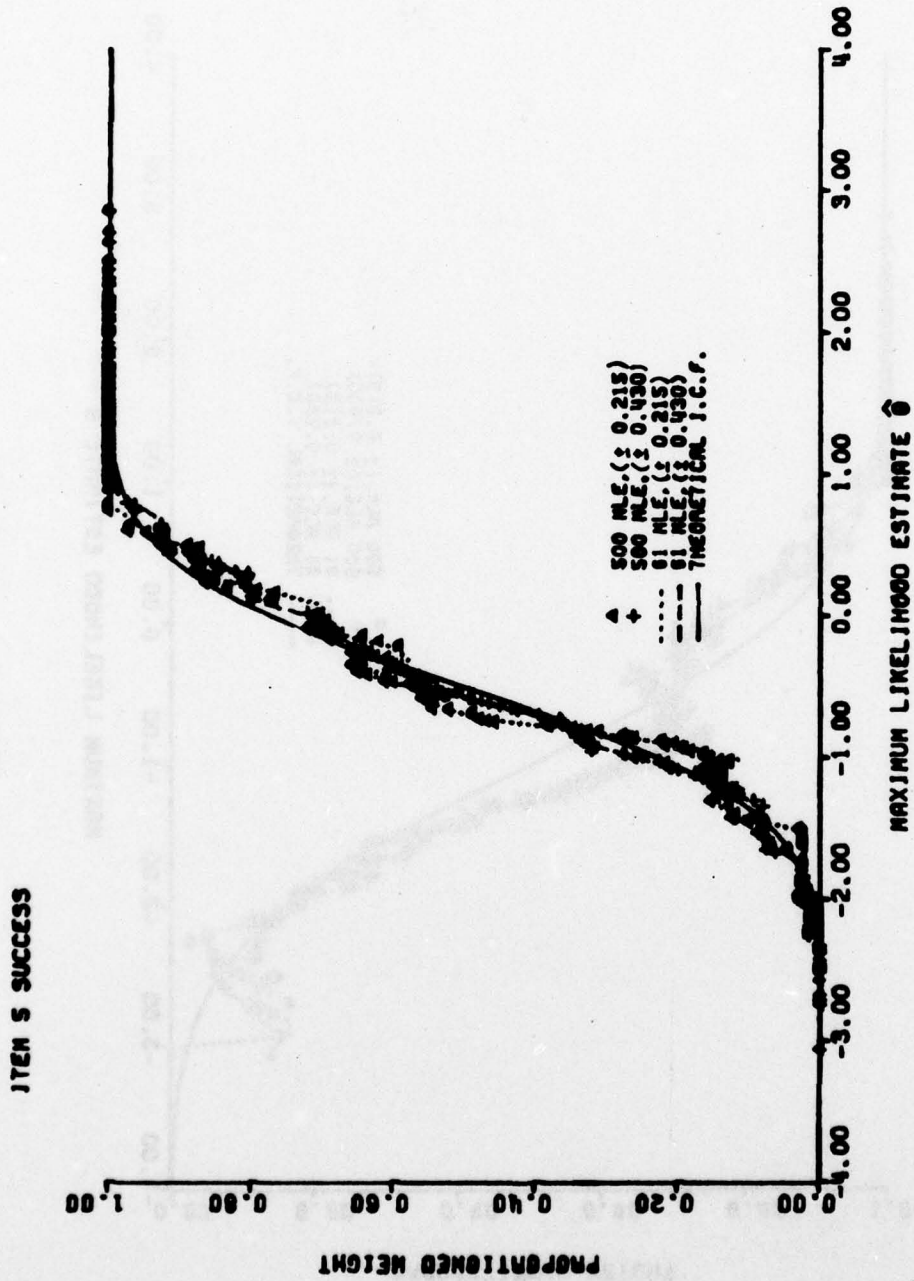


FIGURE 4-1 (Continued)

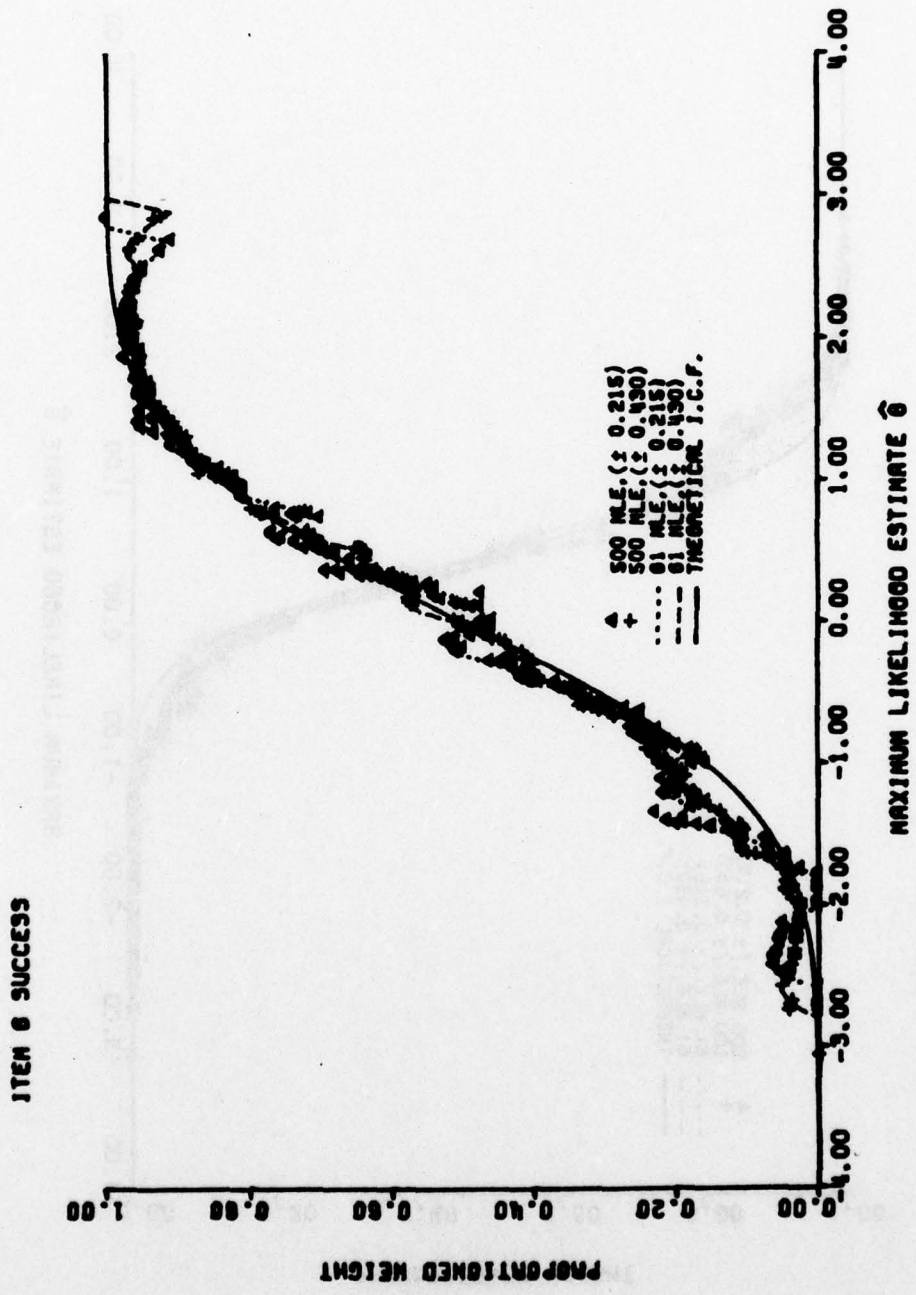


FIGURE 4-1 (Continued)

ITEM 7 SUCCESS

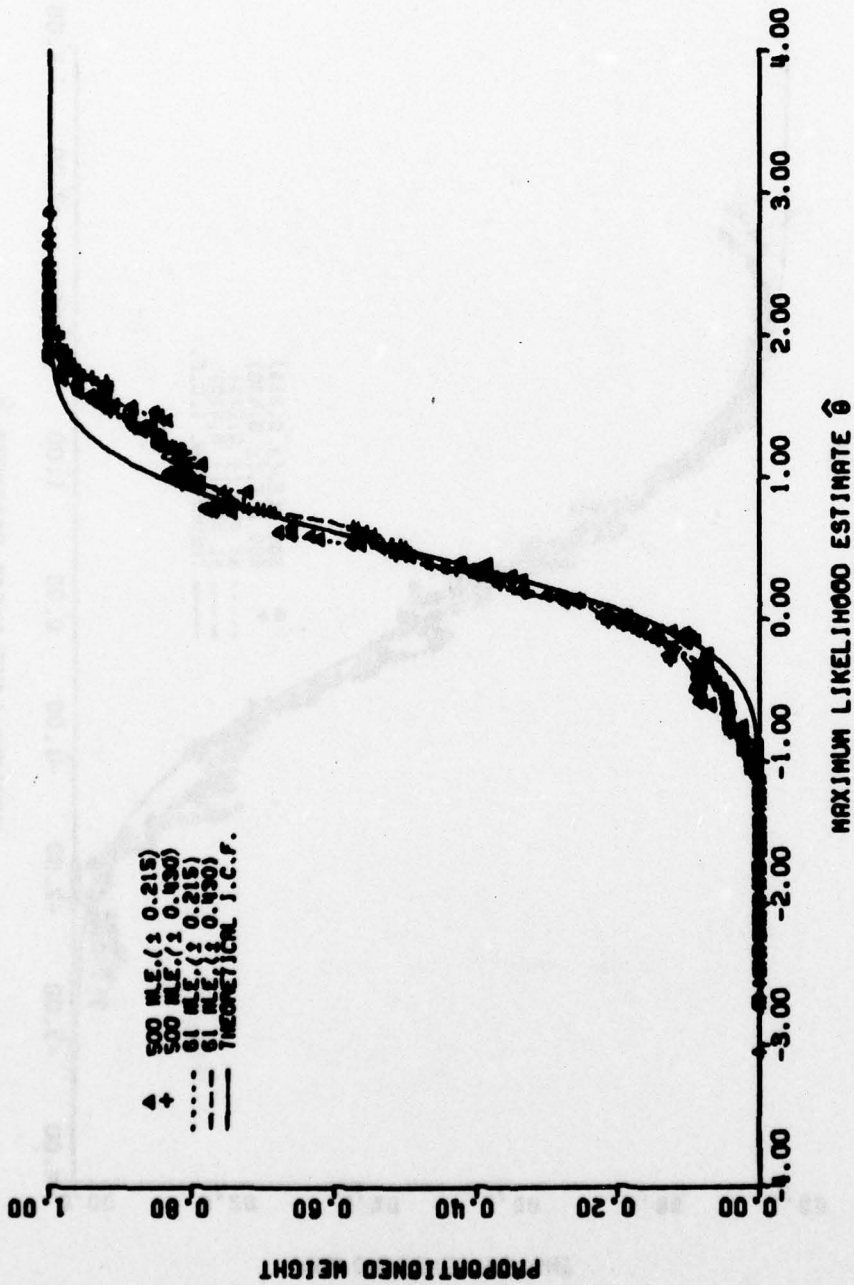


FIGURE 4-1 (Continued)

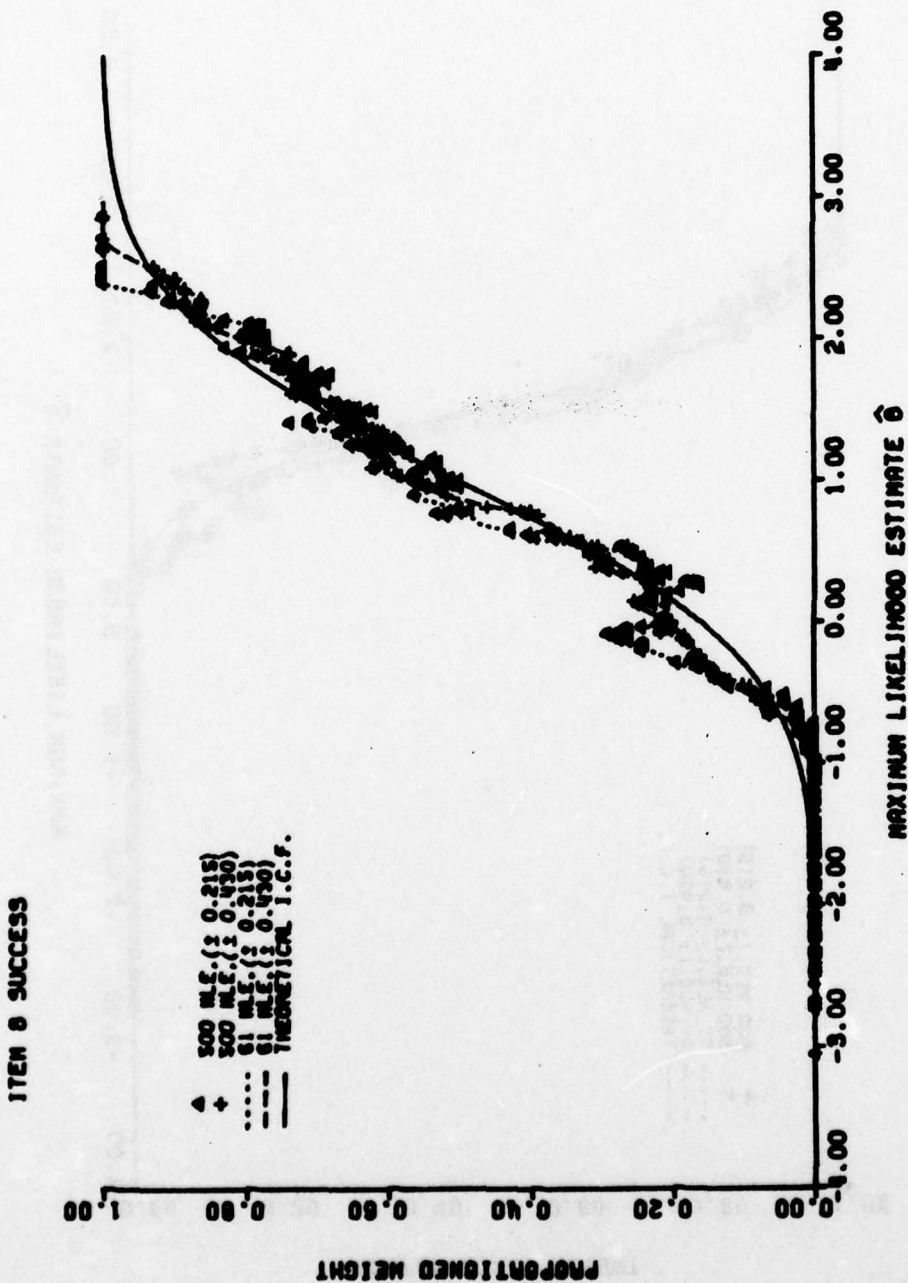


FIGURE 4-1 (Continued)

ITEM 9 SUCCESS

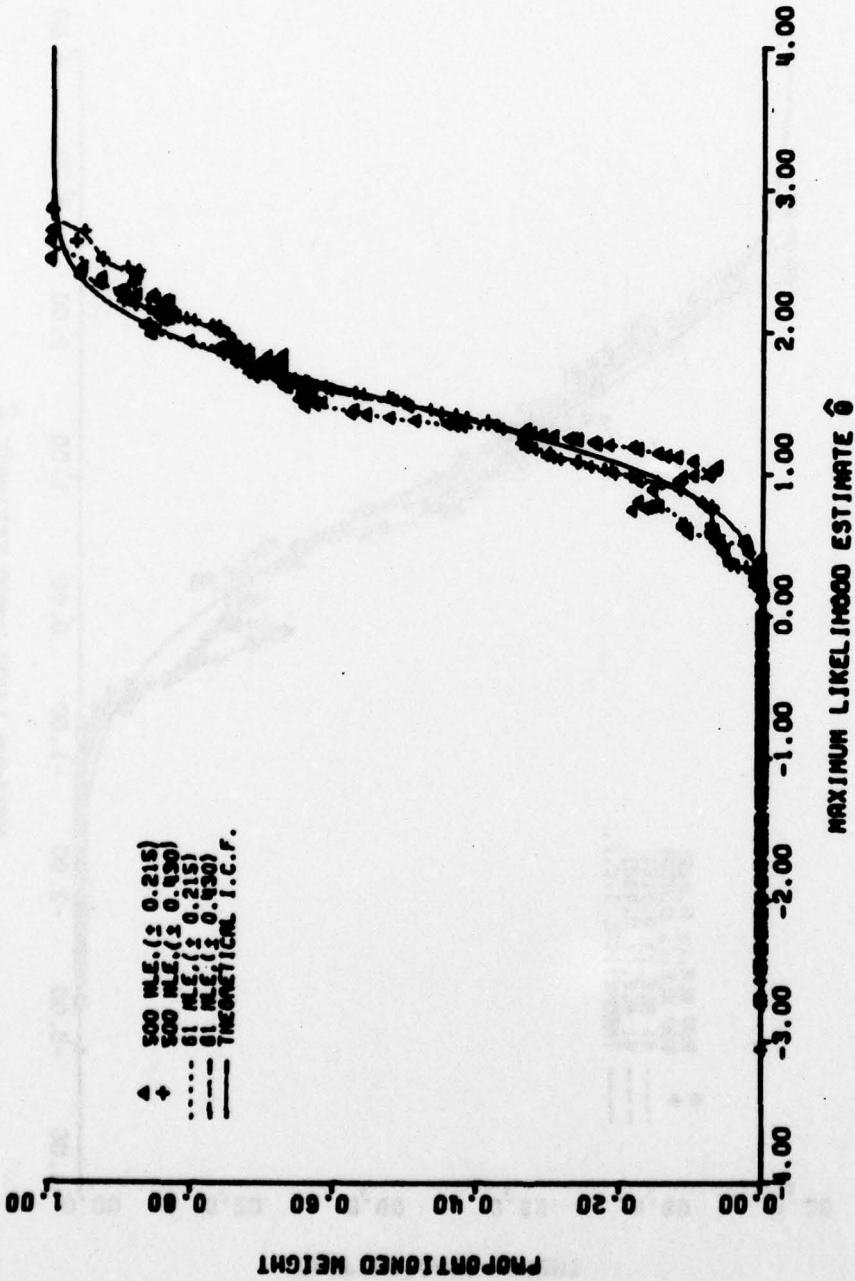


FIGURE 4-1 (Continued)

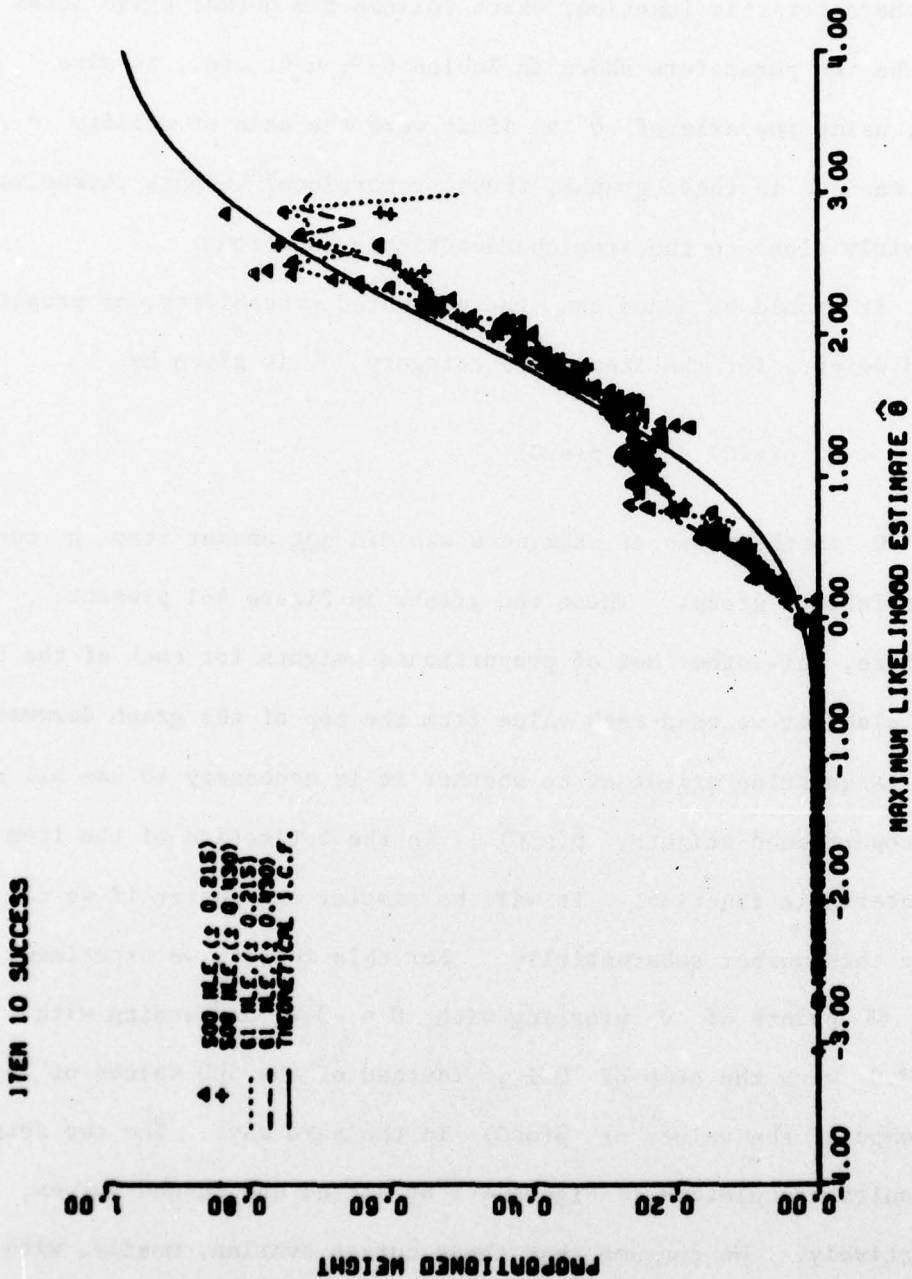


FIGURE 4-1 (Continued)

for these estimates. By virtue of the fact that this percentage is negligibly small, in each graph of Figure 4-1, the theoretical item characteristic function, which follows the normal ogive model with the two parameters shown in Tables 6-3, 6-4, etc., is also drawn, using the axis of  $\hat{\theta}$  as if it were the axis of ability  $\theta$ . As we can see in these graphs, these proportional weights themselves are fairly close to the item characteristic functions.

It should be added that the estimated probability, or proportioned weight, for the item score category 0 is given by

$$(4.6) \quad \hat{p}(s \in \bar{G}) = 1 - \hat{p}(s \in G) ,$$

where  $\bar{G}$  is the group of examinees who did not answer item  $g$  correctly, or the failure group. These ten graphs in Figure 4-1 present, therefore, this other set of proportioned weights for each of the two cases also, if we read each value from the top of the graph downwards.

A question arises as to whether it is necessary to use all the 500 proportioned weights,  $\hat{p}(s \in G)$ , in the estimation of the item characteristic function. It will be simpler and easier if we can reduce this number substantially. For this reason, we experiment using 61 points of  $\hat{\theta}$  starting with  $\hat{\theta} = -3.0$  and ending with  $\hat{\theta} = 3.0$  with the step of 0.1, instead of the 500 values of  $\hat{\theta}_s$ , and compute the values of  $\hat{p}(s \in G)$  in the same way. The two sets of results are plotted in Figure 4-1 by dotted and dashed curves, respectively. We can see that these curves overlap, mostly, with corresponding sets of plots for the 500  $\hat{\theta}_s$ 's, as is naturally expected.



V Results II: Estimated Density Functions of Ability

Figure 5-1 presents the estimated density function,  $\hat{f}(\theta)$ , obtained by the Proportioned Sum Procedure, using the 500  $\hat{\theta}_s$ 's, by dotted line, which is also the denominator of (3.3) with  $\hat{\phi}(\theta|\hat{\theta}_s)$  replaced by its theoretical counterpart,  $\phi(\theta|\hat{\theta}_s)$ , and is divided by the total number of examinees, 500. Note that this estimated density function is the same as the one obtained by the Simple Sum Procedure (Samejima, 1978d) of the Conditional P.D.F. Approach. For comparison, the estimated density function of  $\theta$  obtained by the Pseudo Criterion Degree 4 Case of the Weighted Sum Procedure (Samejima, 1978d) is drawn by broken and dotted line, together with the theoretical density function,  $f(\theta)$ . We can see that the  $\hat{f}(\theta)$  obtained by the Proportioned Sum Procedure is reasonably close to the theoretical density function, although not as close as the one by the Pseudo Criterion Degree 4 Case of the Weighted Sum Procedure.

As an additional information, the same denominator computed for the 61 points of  $\hat{\theta}$  and divided by 61 is presented in the same figure by dashed line. This is by no means an estimate of the density function  $f(\theta)$ , and it is interesting to note that the curve is substantially different from the others.

The mean square error which is obtained from

$$(5.1) \quad \frac{1}{m} \sum_{j=1}^m [\hat{f}(\theta_j) - f(\theta_j)]^2 ,$$

where  $m = 25$  and  $\theta_j$ 's are  $-2.4$  through  $2.4$  with the step  $0.2$  is computed, and it turned out to be  $0.00018$  ( $0.01324$ ), which is a little greater than  $0.00009$  ( $0.00972$ ) for the  $\hat{f}(\theta)$  by the Pseudo Criterion

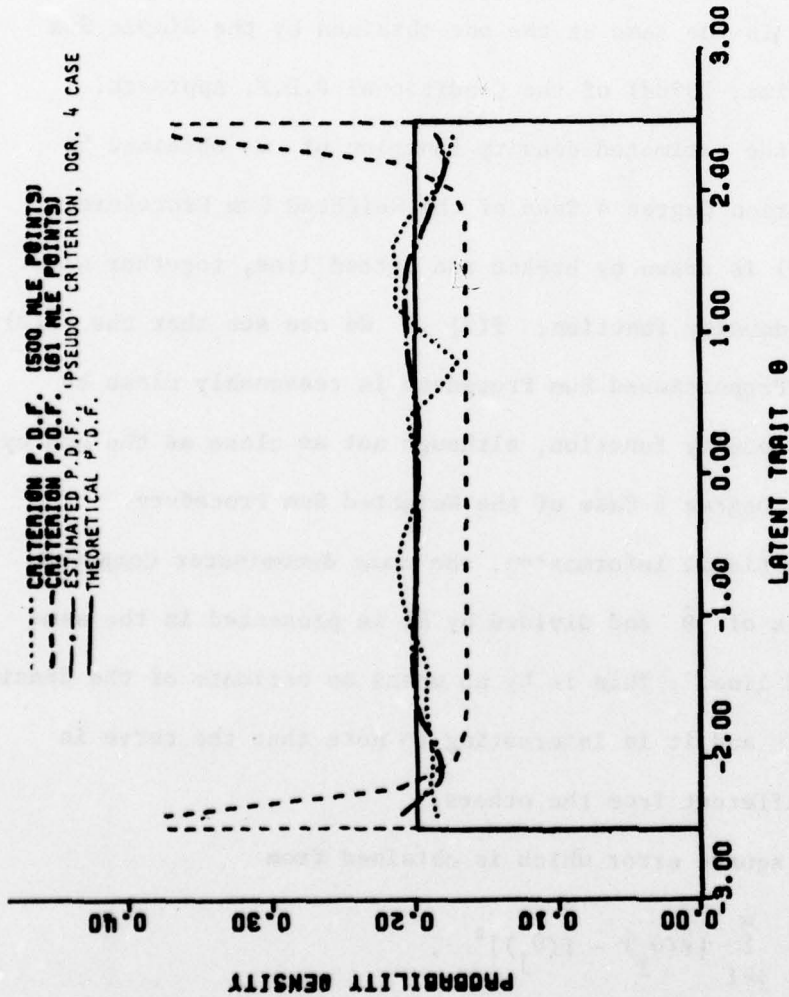


FIGURE 5-1

Estimated Density Function of Ability  $\theta$  Obtained by the Proportioned Sum Procedure of the Conditional P.D.F. Approach, Which Is Based on the 500 True Conditional Density Functions,  $\phi(\theta|\theta)$ , Together with the One Obtained by the Pseudo Criterion Degree 4 Case of the Weighted Sum Procedure and the Theoretical Density Function  $f(\theta)$ . Also Presented Is the One Based on the 61 Points of  $\hat{\theta}$ .

Degree 4 Case of the Weighted Sum Procedure. (The numbers in parentheses are the corresponding square roots of the mean square errors.) For an additional information, if we calculate the value given by (5.1) for the 61 point case, it will be 0.00350 (0.05918).

The estimated shared density function (Samejima, 1978d) for each of the failure and success groups of each item is presented as Figure 5-2, for the two criterion cases of the present study, together with the one obtained by the Simple Sum Procedure, the actual frequency ratios of  $\theta$  and the theoretical shared density. (Again, for an additional information, the corresponding curves for the 61 point case are also presented in the same figure, which are by no means estimated shared densities.) We find that these two curves are fairly close to the one by the Simple Sum Procedure, which indicates that they are not very close to the three curves provided by the Pseudo Criterion Degree 3, 4 and 5 Cases of the Weighted Sum Procedure (Samejima, 1978d).

Figure 5-3 presents the estimated density functions,  $\hat{f}(\theta)$ , obtained by Degree 3 and 4 Cases\* of the Pearson System Method. We can see that these results, especially the one obtained by Degree 3 Case, are very similar to the one obtained by the criterion case.

The means square errors of the  $\hat{f}(\theta)$ 's obtained by Degree 3 and 4 Cases of the Pearson System Method and their square roots were calculated through (5.1), and they are 0.00076 (0.02752) and 0.00105 (0.03242) respectively. As was expected, they are reasonably small, but not as small as those for the criterion case, which turned out

---

\* Refer to Appendix II.

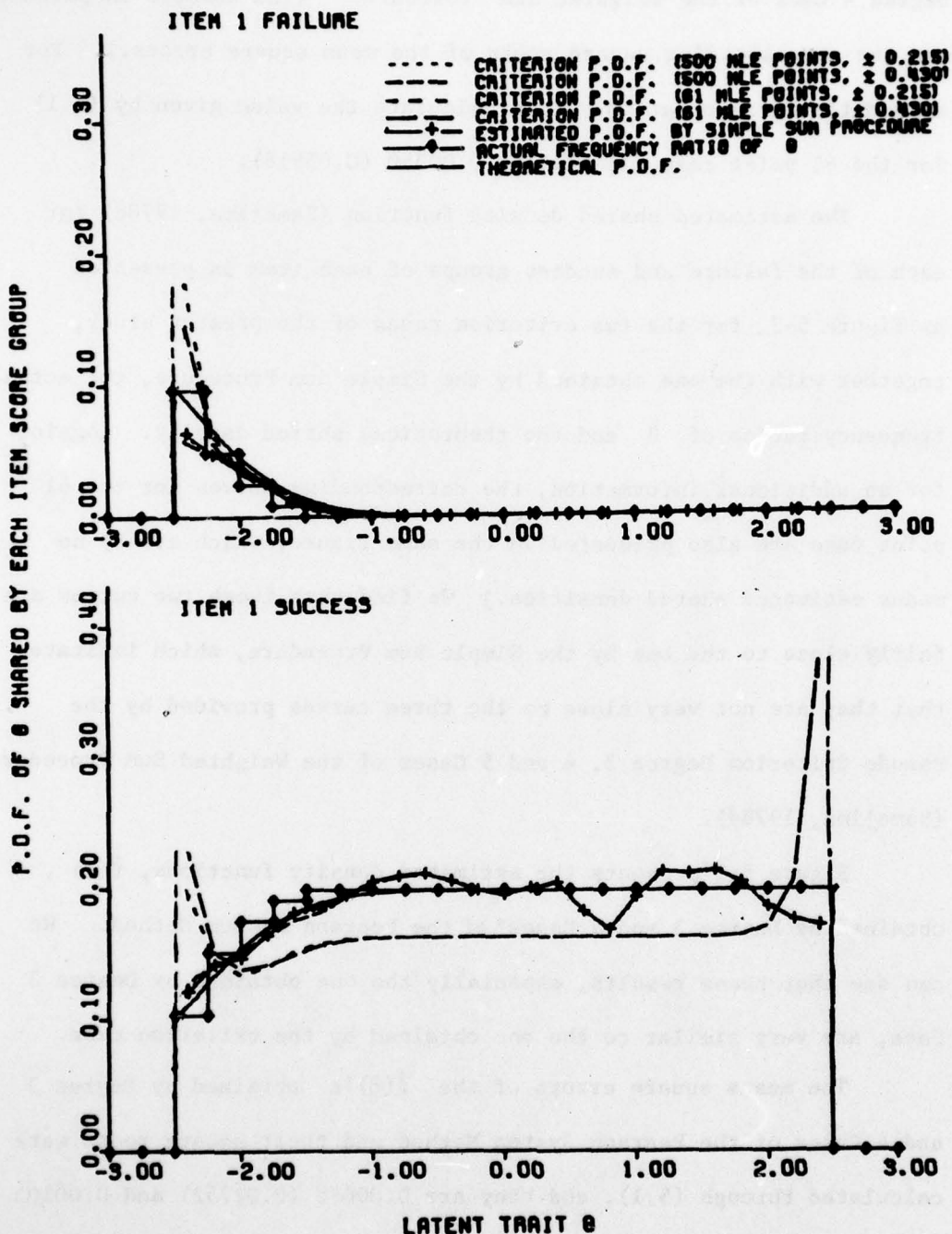


FIGURE 5-2

Comparison of the Four Estimated Shared Density Functions of Ability  $\theta$  of Each of the Two Item Score Groups, Which are Obtained by the Proportioned Sum Procedure, Using the True Conditional Density Function,  $\phi(\theta|\theta_s)$ , with the One Obtained by the Simple Sum Procedure. Theoretical Shared Density Function and Actual Frequency Ratios are Also Presented.

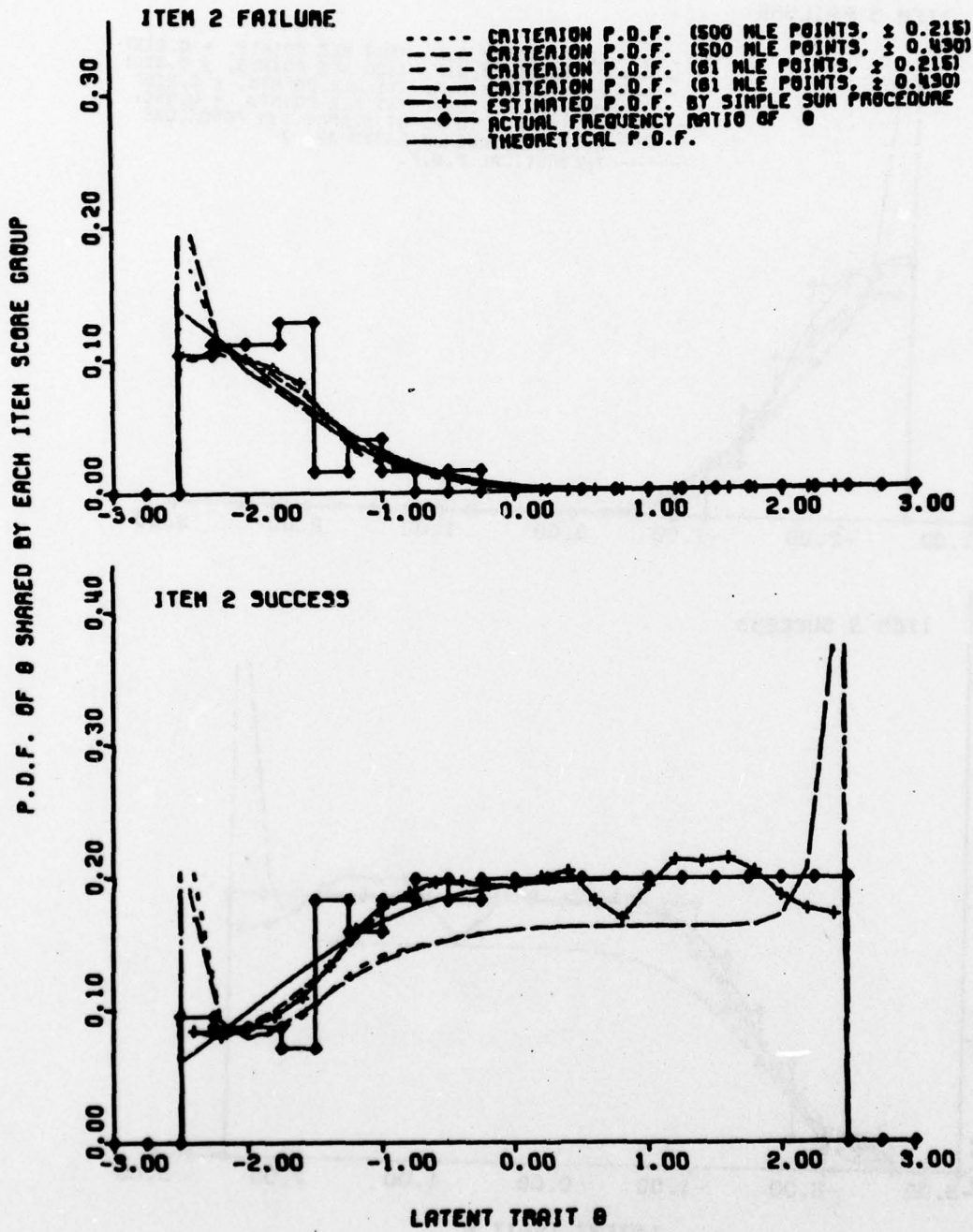


FIGURE 5-2: Criterion Case (Continued)

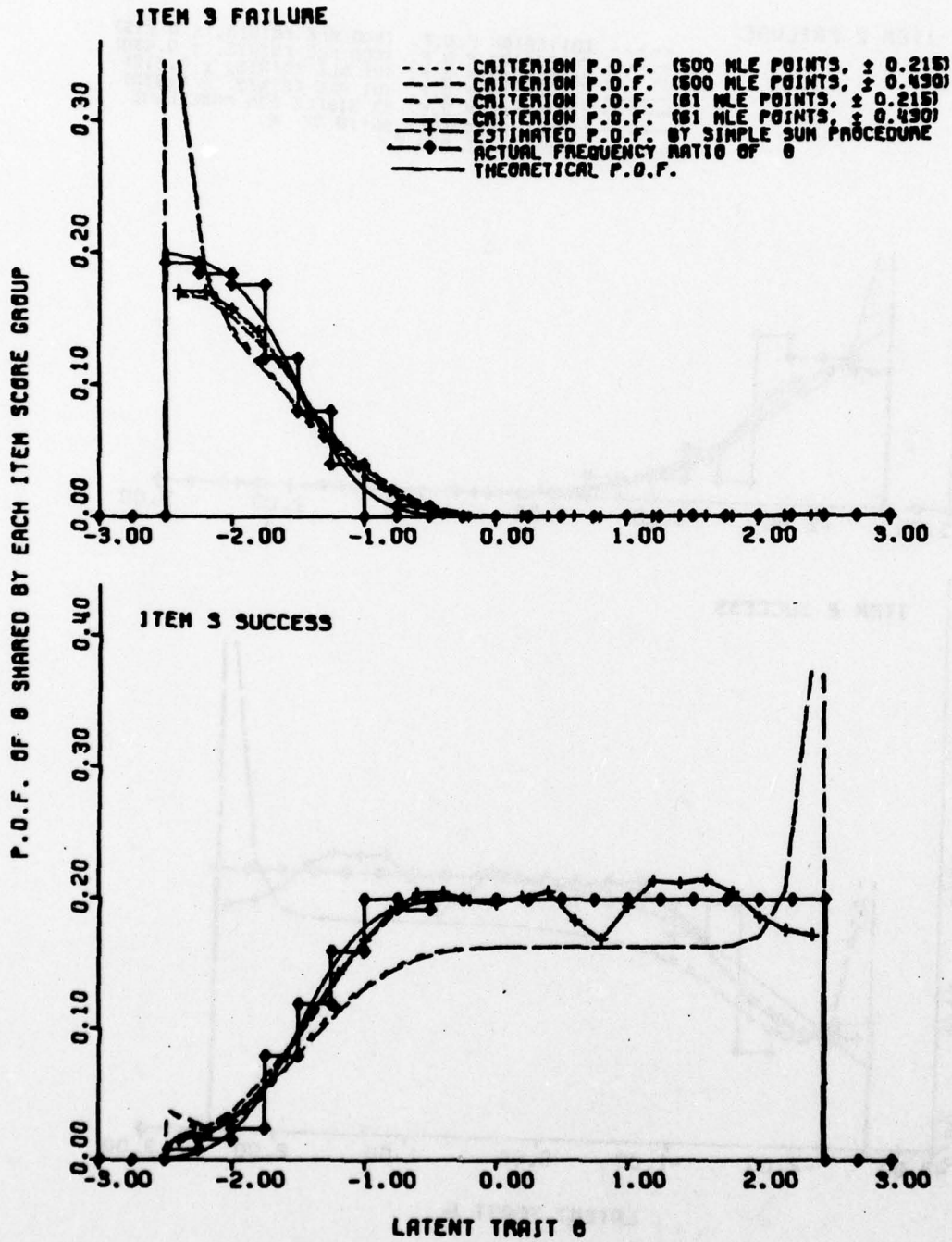


FIGURE 5-2: Criterion Case (Continued)

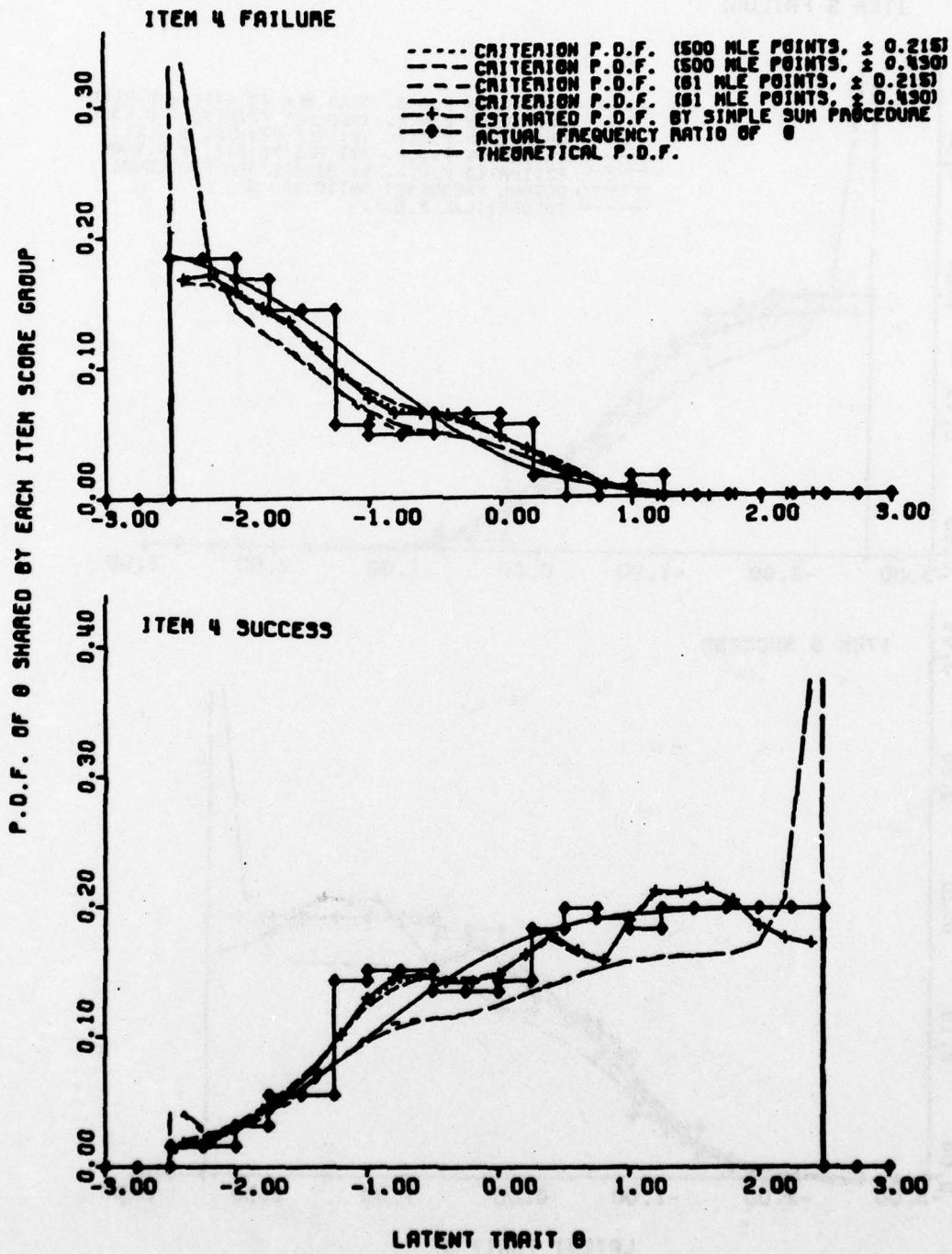
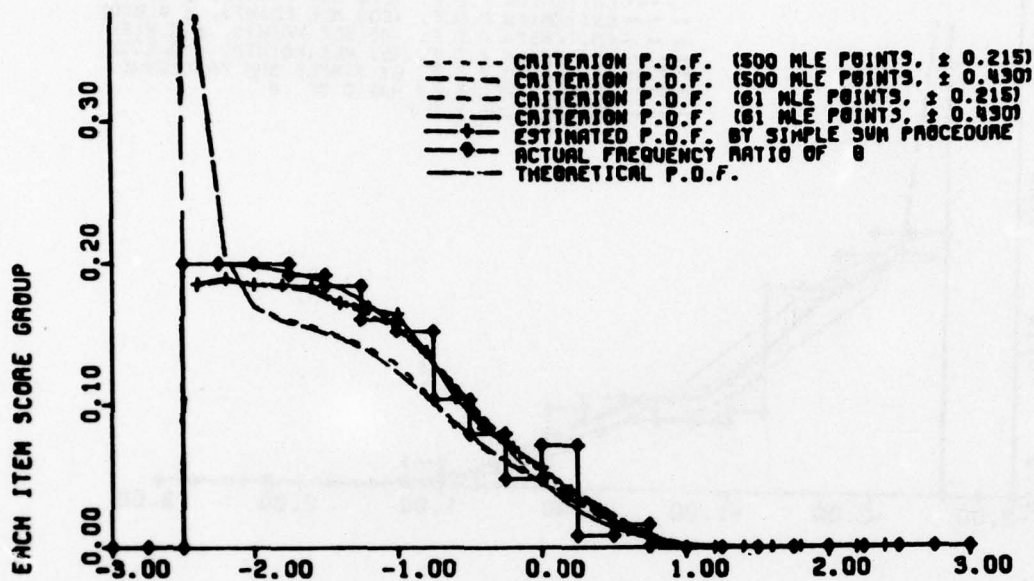


FIGURE 5-2: Criterion Case (Continued)

ITEM 5 FAILURE



ITEM 5 SUCCESS

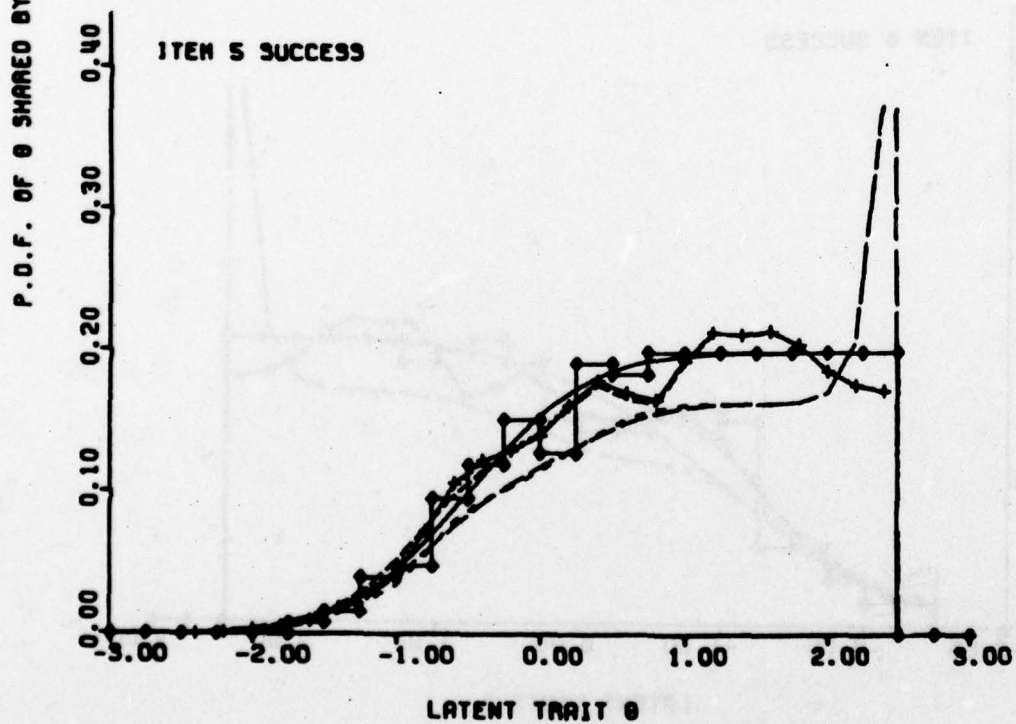


FIGURE 5-2: Criterion Case (Continued)



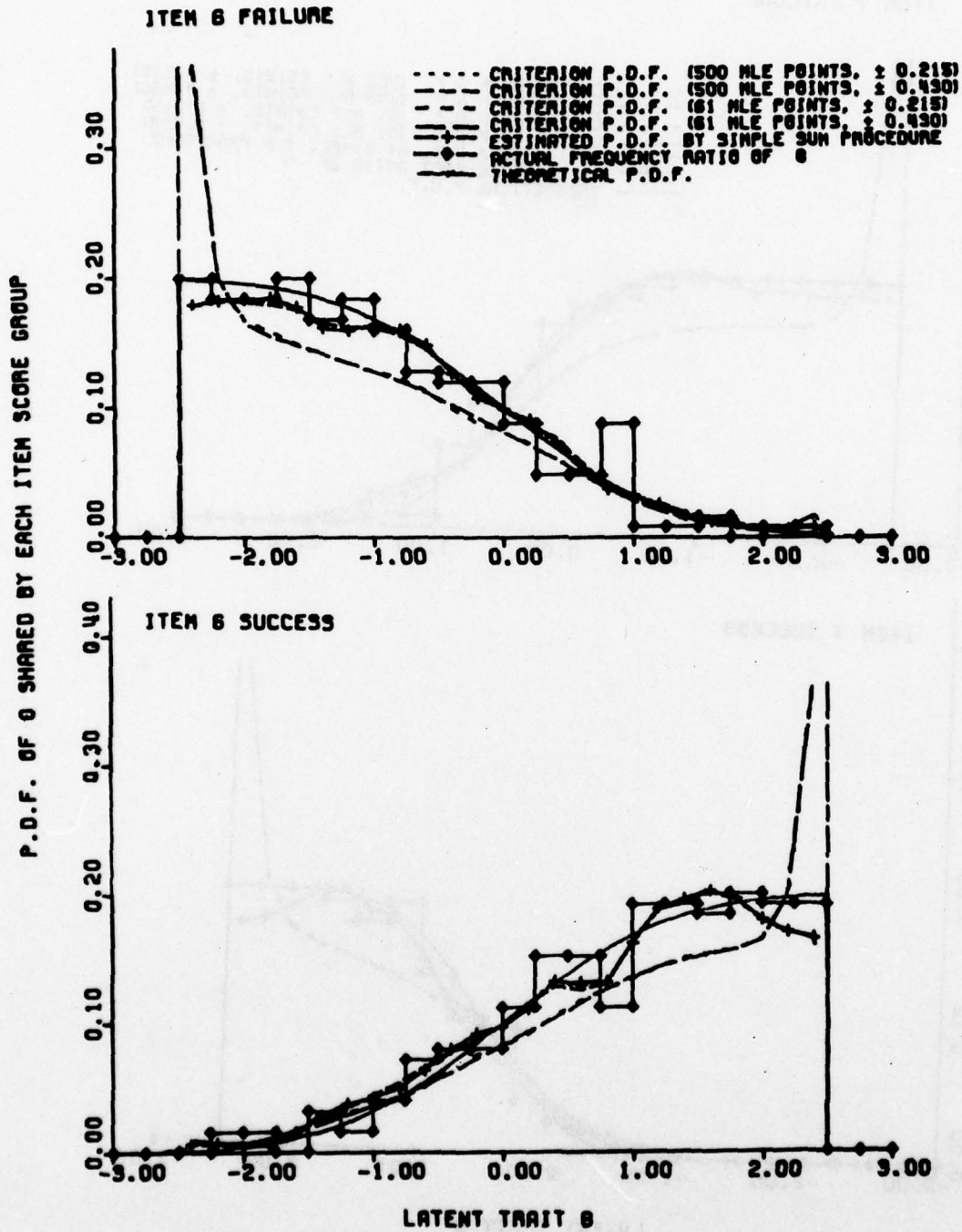


FIGURE 5-2: Criterion Case (Continued)

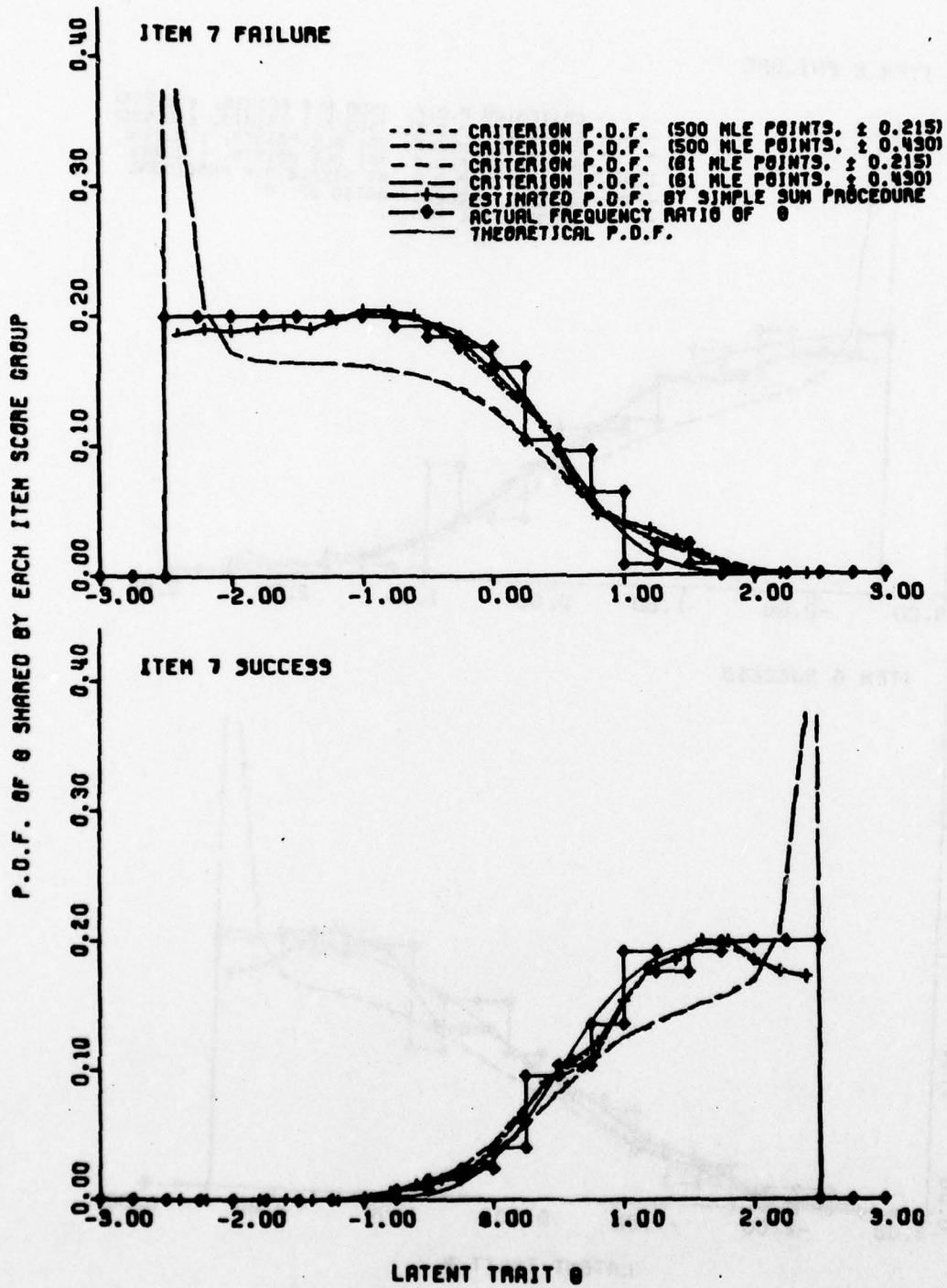


FIGURE 5-2: Criterion Case (Continued)

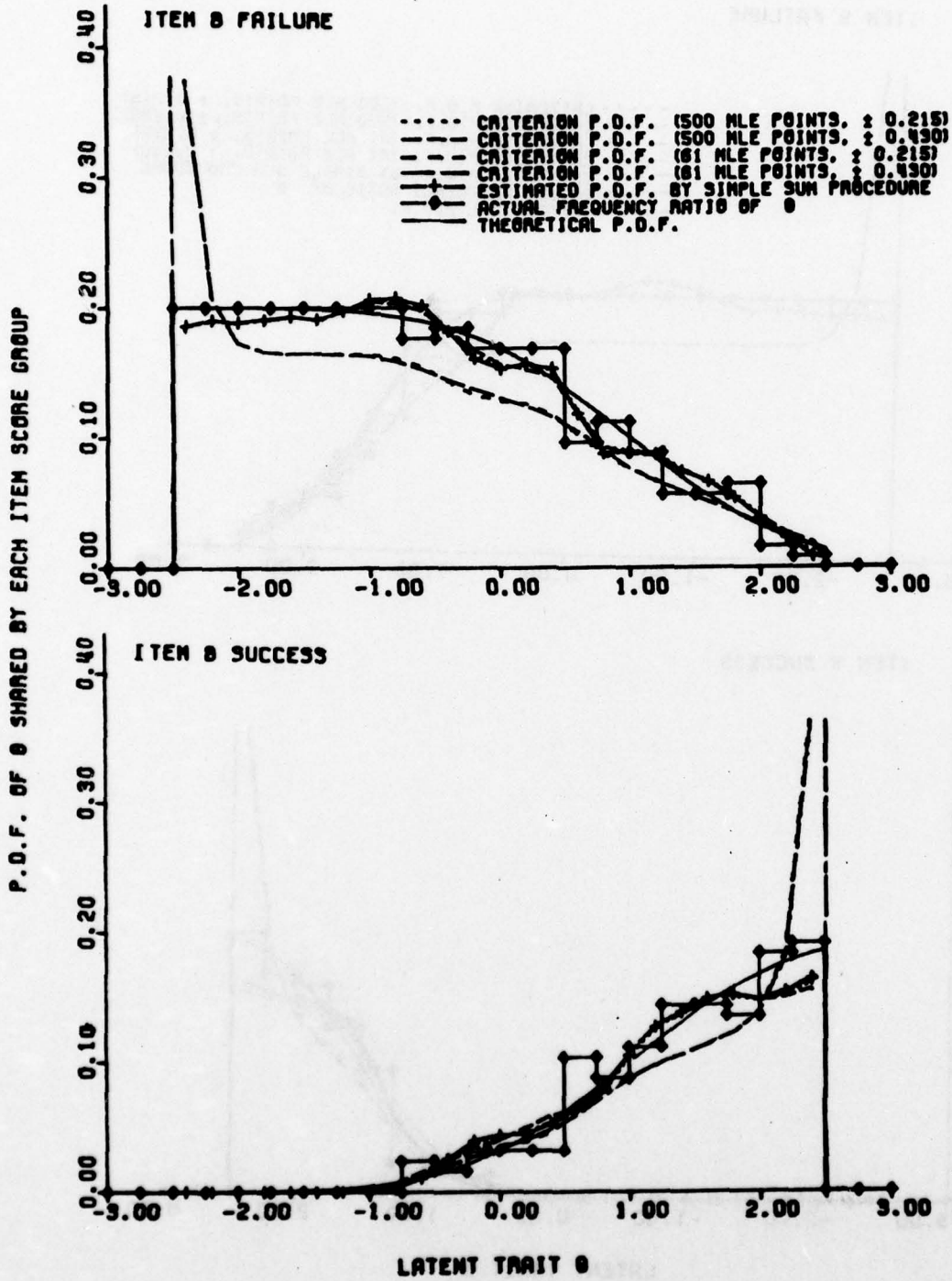


FIGURE 5-2: Criterion Case (Continued)

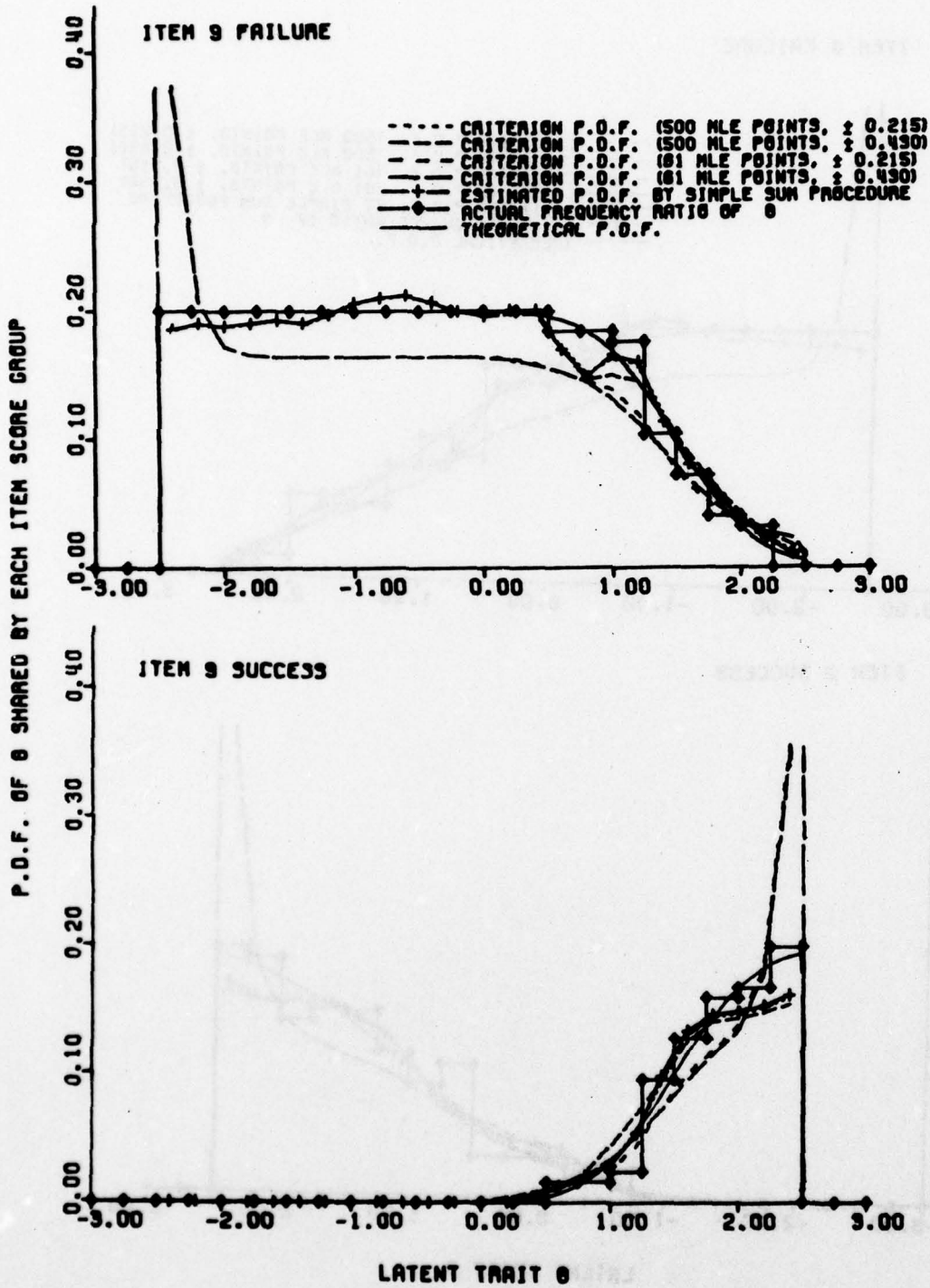


FIGURE 5-2: Criterion Case (Continued)

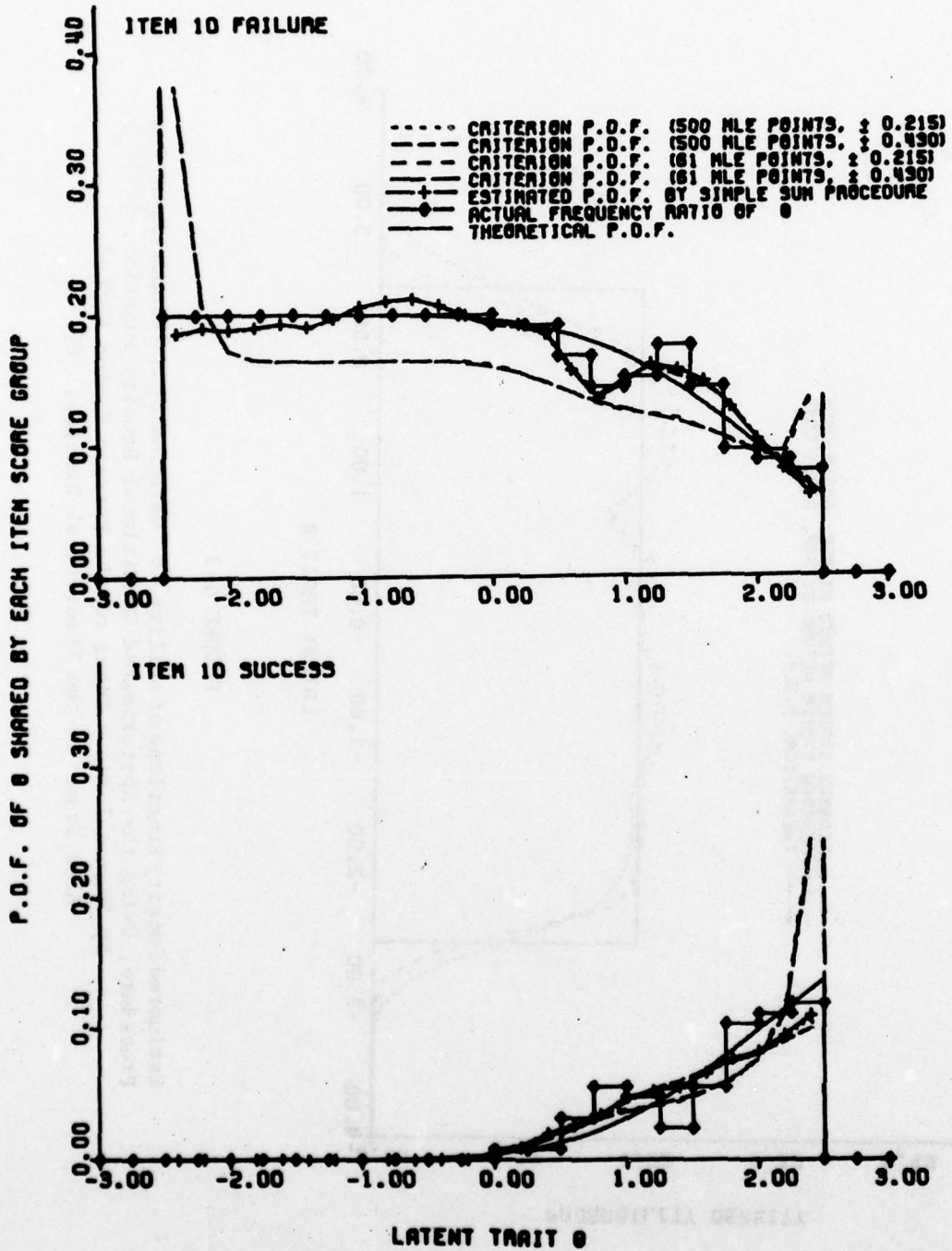


FIGURE 5-2: Criterion Case (Continued)

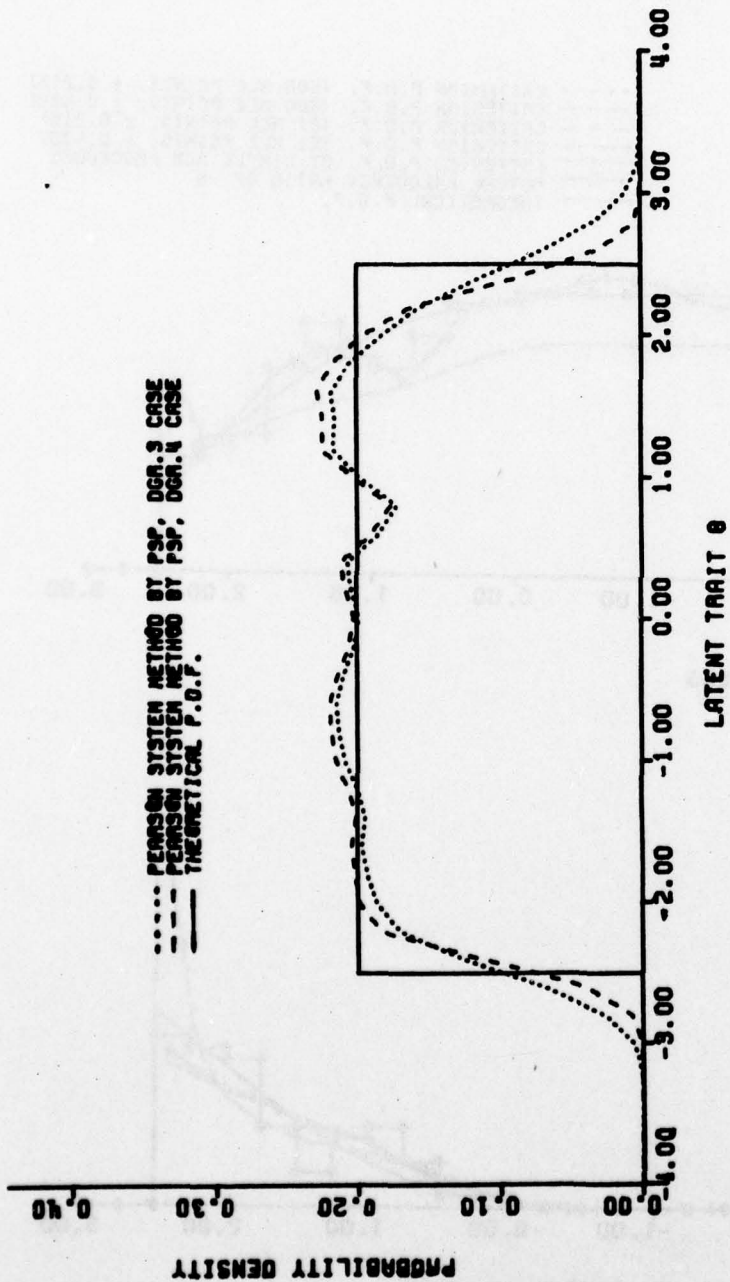


FIGURE 5-3

Estimated Density Functions of Ability  $\theta$  Obtained by the Proportioned Sum Procedure, Using the Approximated Conditional Density Function,  $\hat{\phi}(\theta|\hat{\theta}_s)$ , by Degree 3 and 4 Cases of the Pearson System Method.  
Also Drawn is the Theoretical Density,  $f(\theta)$ .

to be 0,00018 (0,01324), as was mentioned earlier,

The estimated shared density functions by the failure and success groups obtained by Degree 3 and 4 Cases of the Pearson System Method for the two intervals of  $\hat{\theta}_s$  each are presented in Figure 5-4, along with those obtained by the criterion case of the Simple Sum Procedure, the actual frequency ratios and the theoretical shared density functions. We find that the two curves for the different proportioned weights are very close to each other, in each of Degree 3 and 4 Cases.

Figure 5-5 presents the estimated density functions,  $\hat{f}(\theta)$ , obtained by Degree 3 and 4 Cases of the Two-Parameter Beta Method. Just like those obtained by the Pearson System Method, these two curves have "tails" outside the interval of  $\theta$ ,  $(-2,5, 2,5)$ , for which the theoretical density,  $f(\theta)$ , assumes a constant positive value, 0,2. It is noticed that within the interval of  $\theta$  these two curves are very close to each other, and, in fact, more so than those obtained by the Pearson System Method.

The mean square errors for these two estimated density functions from the theoretical density functions, which were computed through (5.1), turned out to be 0,00077 (0,02766) and 0,00064 (0,02529) for Degree 3 and 4 Cases, respectively. They are both comparable to those by Degree 3 Case of the Pearson System Method, i.e., 0,00076 (0,02752).

The estimated shared density functions of  $\theta$  by the failure and success groups obtained by the Two-Parameter Beta Method are presented as Figure 5-6, by the four curves each of which represents Degree 3 Case with the interval  $\hat{\theta}_s \pm 0,215$ , Degree 3 Case with  $\hat{\theta}_s \pm 0,430$ , Degree 4 Case with  $\hat{\theta}_s \pm 0,215$ , or Degree 4 Case with

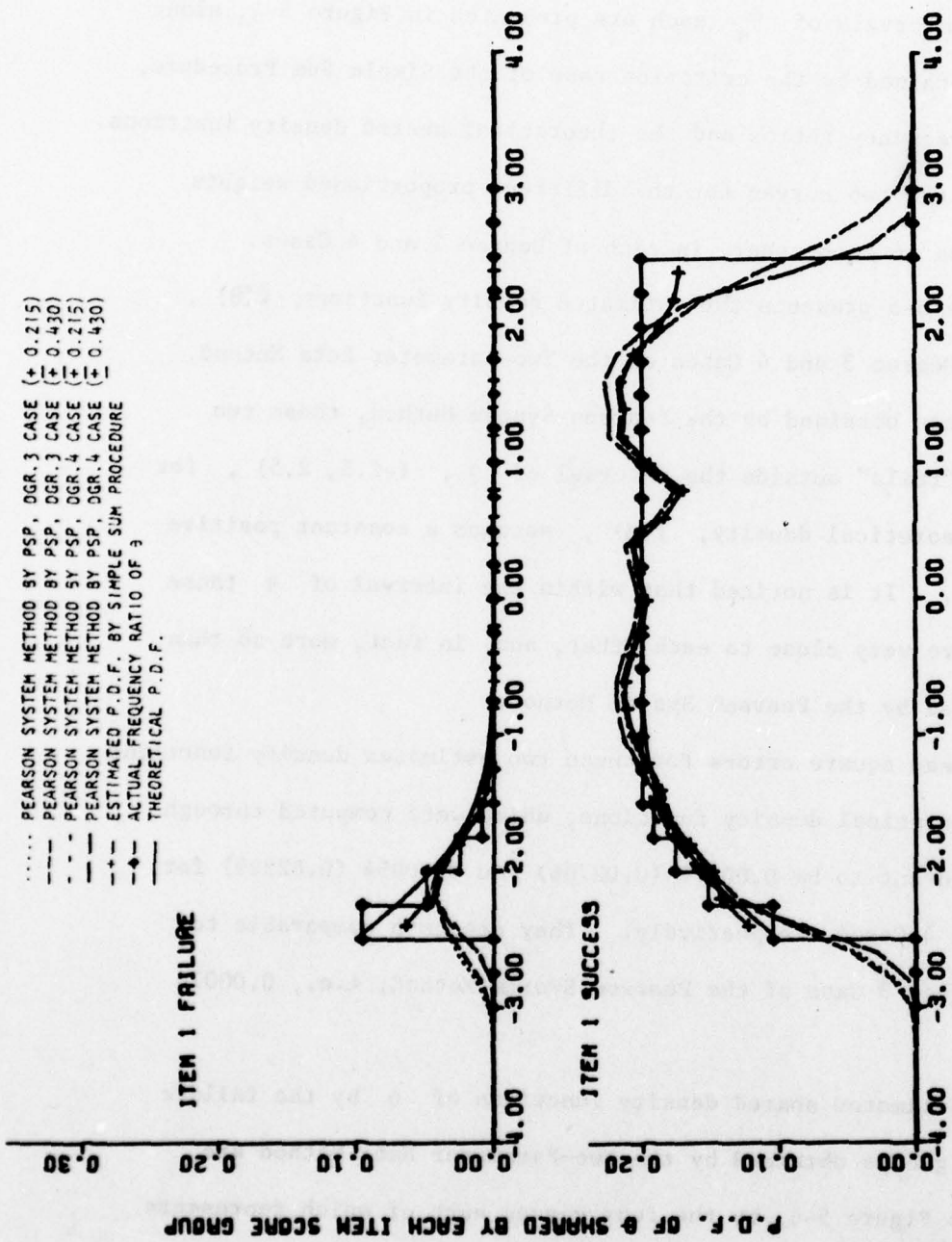


FIGURE 5-4

Four Estimated Shared Density Functions of Ability  $\theta$  of Each of the Two Item Score Groups, Which Are Obtained by the Proportioned Sum Procedure Using the Approximated Conditional Density Function,  $\hat{\phi}(\theta|\hat{\theta}_s)$ , by Degree 3 and 4 Cases of the Pearson System Method. Also Presented Are the One Obtained by Using the True Conditional Density Function,  $\phi(\theta|\theta_s)$ , by the Simple Sum Procedure, Actual Frequency Ratios of Ability  $\theta_s$ , and Theoretical Shared Density Function.



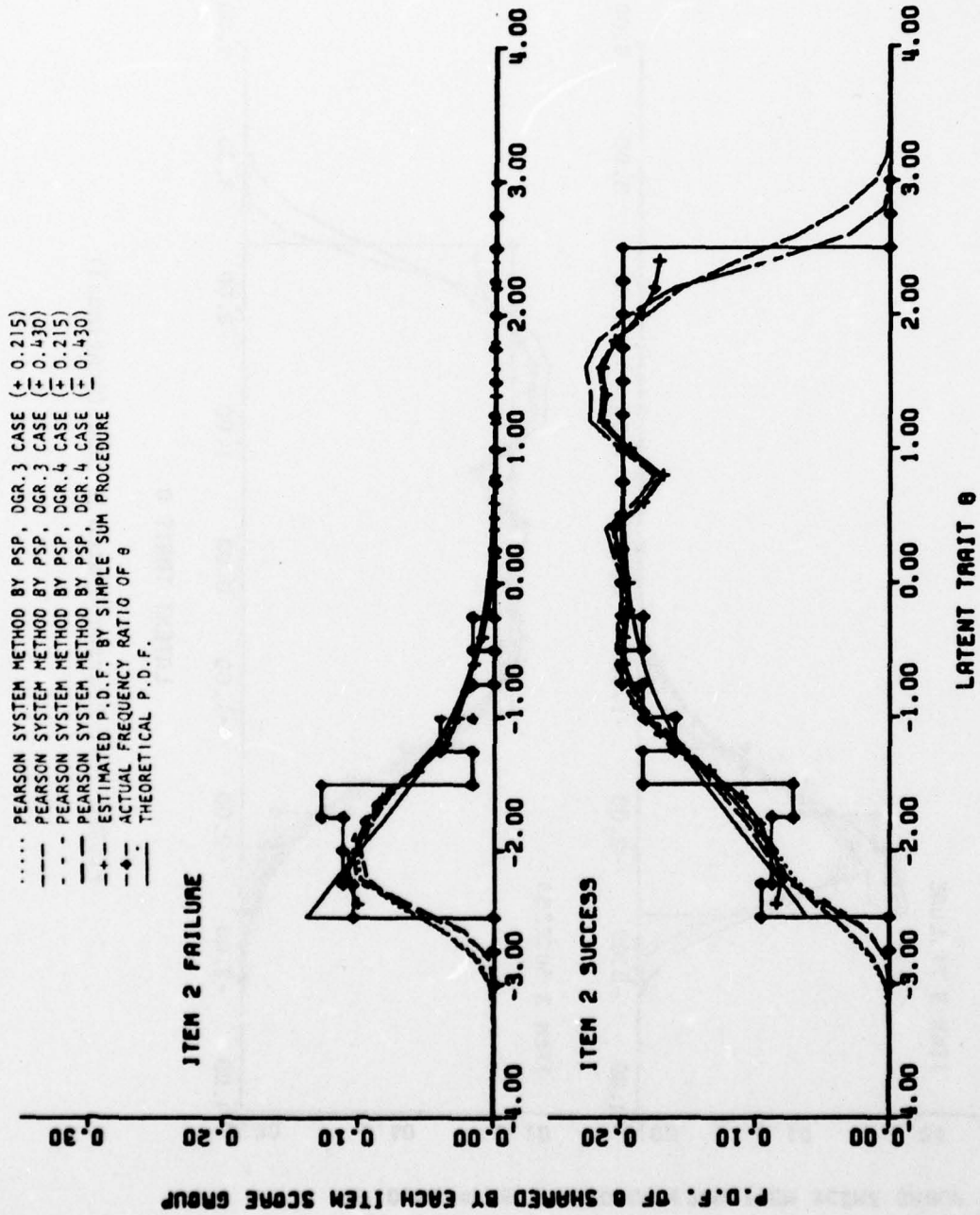


FIGURE 5-4: Pearson System Method (Continued)

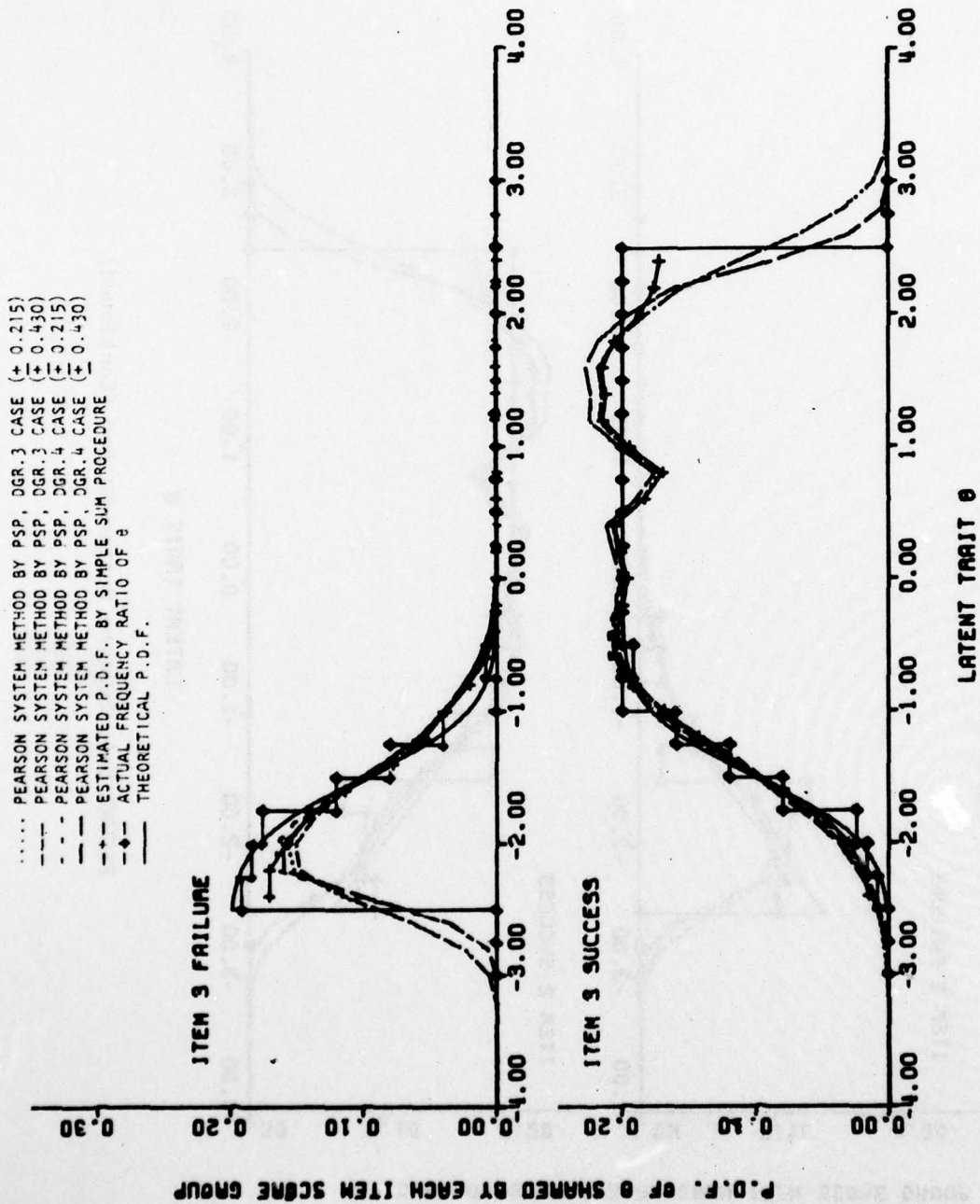


FIGURE 5-4: Pearson System Method (Continued)

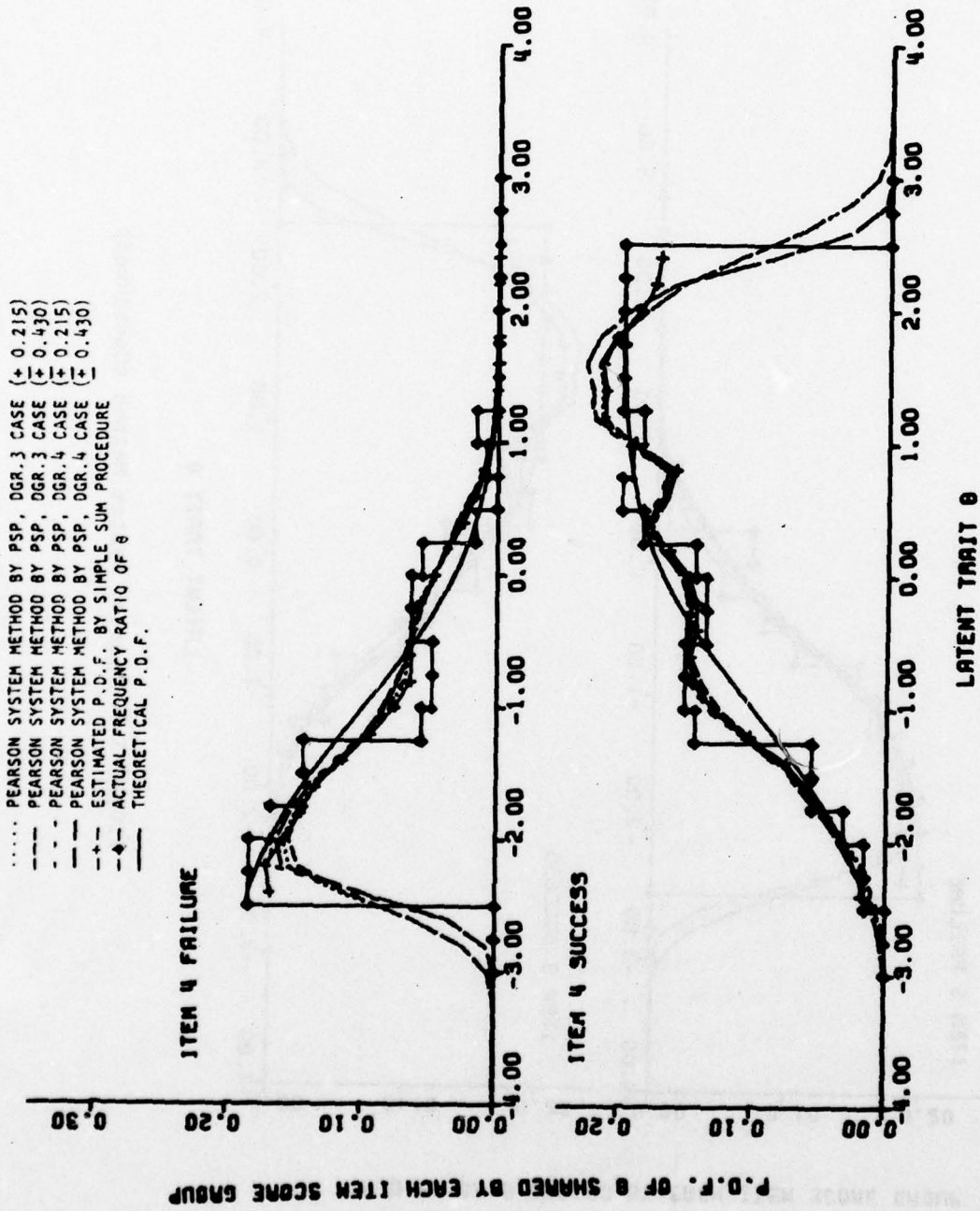


FIGURE 5-4: Pearson System Method (Continued)

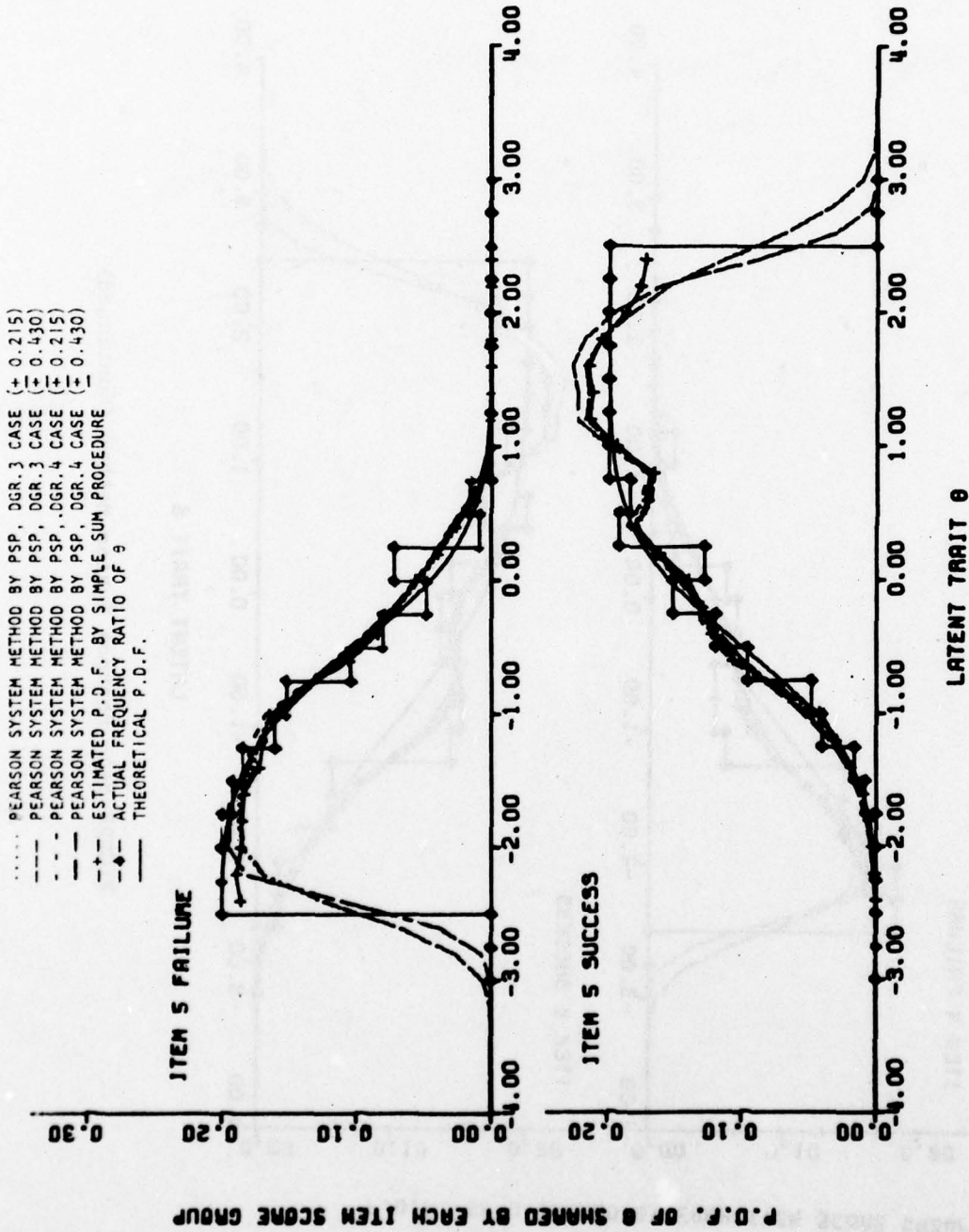


FIGURE 5-4: Pearson System Method (Continued)

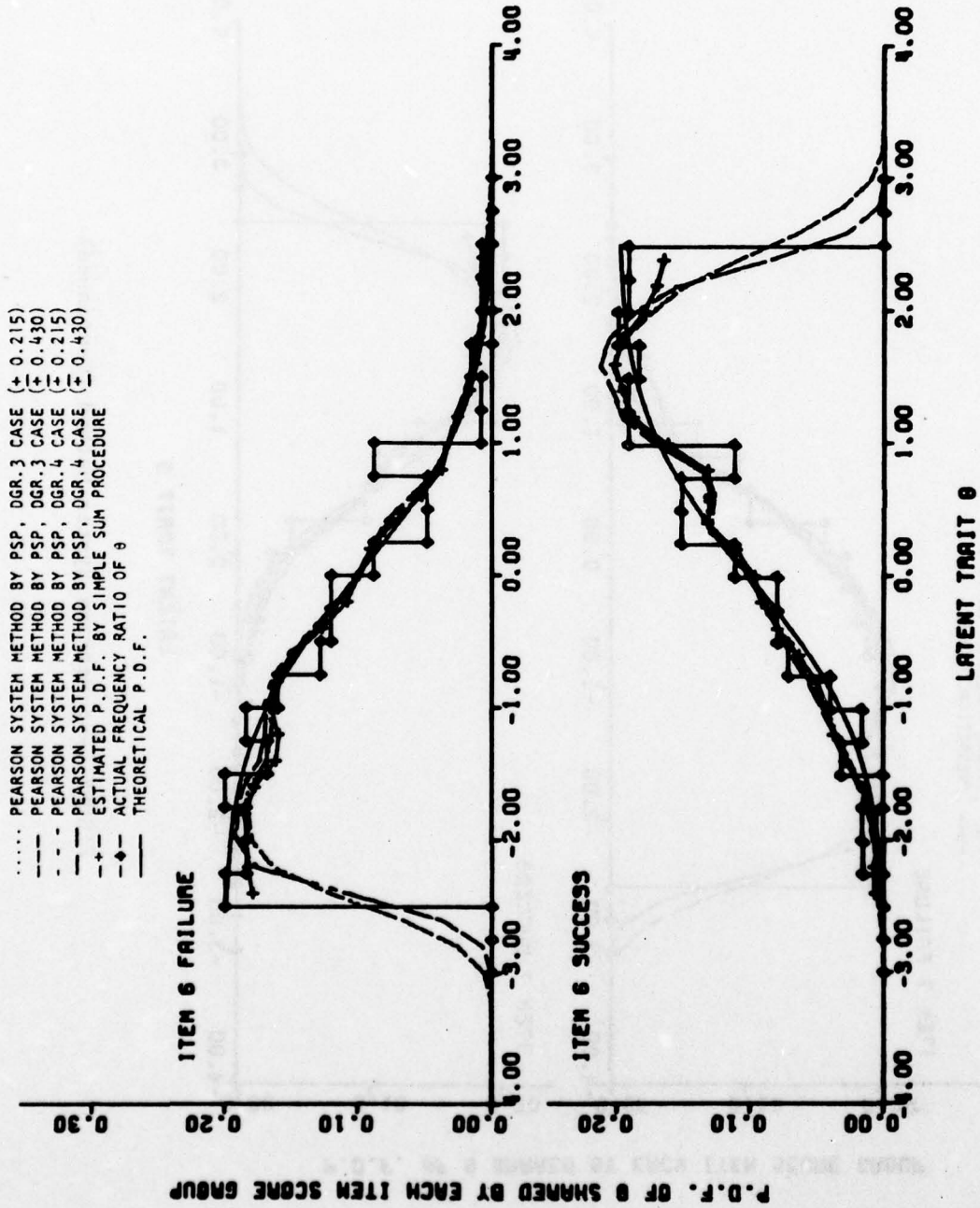


FIGURE 5-4: Pearson System Method (Continued)

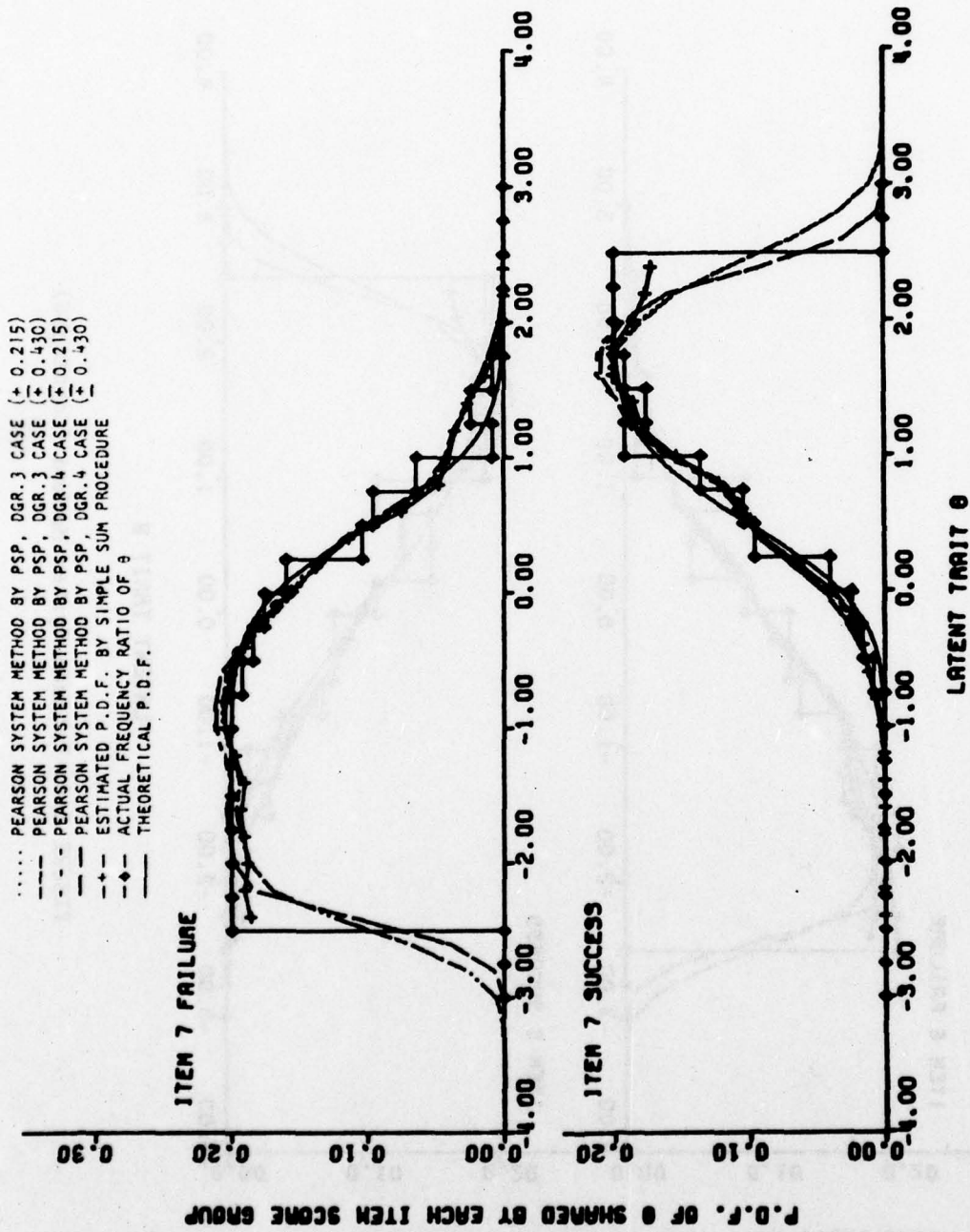


FIGURE 5-4: Pearson System Method (Continued)

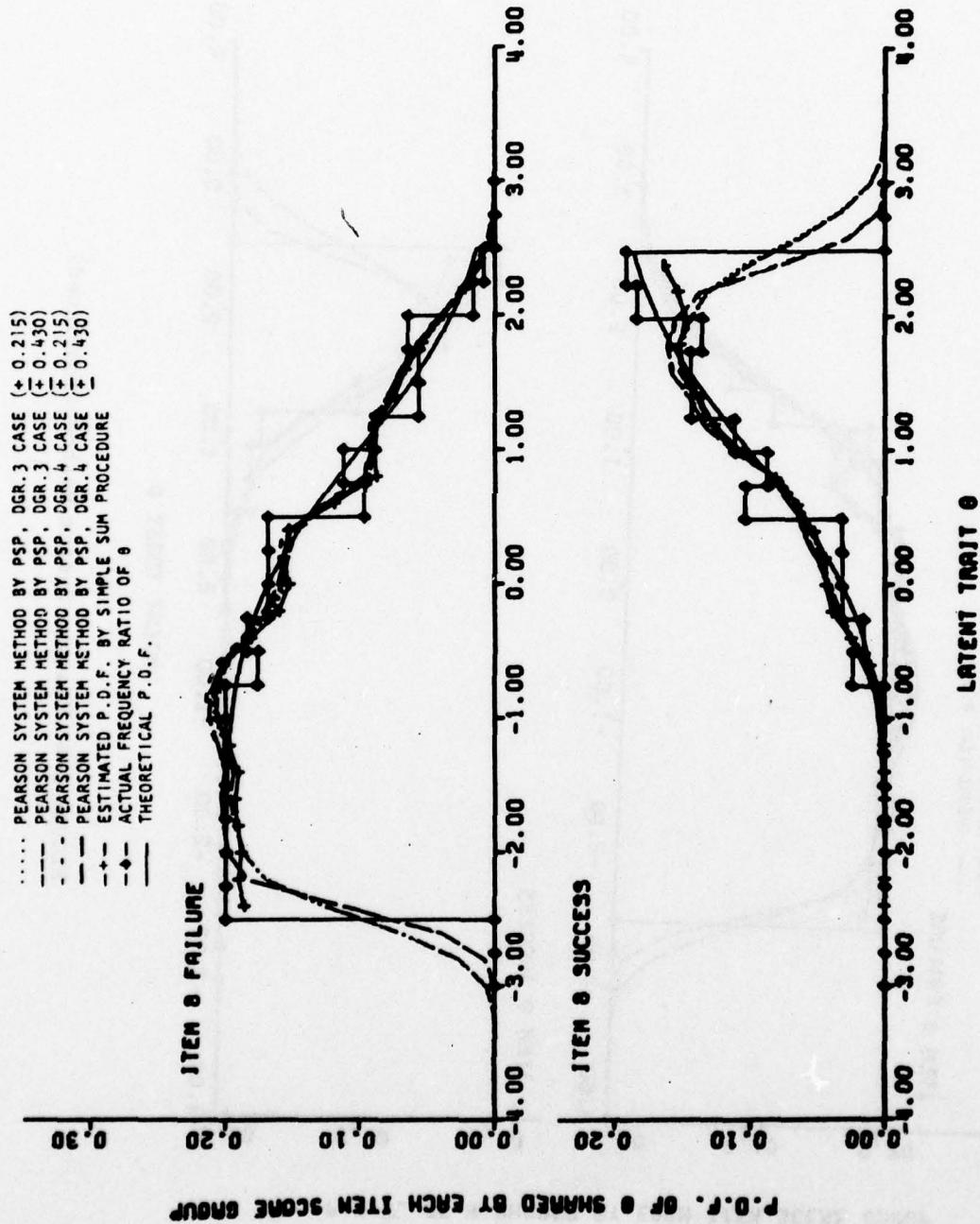


FIGURE 5-4: Pearson System Method (Continued)

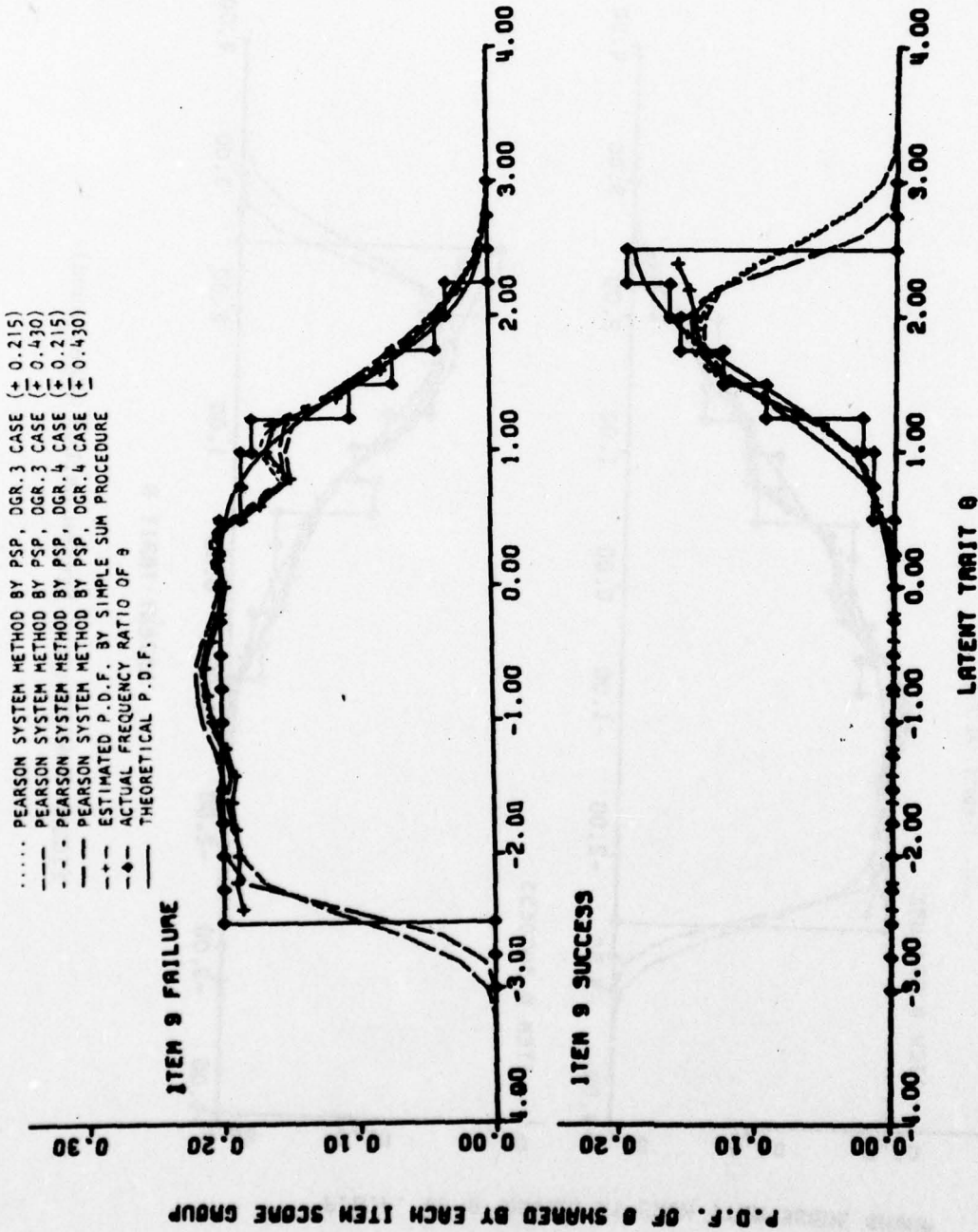


FIGURE 5-4: Pearson System Method (Continued)



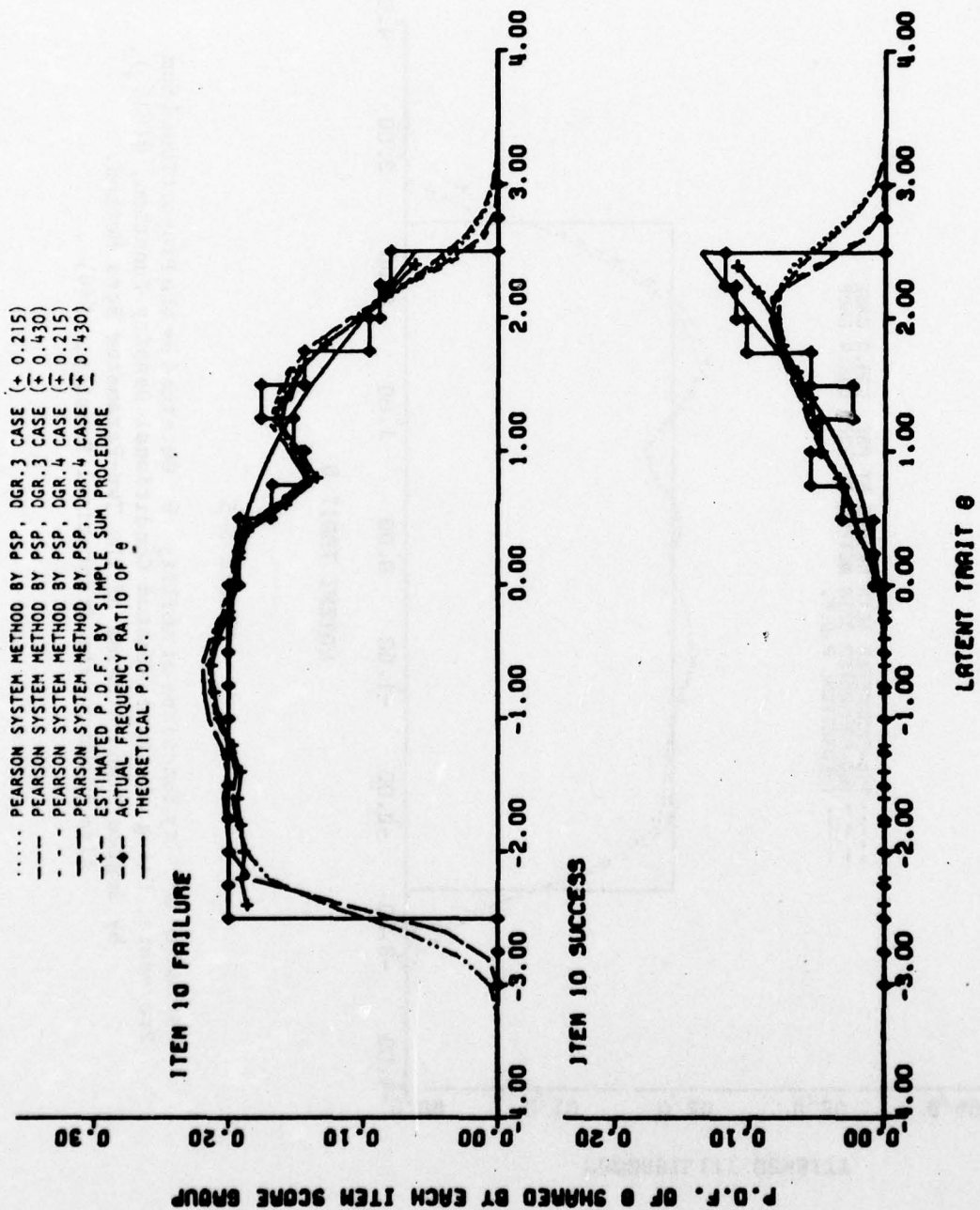


FIGURE 5-4: Pearson System Method (Continued)

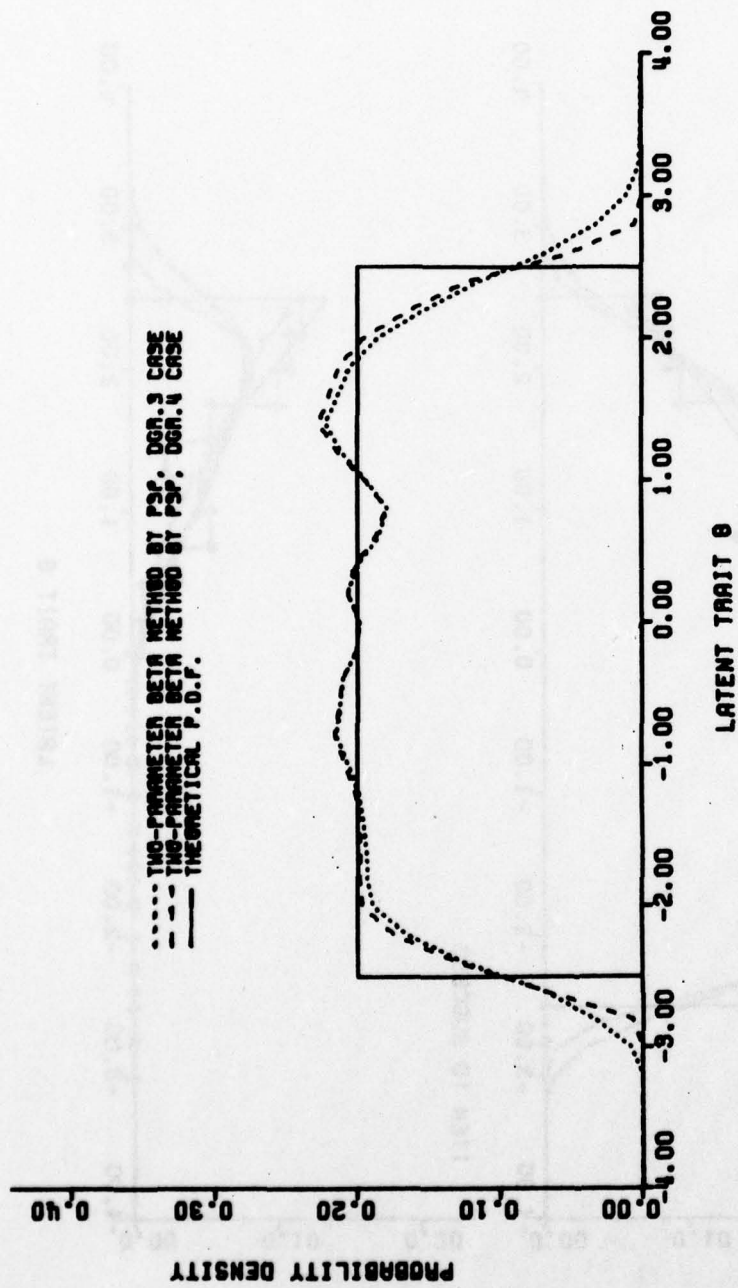


FIGURE 5-5

Estimated Density Functions of Ability  $\theta$  Obtained by the Proportioned Sum Procedure, Using the Approximated Conditional Density Function,  $\hat{\phi}(\theta|\hat{\theta}_s)$ , by Degree 3 and 4 Cases of the Two-Parameter Beta Method.

Also Drawn is the Theoretical Density,  $f(\theta)$ .

$\hat{\theta}_s \pm 0.430$  , Presented in the same figure are those obtained by the criterion case of the Simple Sum Procedure, the actual frequency ratios and the theoretical shared densities. Just as we have observed in the results obtained by the Pearson System Method, in each of Degree 3 and 4 Cases, the two curves for the estimated shared density functions using the two different intervals of  $\hat{\theta}_s$  in obtaining the proportioned weights are very close to each other. It is also noticed that the set of four curves for each item is not too different from the set of those obtained by the Pearson System Method, which are presented in Figure 5-4.

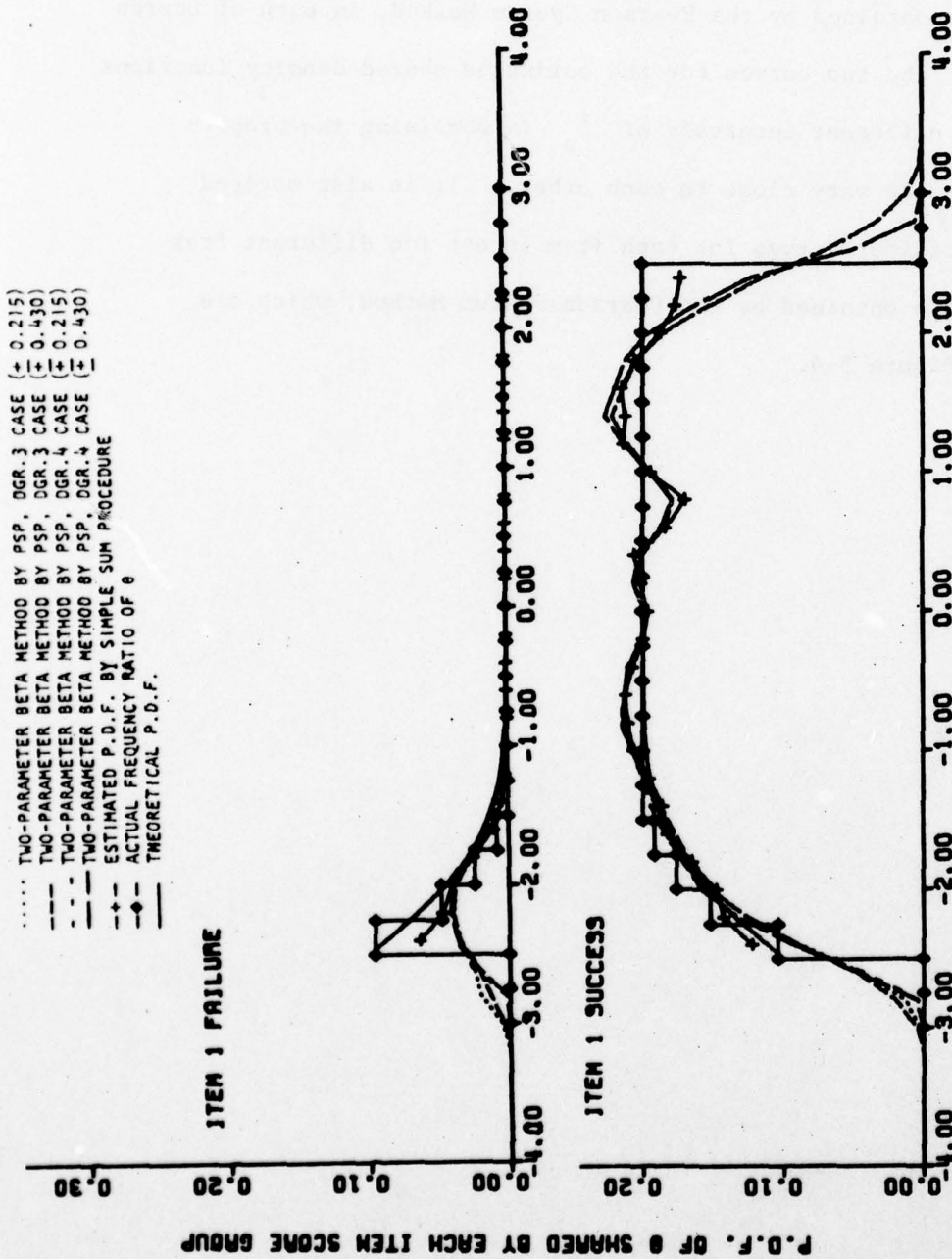


FIGURE 5-6

Four Estimated Shared Density Functions of Ability  $\theta$  of Each of the Two Item Score Groups, Which Are Obtained by the Proportioned Sum Procedure Using the Approximated Conditional Density Function,  $\hat{\phi}(\theta|\hat{\theta}_s)$ , by Degree 3 and 4 Cases of the Two-Parameter Beta Method, Respectively. Also Presented Are the One Obtained by Using the True Conditional Density Function,  $\phi(\theta|\hat{\theta}_s)$ , by the Simple Sum Procedure, Actual Frequency Ratios of Ability  $\theta$ , and Theoretical Shared Density Function.

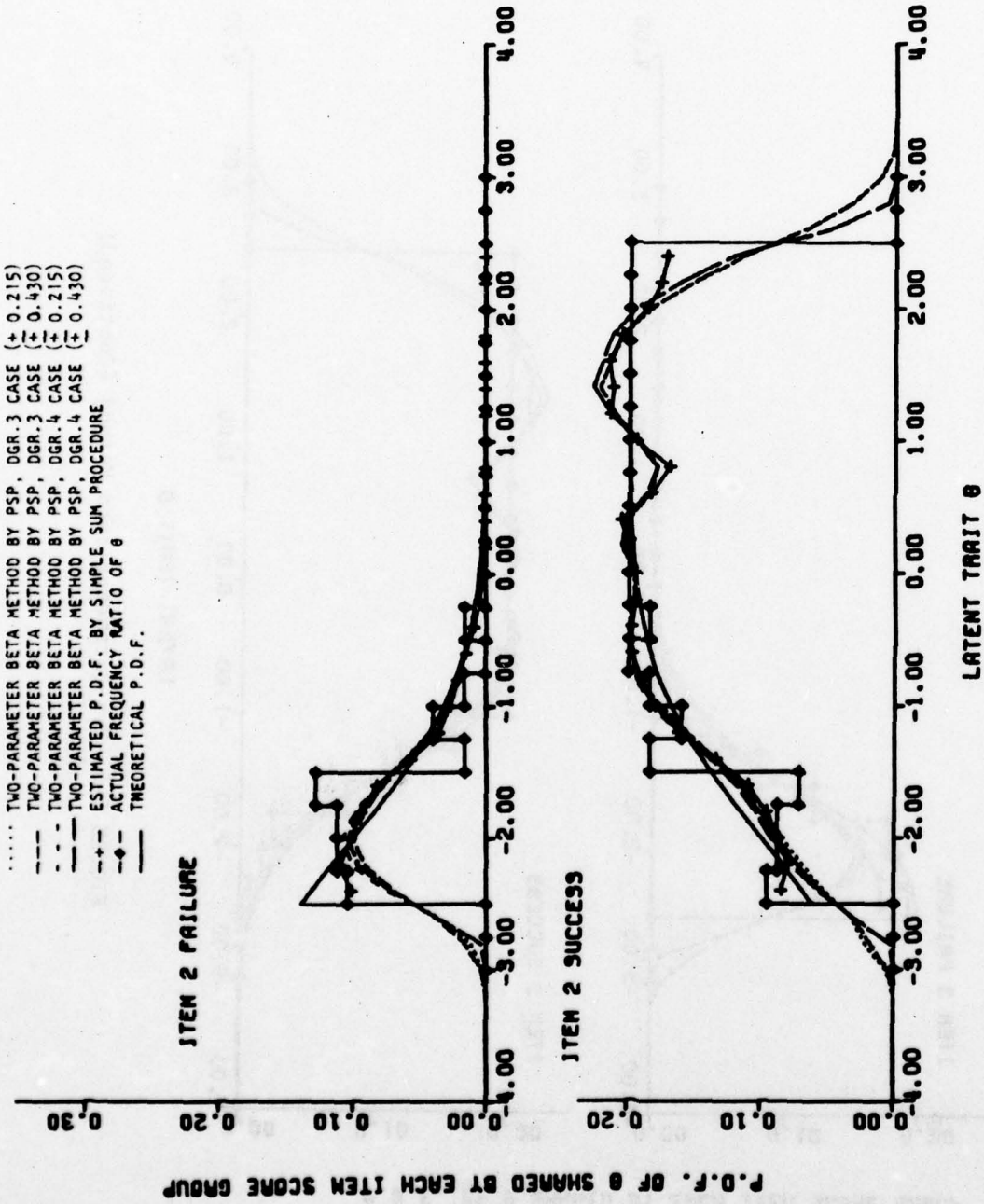


FIGURE 5-6: Two-Parameter Beta Method (Continued)

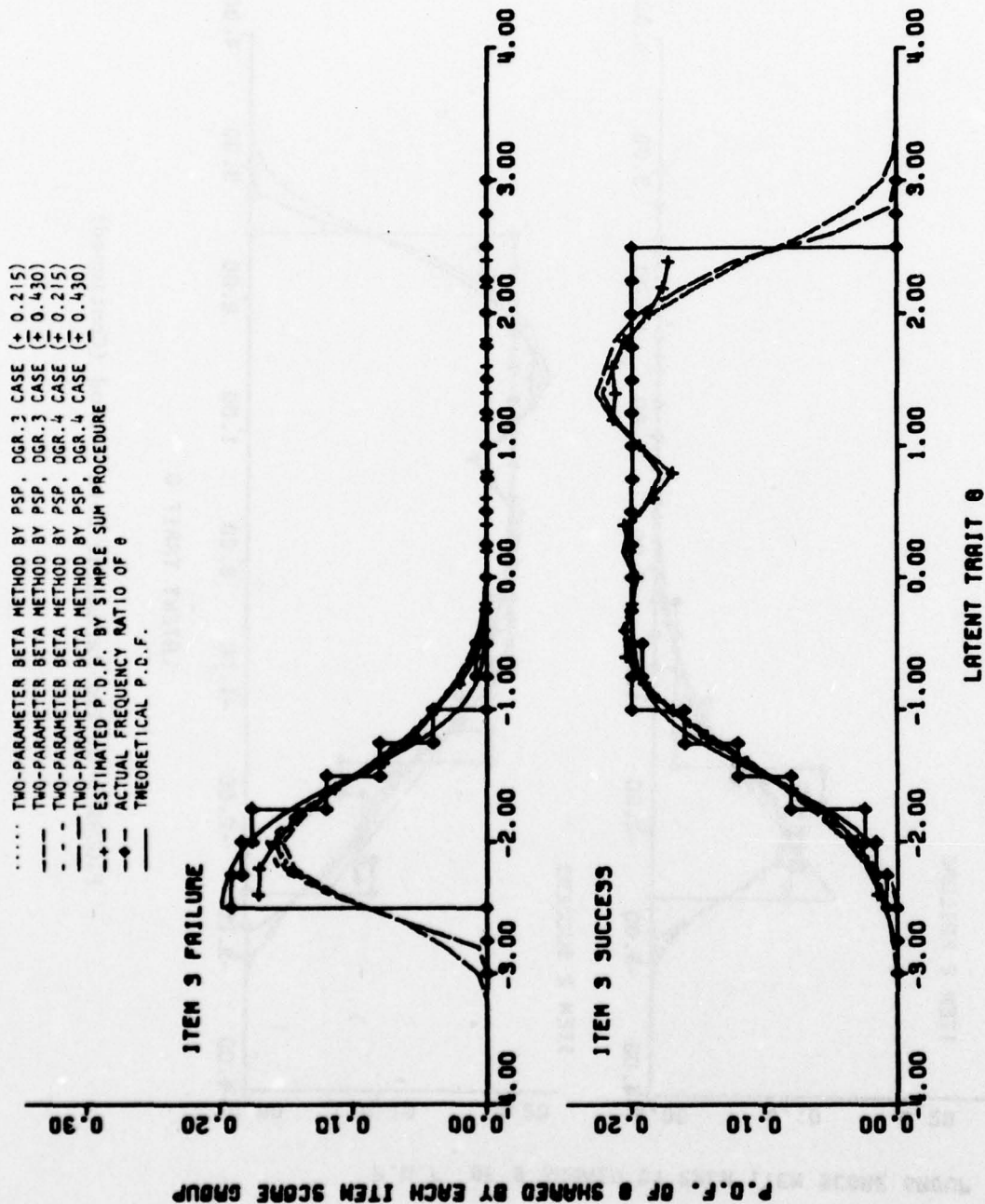


FIGURE 5-6: Two-Parameter Beta Method (Continued)

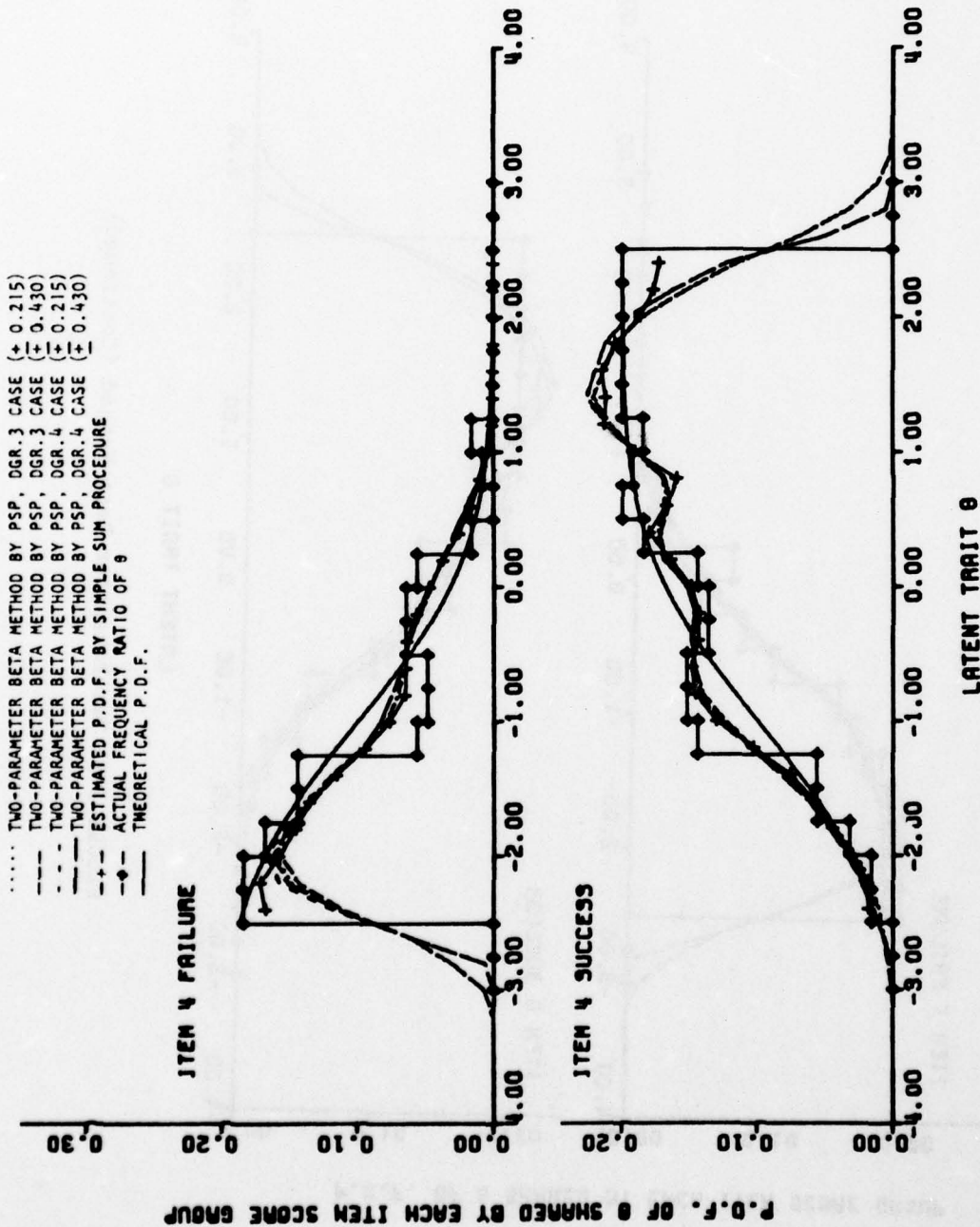


FIGURE 5-6: Two-Parameter Beta Method (Continued)

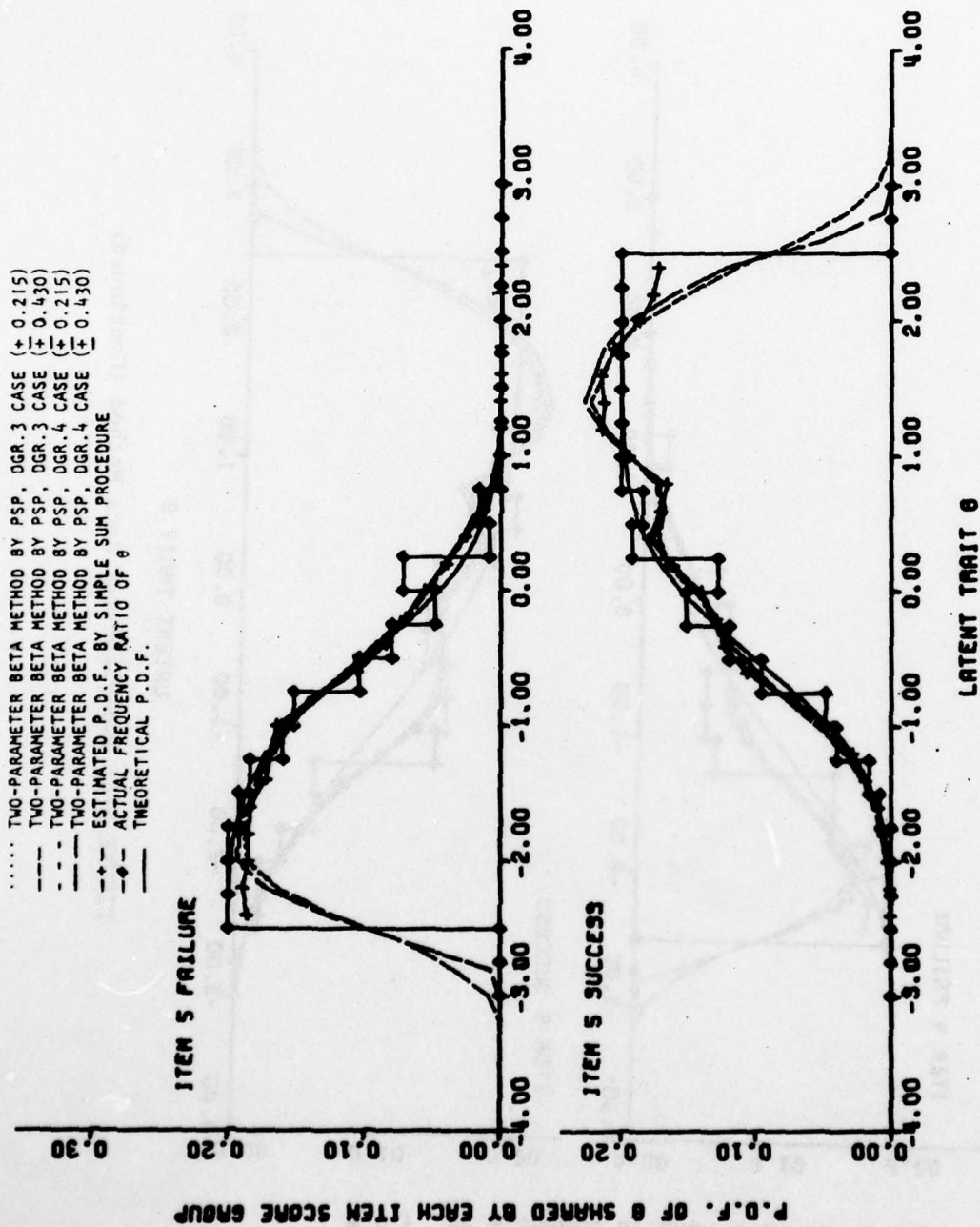


FIGURE 5-6: Two-Parameter Beta Method (Continued)



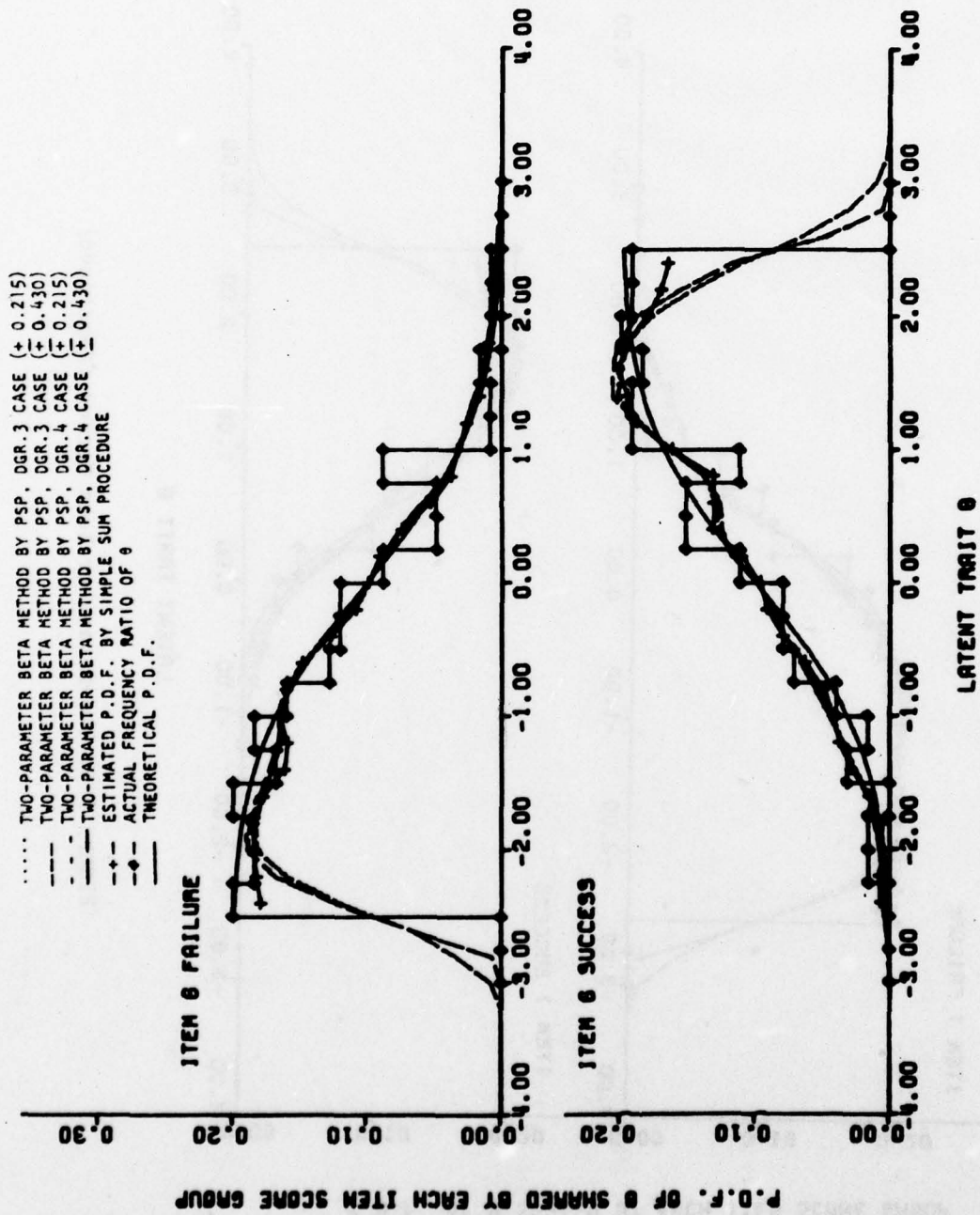


FIGURE 5-6: Two-Parameter Beta Method (Continued)

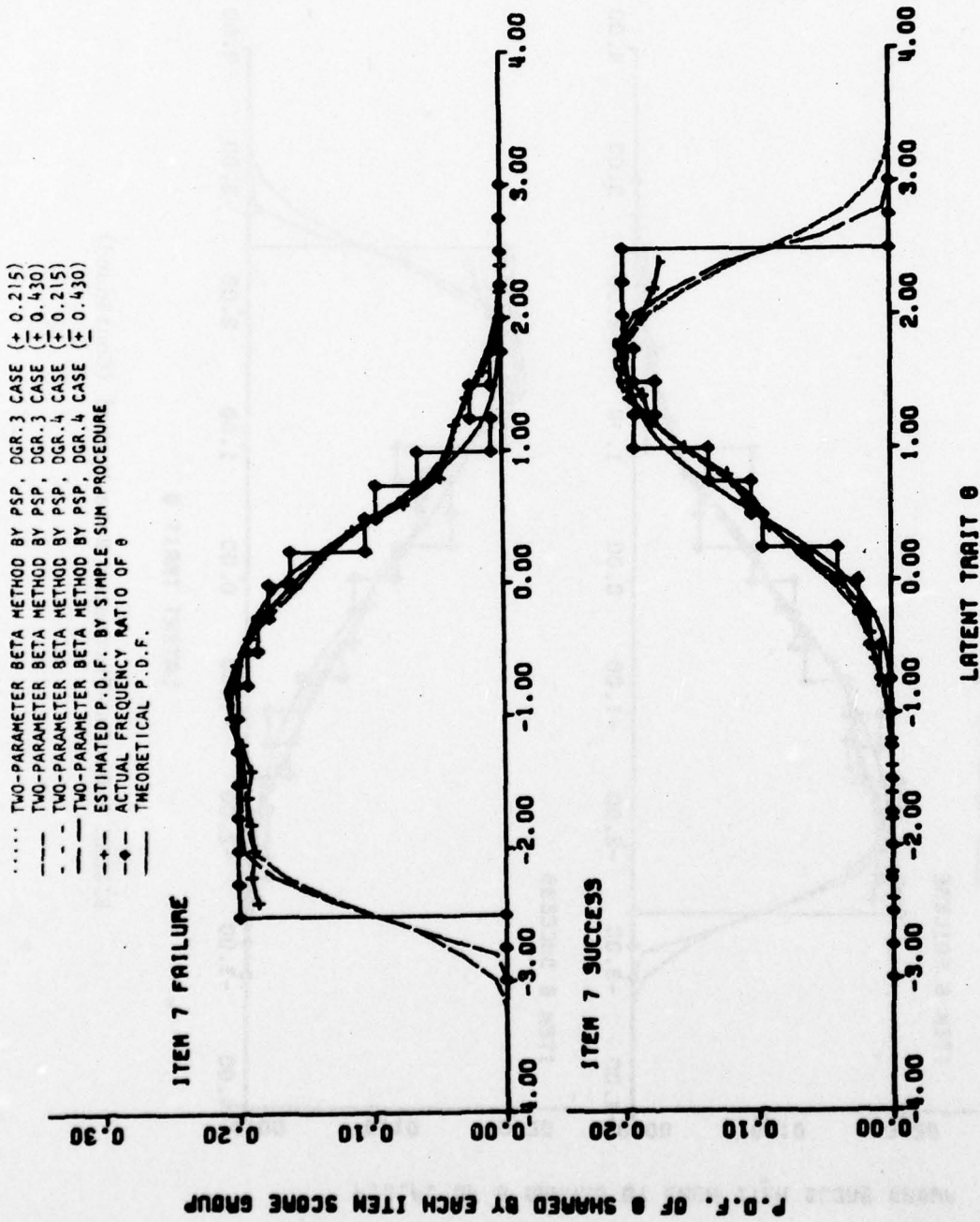


FIGURE 5-6: Two-Parameter Beta Method (Continued)

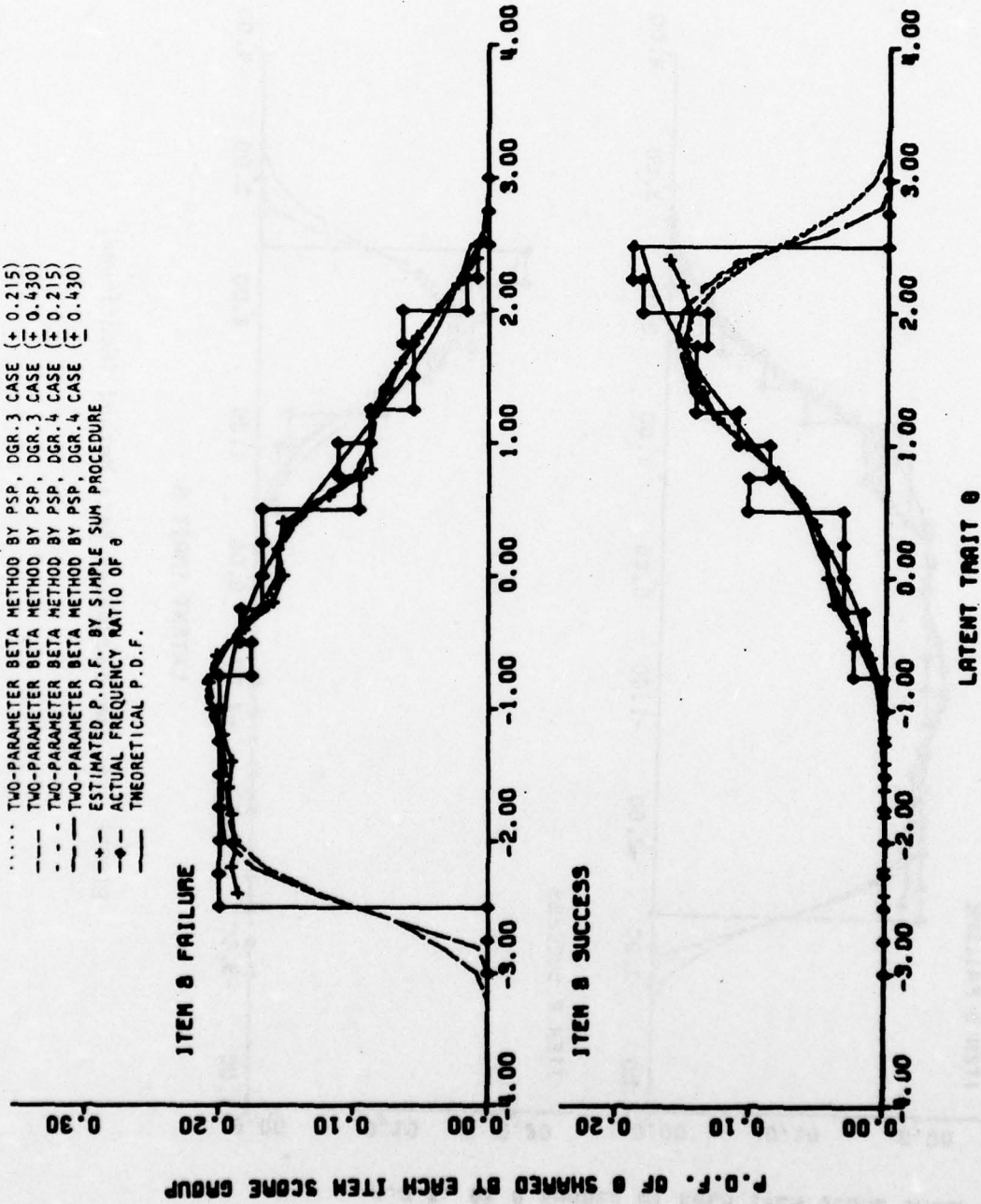


FIGURE 5-6: Two-Parameter Beta Method (Continued)

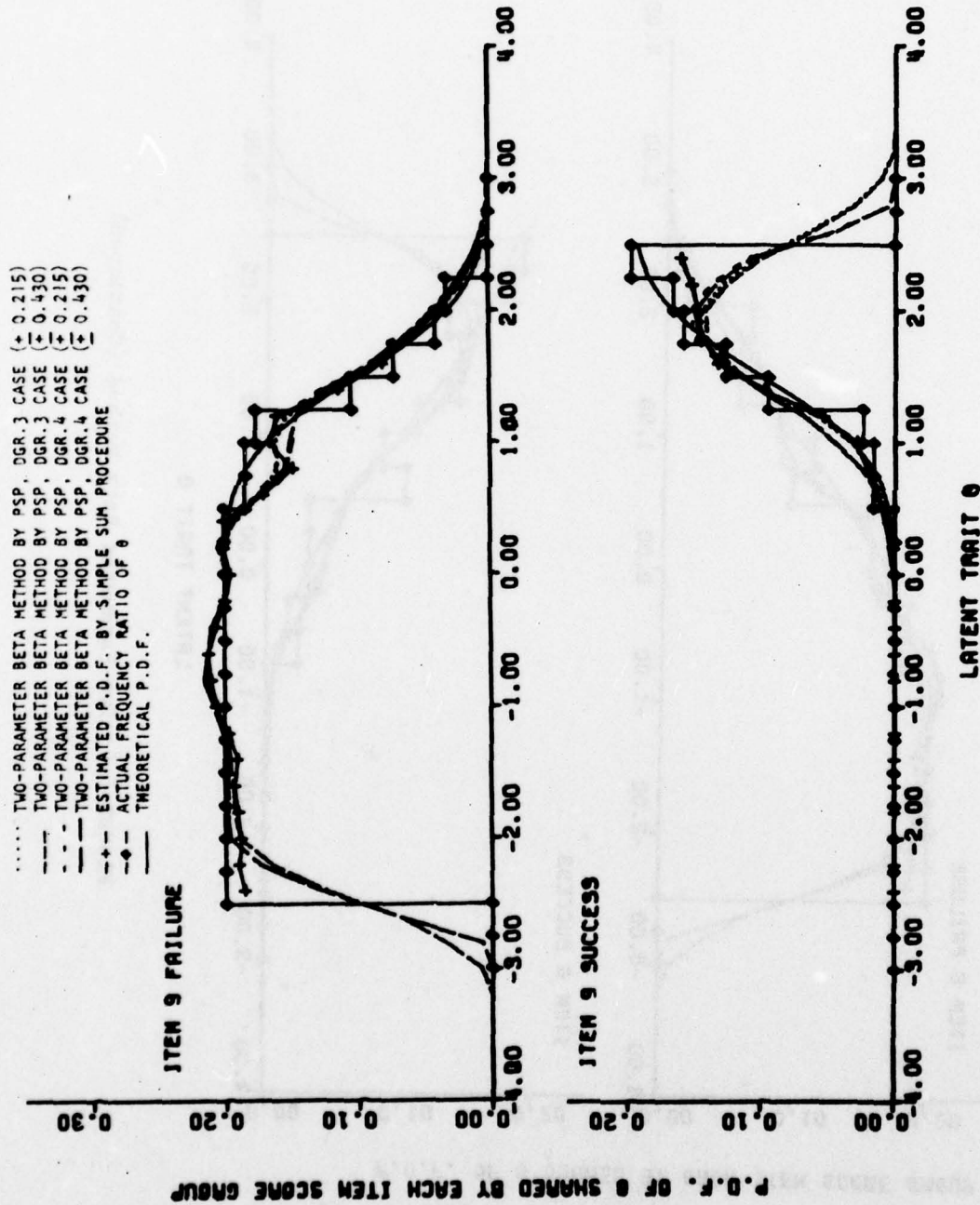


FIGURE 5-6: Two-Parameter Beta Method (Continued)

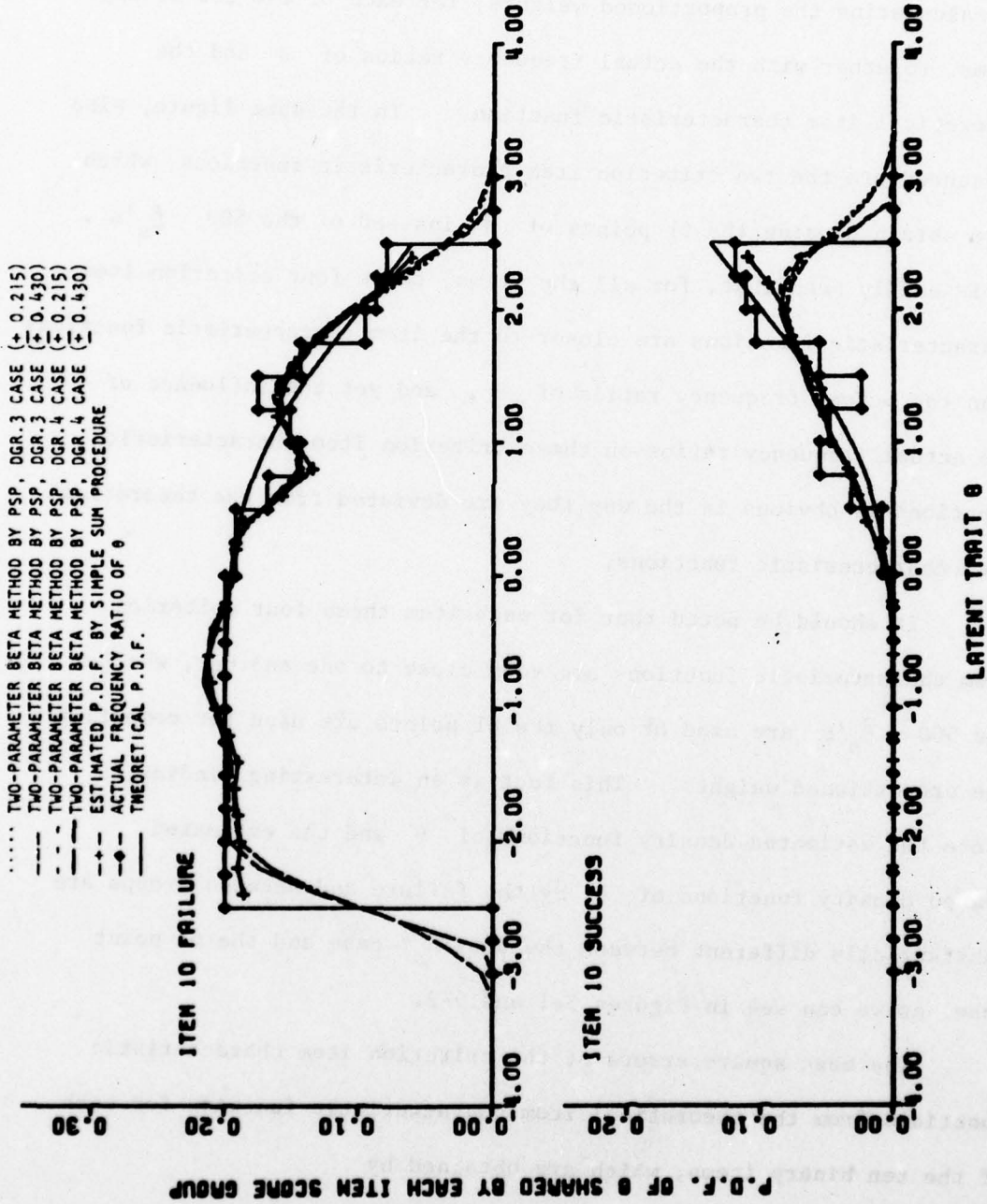


FIGURE 5-6: Two-Parameter Beta Method (Continued)

VI Results III: Estimated Item Characteristic Functions

Figure 6-1 presents the resultant criterion item characteristic functions using two different intervals,  $\hat{\theta}_s \pm 0.215$  and  $\hat{\theta}_s \pm 0.430$ , in calculating the proportioned weights, for each of the ten binary items, together with the actual frequency ratios of  $\theta$  and the theoretical item characteristic function. In the same figure, also presented are the two criterion item characteristic functions, which were obtained using the 61 points of  $\hat{\theta}$  instead of the 500  $\hat{\theta}_s$ 's. It is easily seen that, for all the items, these four criterion item characteristic functions are closer to the item characteristic functions than the actual frequency ratios of  $\theta$ , and yet the influence of the actual frequency ratios on these criterion item characteristic functions is obvious in the way they are deviated from the theoretical item characteristic functions.

It should be noted that for each item these four criterion item characteristic functions are very close to one another, whether the 500  $\hat{\theta}_s$ 's are used or only the 61 points are used for computing the proportioned weights. This fact is an interesting finding, since the estimated density functions of  $\theta$  and the estimated shared density functions of  $\theta$  by the failure and success groups are substantially different between the 500  $\hat{\theta}_s$  case and the 61 point case, as we can see in Figures 5-1 and 5-2.

The mean square errors of the criterion item characteristic functions from the theoretical item characteristic function for each of the ten binary items, which are obtained by

$$(6.1) \quad \frac{1}{m} \sum_{j=1}^m [\hat{P}_g(\theta_j) - P_g(\theta_j)]^2 ,$$

where  $m = 25$  and  $\hat{\theta}_j$ 's are -2.4 through 2.4 with the step of 0.2, were computed for each criterion item characteristic function of each item, and these values and their square roots are presented in Tables 6-1 and 6-2.

It is noticed that among these four sets of mean square errors, or their square roots, there are no substantial differences, which indicates that these four sets of criterion item characteristic functions are just as good as one another. There is even a slight tendency that the two sets for the 61 point case have smaller values than those for the 500  $\hat{\theta}_s$ 's case, which encourages the idea of reducing the number of points of the maximum likelihood estimate for which the conditional density,  $\phi(\theta|\hat{\theta})$ , is estimated and used for the estimation of the operating characteristics.

The discrimination parameter,  $a_g$ , and the difficulty parameter,  $b_g$ , in the normal ogive model\* were estimated by the usual least square method for each of the ten binary items, and are presented as Tables 6-3 and 6-4. It is obvious that these estimated parameters are just as good as those obtained by the Simple Sum Procedure, which are also presented in these tables, if not substantially better. This indicates that they are also comparable with those obtained from the Pseudo Criterion Item Characteristic Functions of Degree 3, 4 and 5 Cases, as we can see in one of the previous studies (Samejima, 1978d). As before, in the present results of the Proportioned Sum Procedure, in general, the estimates of the difficulty parameters  $b_g$  are closer to the true values than those of the discrimination parameters  $a_g$ .

---

\* Refer to Appendix I.

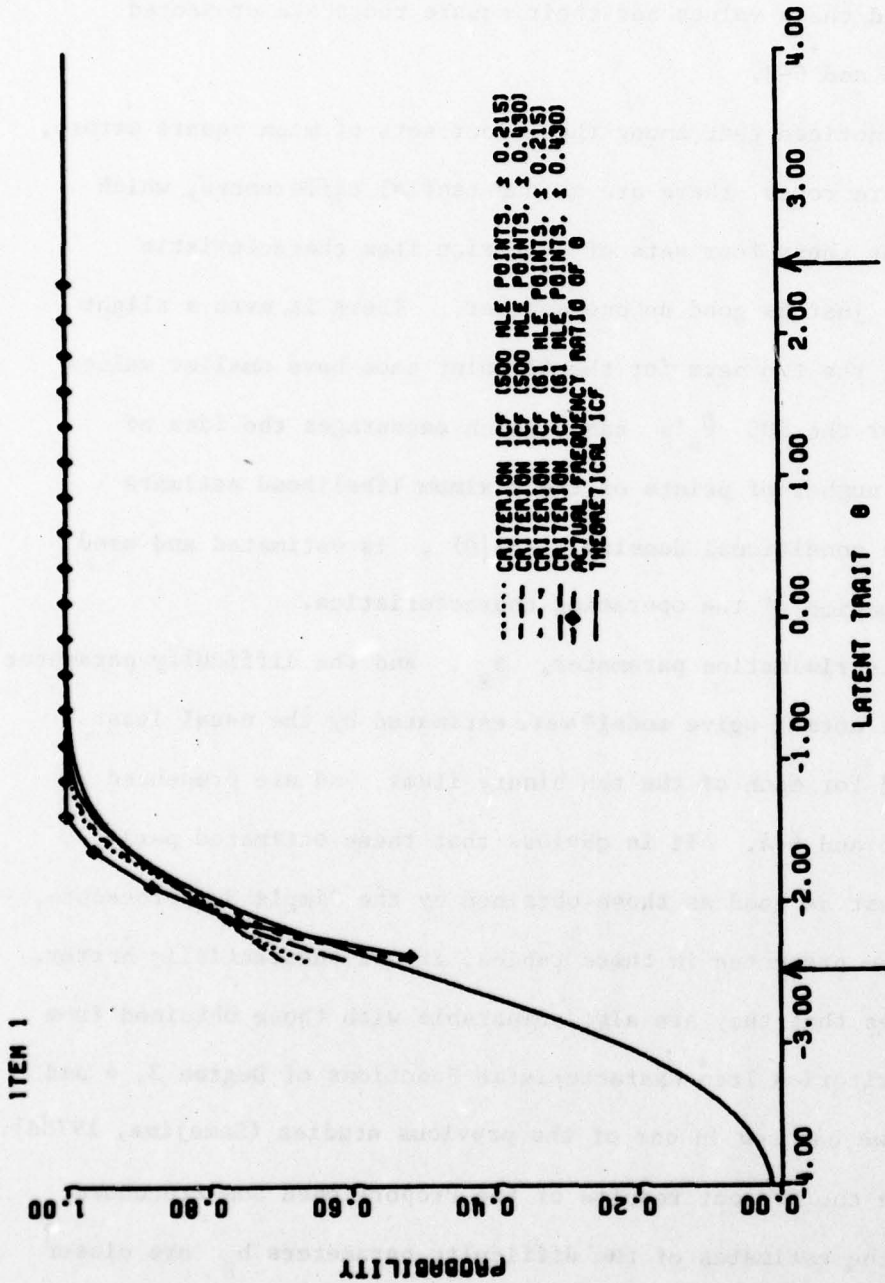


FIGURE 6-1

Comparison of the Four Criterion Item Characteristic Functions by the Proportioned Sum Procedure of Conditional P.D.F. Approach with the Actual Frequency Ratios and Theoretical Item Characteristic Function.



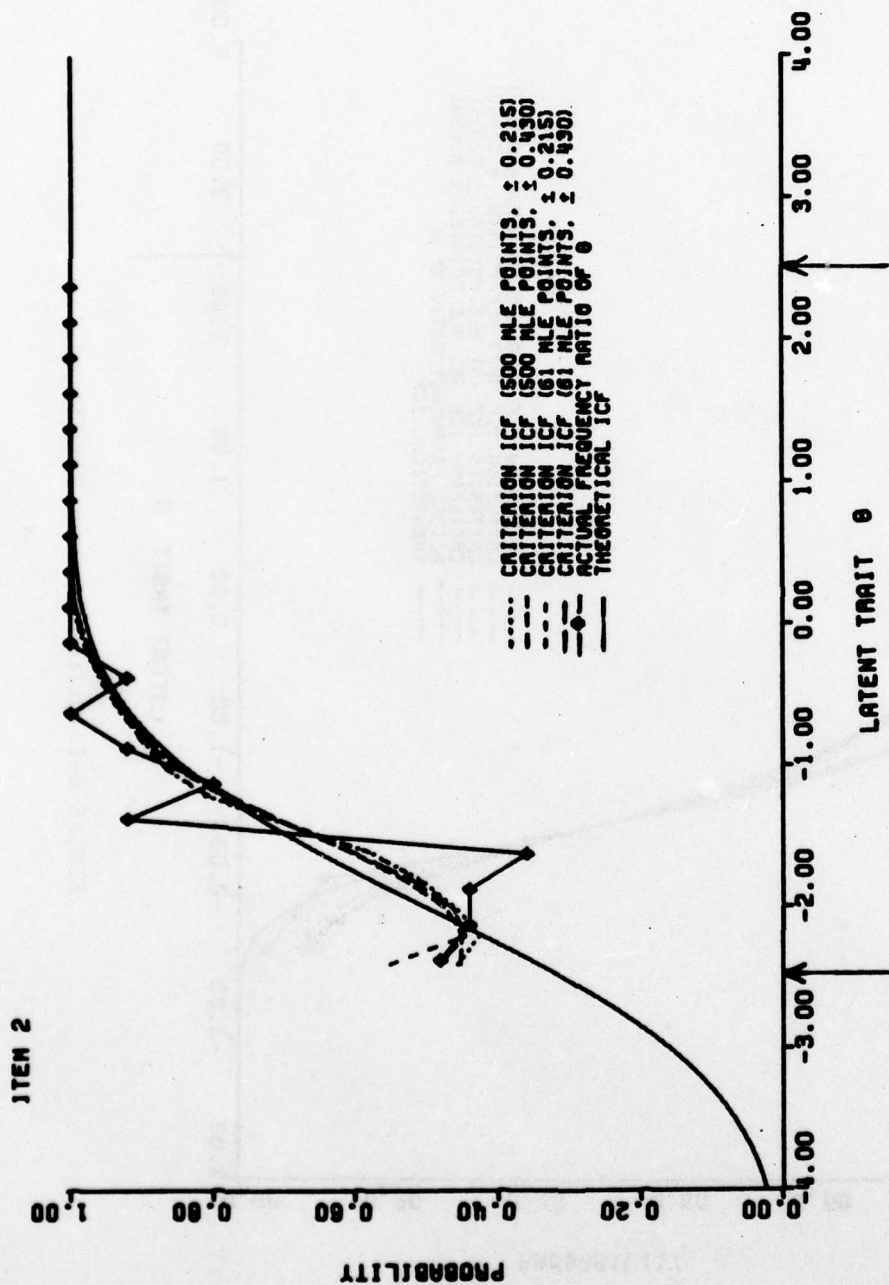


FIGURE 6-1: Criterion Case (Continued)

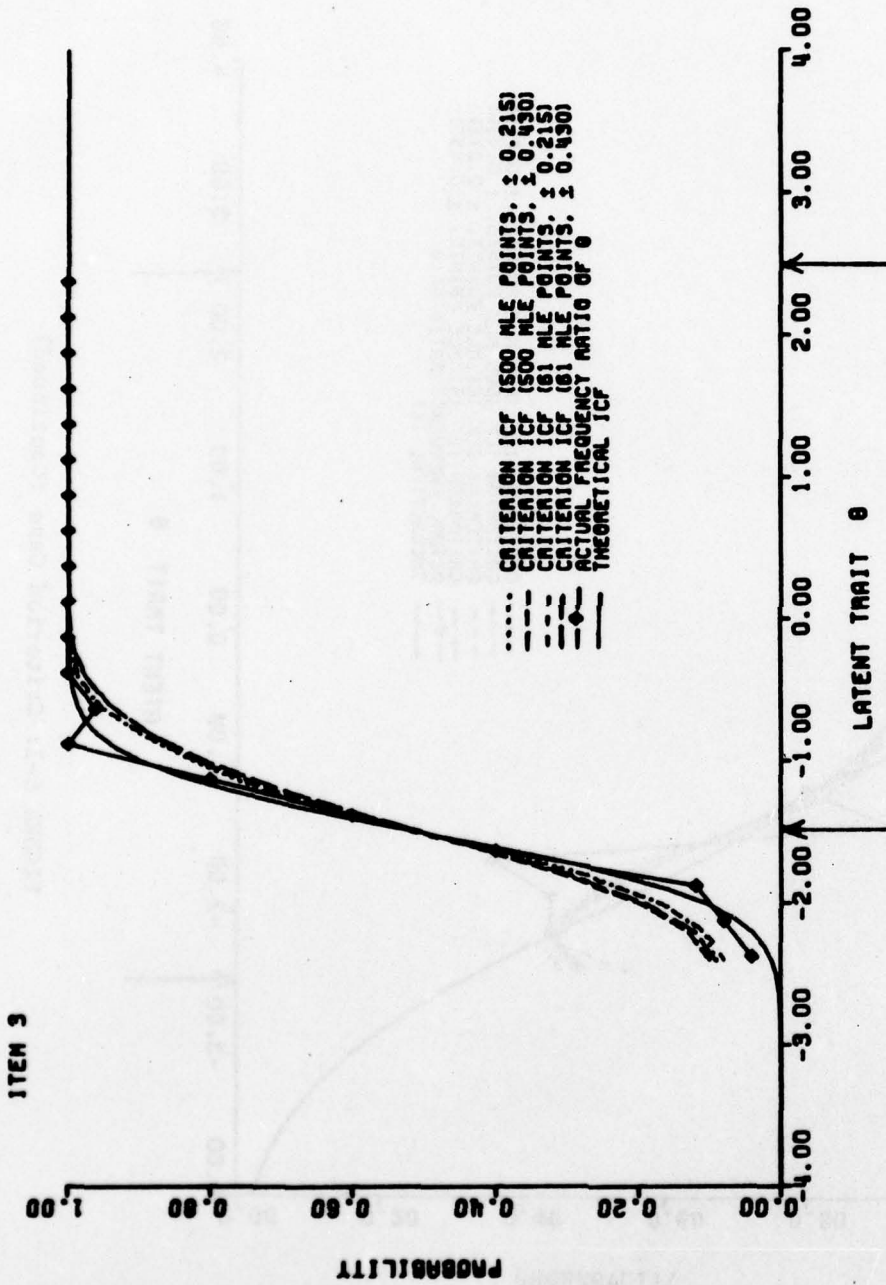


FIGURE 6-1: Criterion Case (Continued)

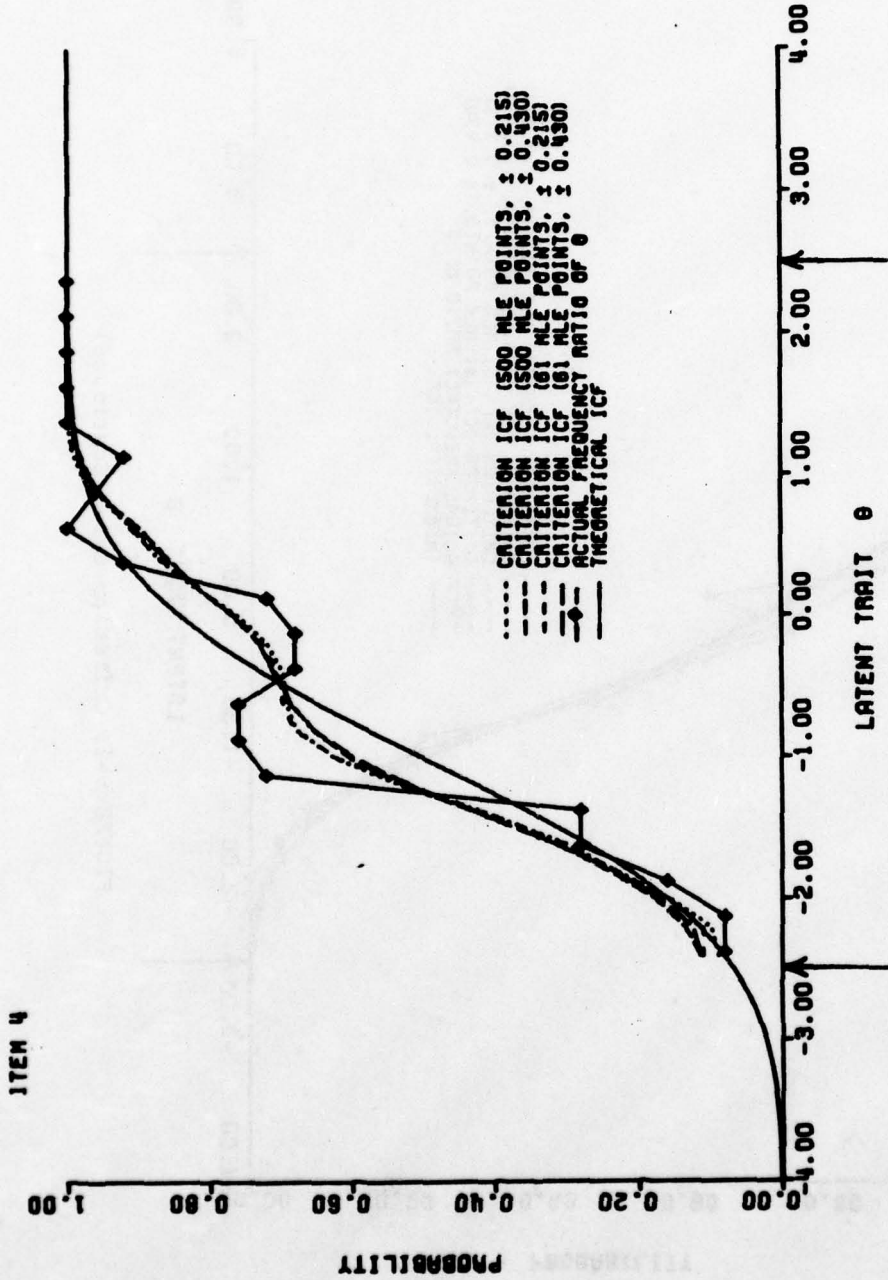


FIGURE 6-1: Criterion Case (Continued)

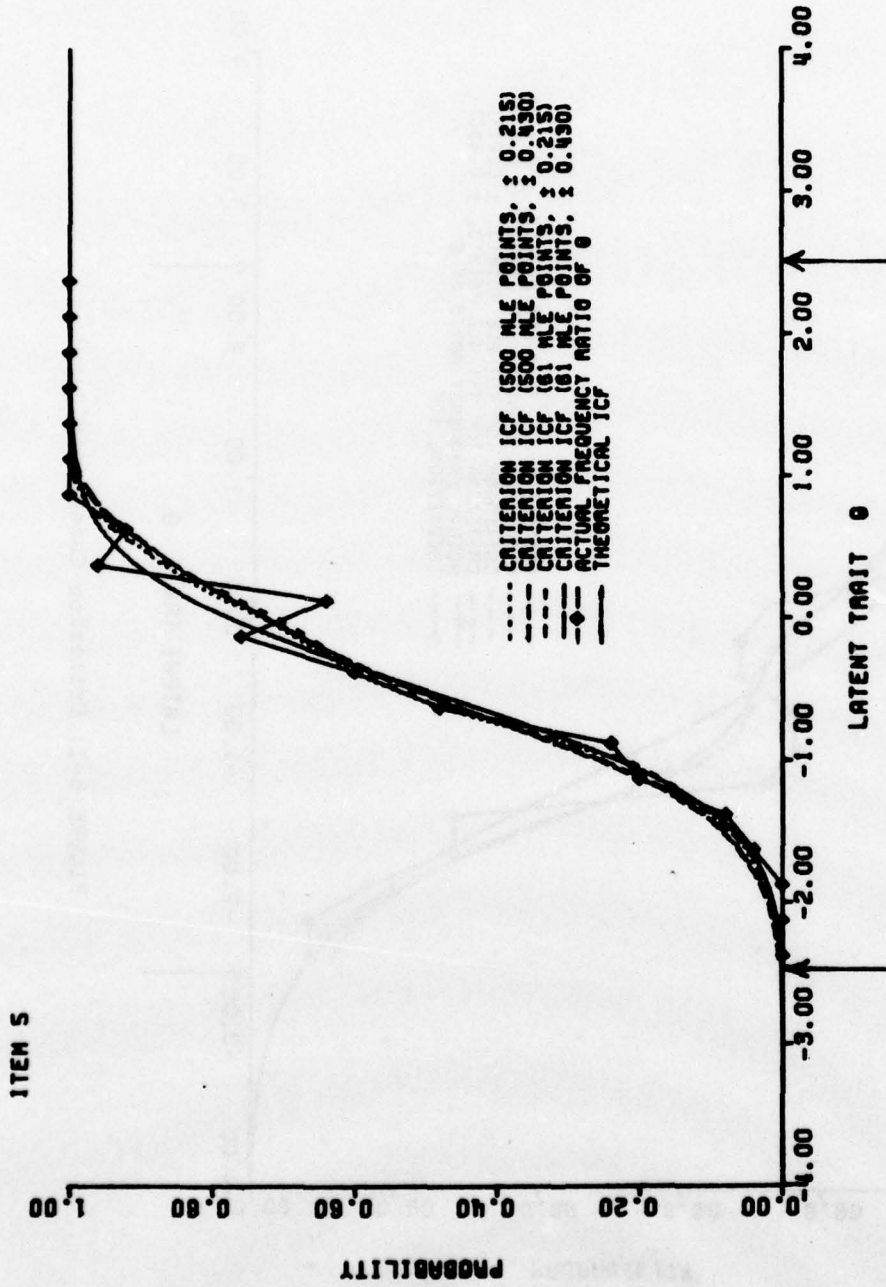


FIGURE 6-1: Criterion Case (Continued)

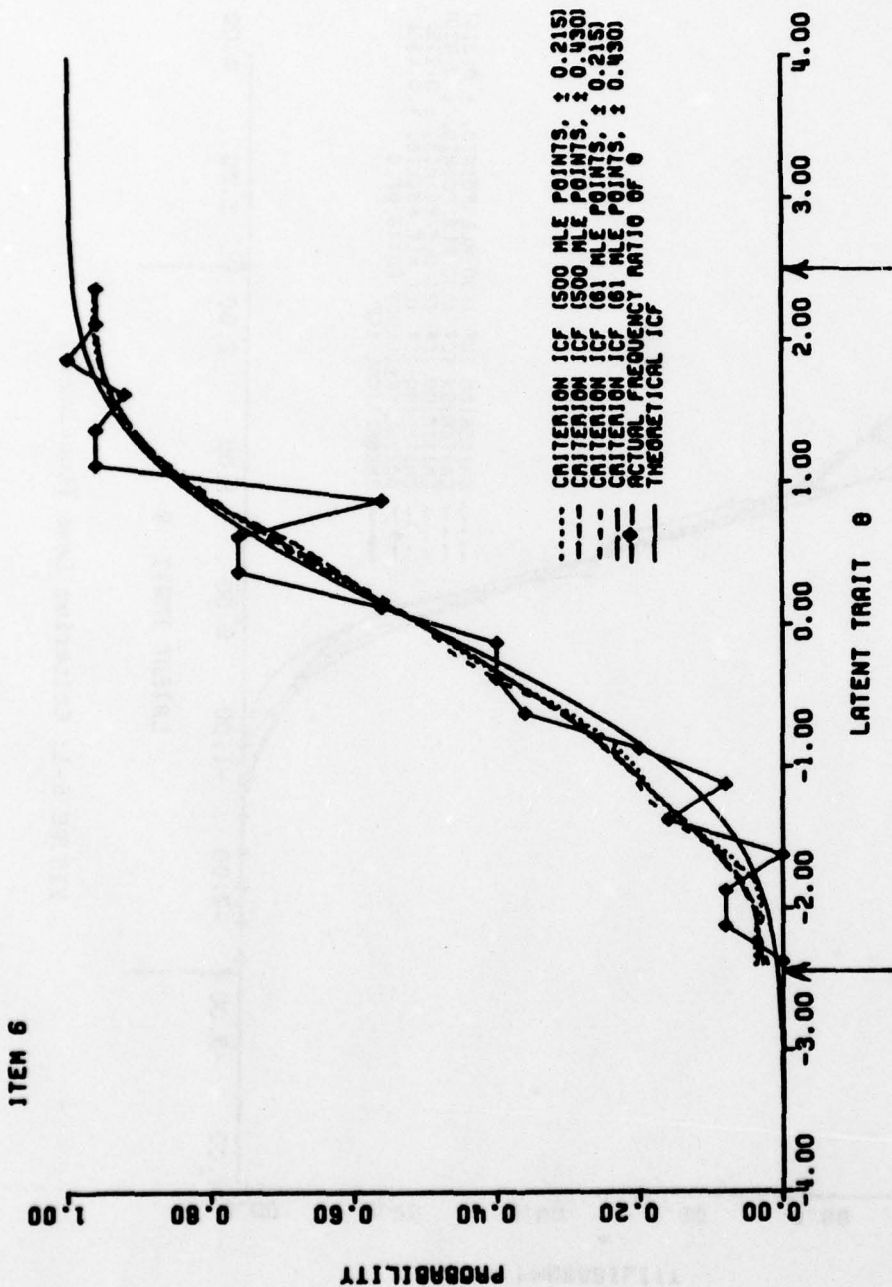


FIGURE 6-1: Criterion Case (Continued)

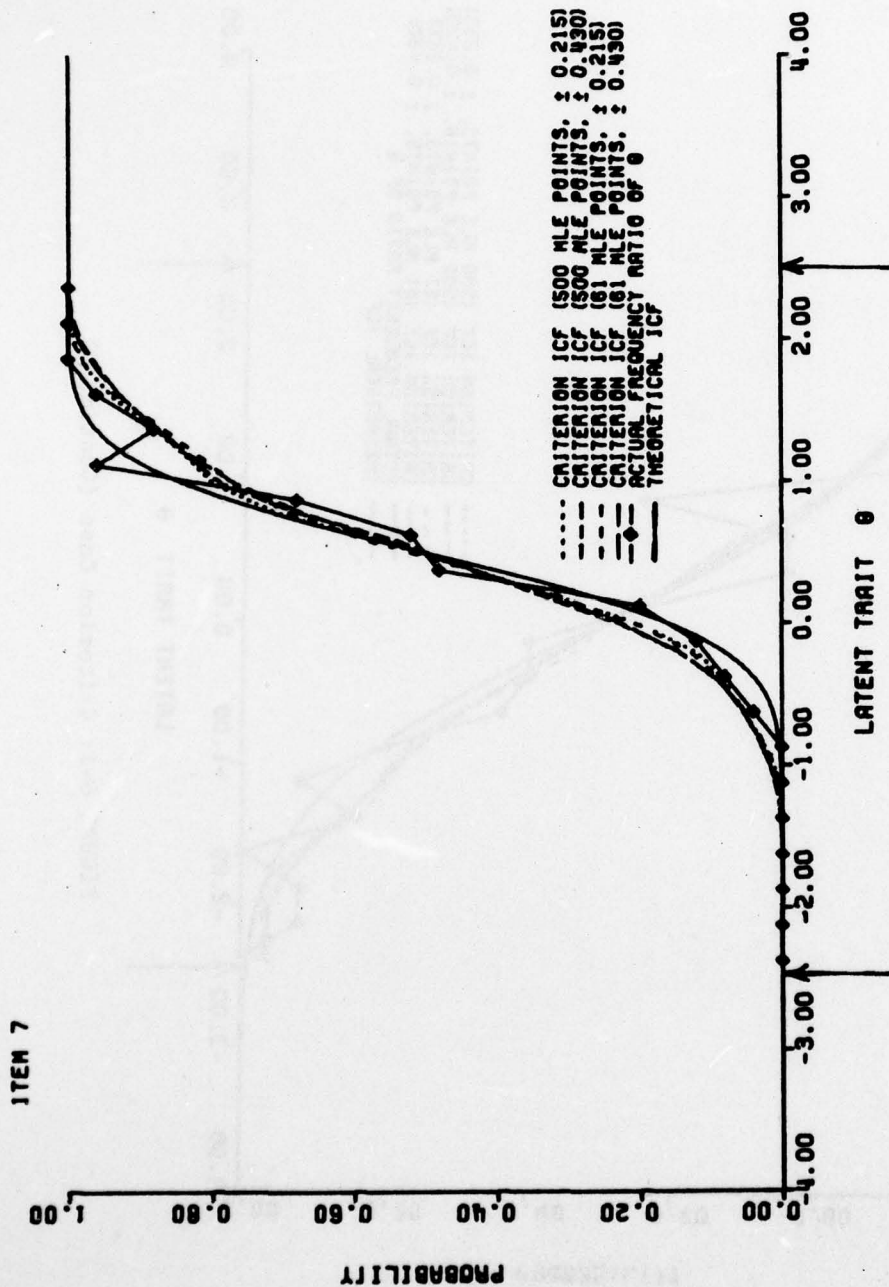


FIGURE 6-1: Criterion Case (Continued)

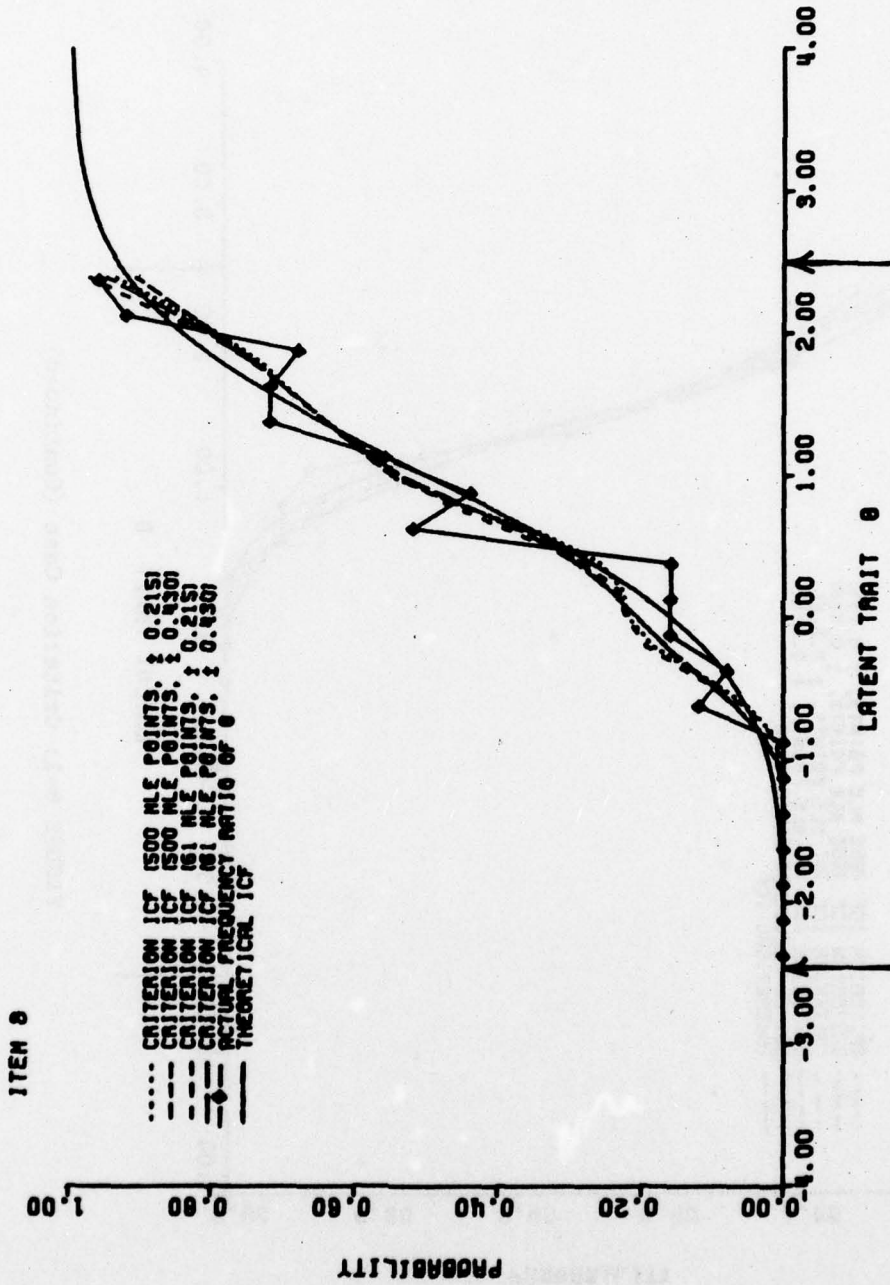


FIGURE 6-1: Criterion Case (Continued)

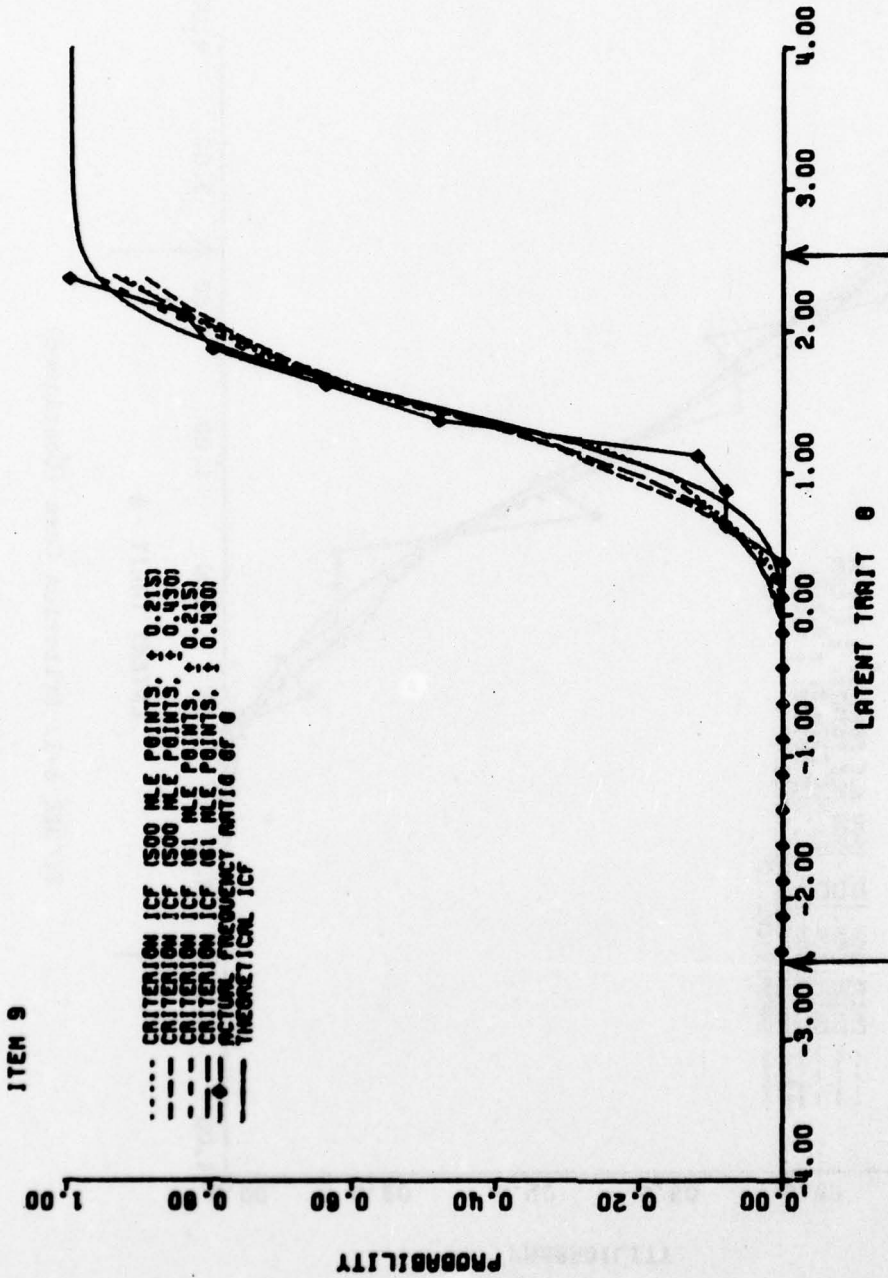


FIGURE 6-1: Criterion Case (Continued)



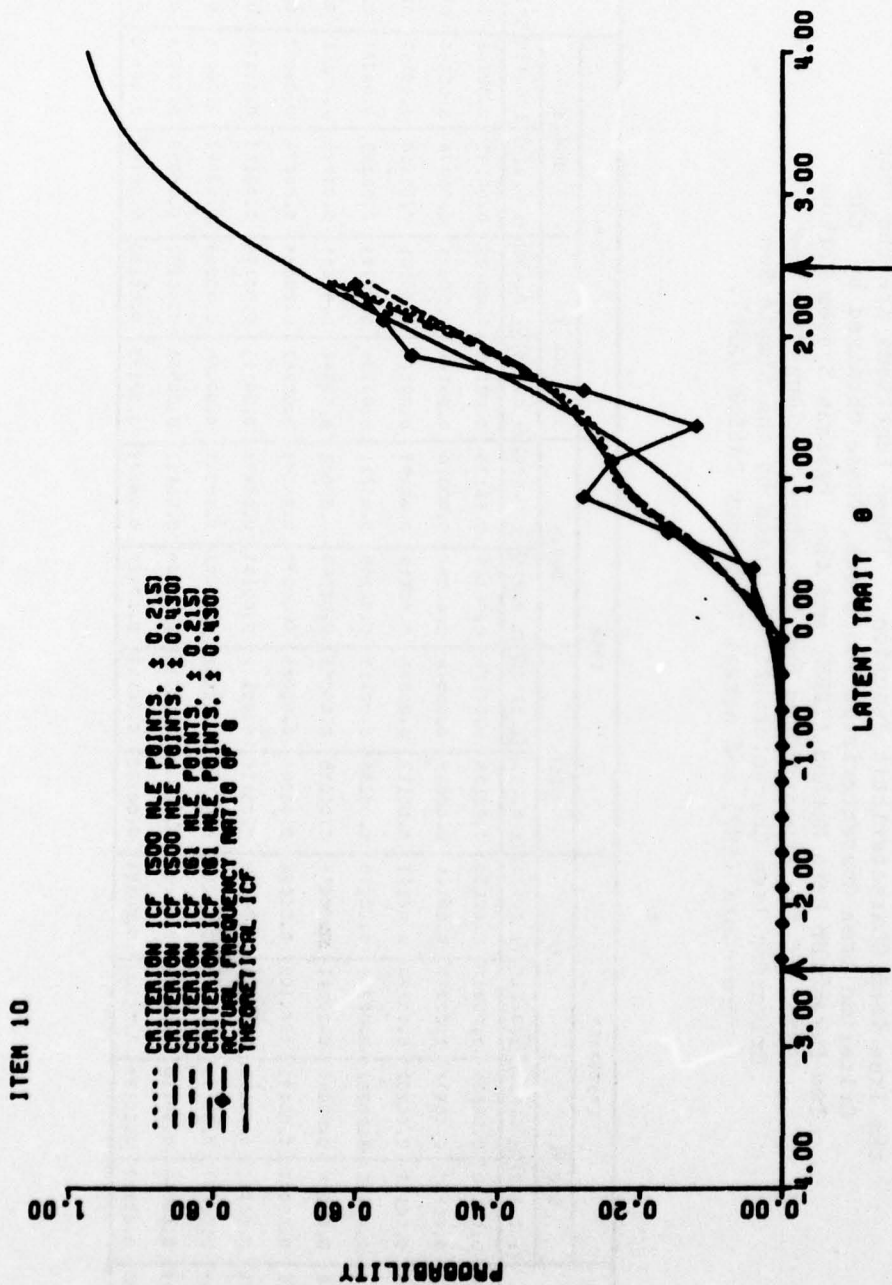


FIGURE 6-1: Criterion Case (Continued)

TABLE 6-1

Mean Square Errors of the Estimated Item Characteristic Functions Against the True Item Characteristic Function. These Functions Are the Four Criterion Item Characteristic Functions, Those Obtained by the Two-Parameter Beta Method (TPBM) and the Pearson System Method (PSM), of the Proportioned Sum Procedure, Along with the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP) and Actual Frequency Ratios (AFR).

ITEM	AFR	CRITERION						TPBM				PSM				SSP CRITERION		
		500 MLE		61 PLE		DGR.3		DGR.4		DGR.3		DGR.4		DGR.3			DGR.4	
		(± 0.215)	(± 0.430)	(± 0.215)	(± 0.430)	(± 0.215)	(± 0.430)	(± 0.215)	(± 0.430)	(± 0.215)	(± 0.430)	(± 0.215)	(± 0.430)	(± 0.215)	(± 0.430)		(± 0.215)	(± 0.430)
1	0.00068	0.00080	0.00109	0.00012	0.00030	0.00126	0.00129	0.00129	0.00129	0.00133	0.00133	0.00135	0.00129	0.00129	0.00129	0.00195	0.00148	0.00070
2	0.000782	0.00110	0.00076	0.00111	0.00095	0.00076	0.00087	0.00076	0.00087	0.00094	0.00076	0.00070	0.00075	0.00075	0.00076	0.00076	0.00065	0.00121
3	0.000649	0.00118	0.00202	0.00120	0.00197	0.00117	0.00201	0.00201	0.00110	0.00116	0.00195	0.00110	0.00201	0.00201	0.00108	0.00197	0.00098	0.00098
4	0.000940	0.00251	0.00222	0.00245	0.00210	0.00233	0.00215	0.00249	0.00231	0.00236	0.00249	0.00231	0.00215	0.00215	0.00252	0.00231	0.00265	0.00265
5	0.00242	0.00044	0.00060	0.00061	0.00054	0.00038	0.00055	0.00043	0.00043	0.00041	0.00043	0.00062	0.00041	0.00041	0.00045	0.00061	0.00055	0.00055
6	0.000601	0.00072	0.00053	0.00103	0.00056	0.00065	0.00085	0.00072	0.00093	0.00067	0.00093	0.00072	0.00085	0.00085	0.00074	0.00052	0.00085	0.00085
7	0.00179	0.00117	0.00179	0.00164	0.00183	0.00111	0.00171	0.00115	0.00115	0.00111	0.00178	0.00115	0.00170	0.00170	0.00115	0.00178	0.00120	0.00120
8	0.00368	0.00087	0.00079	0.00166	0.00069	0.00083	0.00082	0.00081	0.00081	0.00077	0.00081	0.00077	0.00083	0.00083	0.00087	0.00086	0.00105	0.00105
9	0.00135	0.00040	0.00116	0.00028	0.00060	0.00047	0.00134	0.00039	0.00039	0.00051	0.00039	0.00051	0.00032	0.00032	0.00058	0.00134	0.00051	0.00051
10	0.00325	0.00127	0.00135	0.00103	0.00165	0.00135	0.00149	0.00127	0.00138	0.00135	0.00127	0.00138	0.00150	0.00150	0.00142	0.00170	0.00124	0.00124

TABLE 6-2

Square Roots of the Mean Square Errors of the Estimated Item Characteristic Functions Against the True Item Characteristic Function. These Functions Are Criterion Item Characteristic Functions, Those Obtained by the Two-Parameter Beta Method (TPBM) and the Pearson System Method (PSM), of the Proportioned Sum Procedure, Along with the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP) and Actual Frequency Ratios (AFR).

ITEM	AFR	CRITERION				TPBM				PSM					
		500 MLE (± 0.215)	61 MLE (± 0.215)	DGR.3 (± 0.215)	DGR.4 (± 0.215)	DGR.3 (± 0.215)	DGR.4 (± 0.215)	DGR.3 (± 0.215)	DGR.4 (± 0.215)	DGR.3 (± 0.215)	DGR.4 (± 0.215)	DGR.3 (± 0.215)	DGR.4 (± 0.215)	SSP CRITERION	
1	0.02610	0.02830	0.03301	0.01149	0.01728	0.03546	0.03595	0.03597	0.03681	0.03666	0.03685	0.03585	0.04421	0.03842	0.02650
2	0.02643	0.03313	0.02762	0.04801	0.03225	0.03085	0.02750	0.02955	0.02648	0.03062	0.02746	0.02746	0.02756	0.02621	0.03473
3	0.02223	0.03429	0.04456	0.03465	0.04436	0.03414	0.04480	0.03310	0.04413	0.03404	0.04479	0.04479	0.03292	0.04442	0.03138
4	0.05656	0.05013	0.04820	0.04947	0.04557	0.04830	0.04635	0.04995	0.04802	0.04854	0.04640	0.04640	0.05020	0.04805	0.05187
5	0.04923	0.02054	0.02452	0.02476	0.02331	0.01943	0.02349	0.02062	0.02481	0.02020	0.02352	0.02352	0.02132	0.02479	0.02421
6	0.07154	0.02680	0.03051	0.03212	0.03057	0.02554	0.02918	0.02685	0.03050	0.02586	0.02916	0.02916	0.02714	0.03027	0.02582
7	0.04236	0.03420	0.04259	0.04053	0.04262	0.03327	0.04130	0.03397	0.04224	0.03332	0.04127	0.04127	0.03398	0.04219	0.03458
8	0.06670	0.02952	0.02807	0.03252	0.02830	0.02884	0.02871	0.02840	0.02783	0.02927	0.02874	0.02874	0.02942	0.02538	0.03306
9	0.03671	0.02011	0.03405	0.01662	0.02453	0.02172	0.03658	0.01984	0.03480	0.02265	0.03635	0.03635	0.02417	0.03666	0.02260
10	0.05701	0.03558	0.03673	0.03205	0.03236	0.03676	0.03865	0.03561	0.03717	0.03674	0.03877	0.03877	0.03767	0.04118	0.03526

TABLE 6-3

Estimated Item Discrimination Parameters,  $\hat{a}_g$ , Obtained from the Four Criterion Item Characteristic Functions of the Proportioned Sum Procedure of the Conditional P.D.F. Approach, with the One from the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP), for Each of the Ten Binary Items. Estimation Is Based on the Least Squares Principle, Using, at Most, 25 Points of  $\theta$  in the Interval,  $[-2.4, 2.4]$ , Excluding the Points for Which  $P_g(\theta)$  Is Less Than 0.05 or Greater Than 0.95.

ITEM	TRUE	CRITERION				SSP CRITERION
		500 MLE		61 MLE		
		$\pm 0.215$	$\pm 0.430$	$\pm 0.215$	$\pm 0.430$	
1	1.5	1.330 <sub>5</sub>	1.064 <sub>5</sub>	1.683 <sub>5</sub>	1.327 <sub>5</sub>	1.400 <sub>5</sub>
2	1.0	1.023	0.970	0.945	0.939	1.024
3	2.5	1.740	1.642	1.755	1.660	1.788
4	1.0	0.873	0.821	0.854	0.836	0.868
5	1.5	1.377	1.286	1.372	1.303	1.368
6	1.0	0.894	0.882	0.872	0.872	0.895
7	2.0	1.475	1.415	1.420	1.405	1.473
8	1.0	0.910	0.873	0.885	0.916	0.886
9	2.0	1.708	1.522	1.741	1.653	1.716
10	1.0	0.757	0.740	0.797	0.769	0.725

THE NUMBER OF INTERVALS USED IN ESTIMATION IS SHOWN AS A SUBSCRIPT WHEN IT IS LESS THAN 6.

TABLE 6-4

Estimated Item Difficulty Parameters,  $\hat{b}_g$ , Obtained from the Four Criterion Item Characteristic Functions of the Proportioned Sum Procedure of the Conditional P.D.F. Approach, with the One from the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP), for Each of the Ten Binary Items. Estimation Is Based On the Least Squares Principle, Using, at Most, 25 Points of  $\theta$  in the Interval,  $[-2.4, 2.4]$ , Excluding the Points for Which  $\hat{P}_g(\theta)$  Is Less Than 0.05 or Greater Than 0.95.

ITEM	TRUE	CRITERION				SSP CRITERION
		500 MLE		61 MLE		
		$\pm 0.215$	$\pm 0.430$	$\pm 0.215$	$\pm 0.430$	
1	-2.5	-2.692 <sub>5</sub>	-2.857 <sub>5</sub>	-2.467 <sub>5</sub>	-2.638 <sub>5</sub>	-2.651 <sub>5</sub>
2	-2.0	-2.001	-2.031	-2.075	-2.036	-2.002
3	-1.5	-1.512	-1.526	-1.509	-1.513	-1.507
4	-1.0	-0.997	-1.030	-1.016	-1.022	-1.005
5	-0.5	-0.471	-0.468	-0.454	-0.467	-0.472
6	0.0	-0.063	-0.061	-0.083	-0.070	-0.075
7	0.5	0.522	0.522	0.533	0.527	0.527
8	1.0	0.976	0.977	0.951	0.940	0.981
9	1.5	1.518	1.511	1.490	1.496	1.502
10	2.0	2.116	2.157	2.048	2.080	2.118

THE NUMBER OF INTERVALS USED IN ESTIMATION IS SHOWN AS A SUBSCRIPT WHEN IT IS LESS THAN 6.

As was the case of the Simple Sum and Weighted Sum Procedures, we notice that the estimates of the discrimination parameters  $a_g$  are further from the true values when  $a_g$ 's are large, like for items 3, 7 and 9. It is interesting to note that, although the parameter estimation tends to be more inaccurate for very difficult and very easy items in most cases, this tendency is ameliorated when the 61 point case is used, in both the discrimination and difficulty parameter estimations.

Figure 6-2 presents the resulting estimated item characteristic functions obtained by Degree 3 and 4 Cases of the Pearson System Method, each of which has two results using the intervals,  $\hat{\theta}_g \pm 0.215$  and  $\hat{\theta}_g \pm 0.430$ , in calculating the proportioned weights, respectively. In the same figure, for each item, also presented are the criterion item characteristic function obtained by the Simple Sum Procedure, the actual frequency ratios of  $\theta$  and the theoretical item characteristic function. We notice that these four curves for each item are fairly close to one another, and also to the criterion item characteristic function obtained by the Simple Sum Procedure, for the interval of  $\theta$ , (-2.5, 2.5). It is also noted that these estimated item characteristic functions are similar to the corresponding criterion item characteristic functions, if we compare Figure 6-2 with Figure 6-1.

The mean square errors and their square roots, which were calculated through (6.1), are shown in Tables 6-1 and 6-2, respectively. We can see that these values are very close to those for the criterion item characteristic functions, which are presented in the same tables. There are no tendencies that those obtained by Degree 4 Case are smaller

than those obtained by Degree 3 Case of the Pearson System Method, or vice versa.

Tables 6-5 and 6-6 present the estimated discrimination parameter,  $\hat{a}_g$ , and difficulty parameter,  $\hat{b}_g$ , by the least square method, for each of the ten binary items, in comparison with those obtained from the criterion item characteristic function by the Simple Sum Procedure. These results are very similar to those obtained from the criterion item characteristic functions, and the same tendencies as were described earlier for the estimates from the criterion item characteristic functions are also observed among these estimates, although the latter tend to be slightly more inaccurate than the former. It is interesting to note that the estimates of the discrimination parameters in Degree 3 Case with the interval  $\hat{\theta}_s \pm 0.215$  tend to be better than the others, although the differences are small.

Figure 6-3 presents the estimated item characteristic functions of each of the ten binary items obtained by Degree 3 and 4 Cases of the Two-Parameter Beta Method, each having two functions by using the two intervals,  $\hat{\theta}_s \pm 0.215$  and  $\hat{\theta}_s \pm 0.430$  for calculating the proportioned weights, along with the criterion item characteristic function by the Simple Sum Procedure of the Conditional P.D.F. Approach, the actual frequency ratios of  $\theta$  and the theoretical item characteristic function. Again, these resultant functions are close to one another, to the criterion item characteristic function by the Simple Sum Procedure, and to the criterion item characteristic functions by the present procedure, which are presented in Figure 6-1, for the interval of  $\theta$ ,  $(-2.5, 2.5)$ .

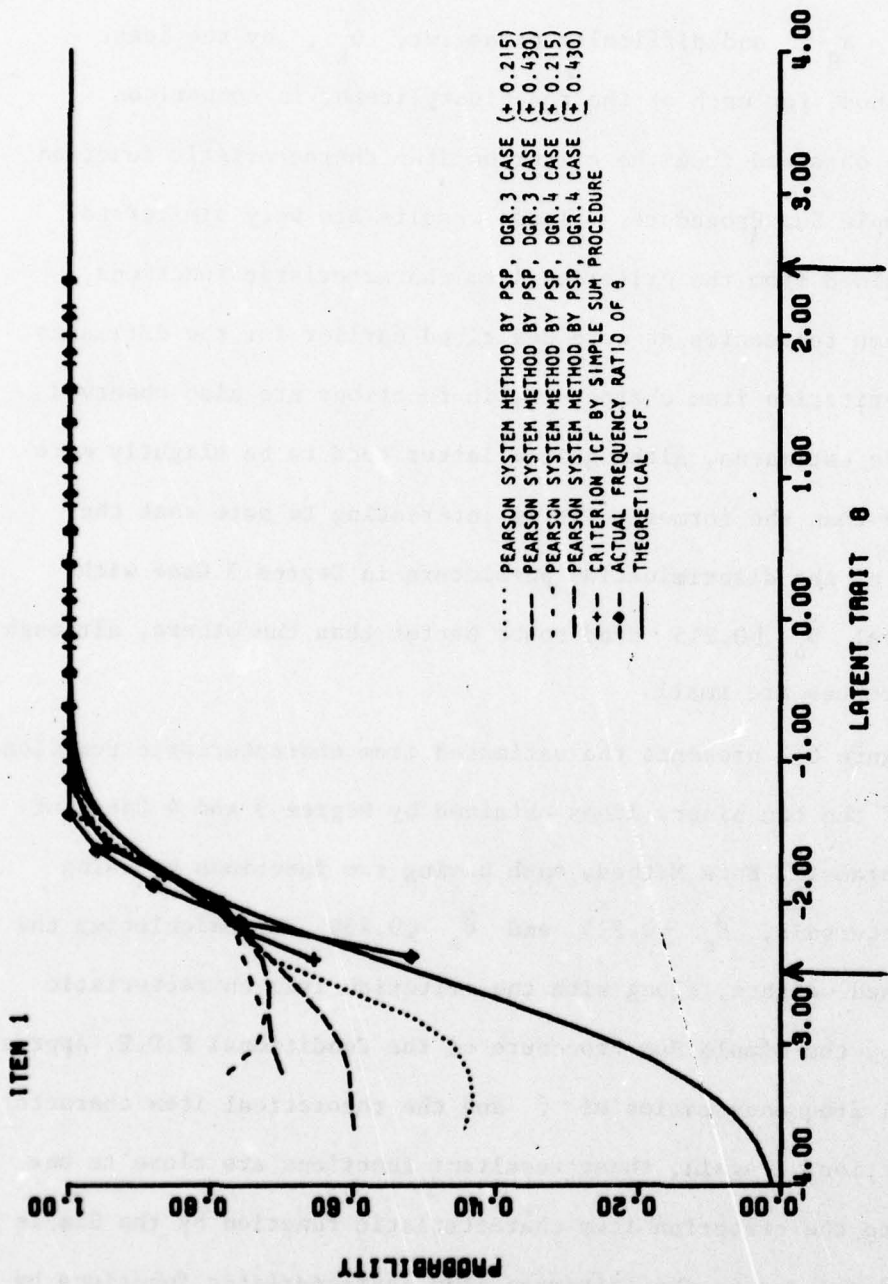


FIGURE 6-2

Comparison of the Four Estimated Item Characteristic Functions of Degree 3 and 4 Cases by the Pearson System Method of the Proportioned Sum Procedure with the Criterion Item Characteristic Function by the Simple Sum Procedure. Actual Frequency Ratios and Theoretical Item Characteristic Function Are Also Presented for Comparison.



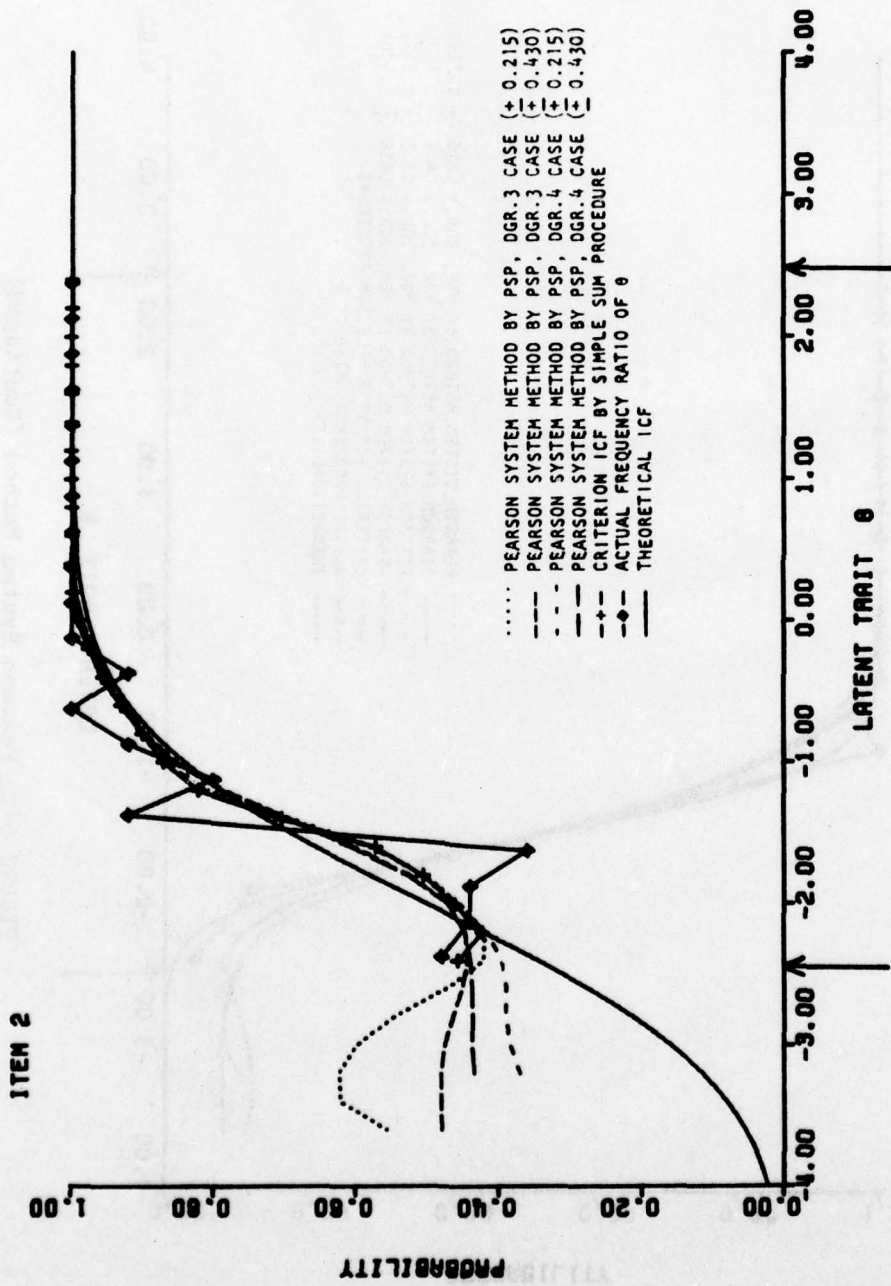


FIGURE 6-2: Pearson System Method (Continued)

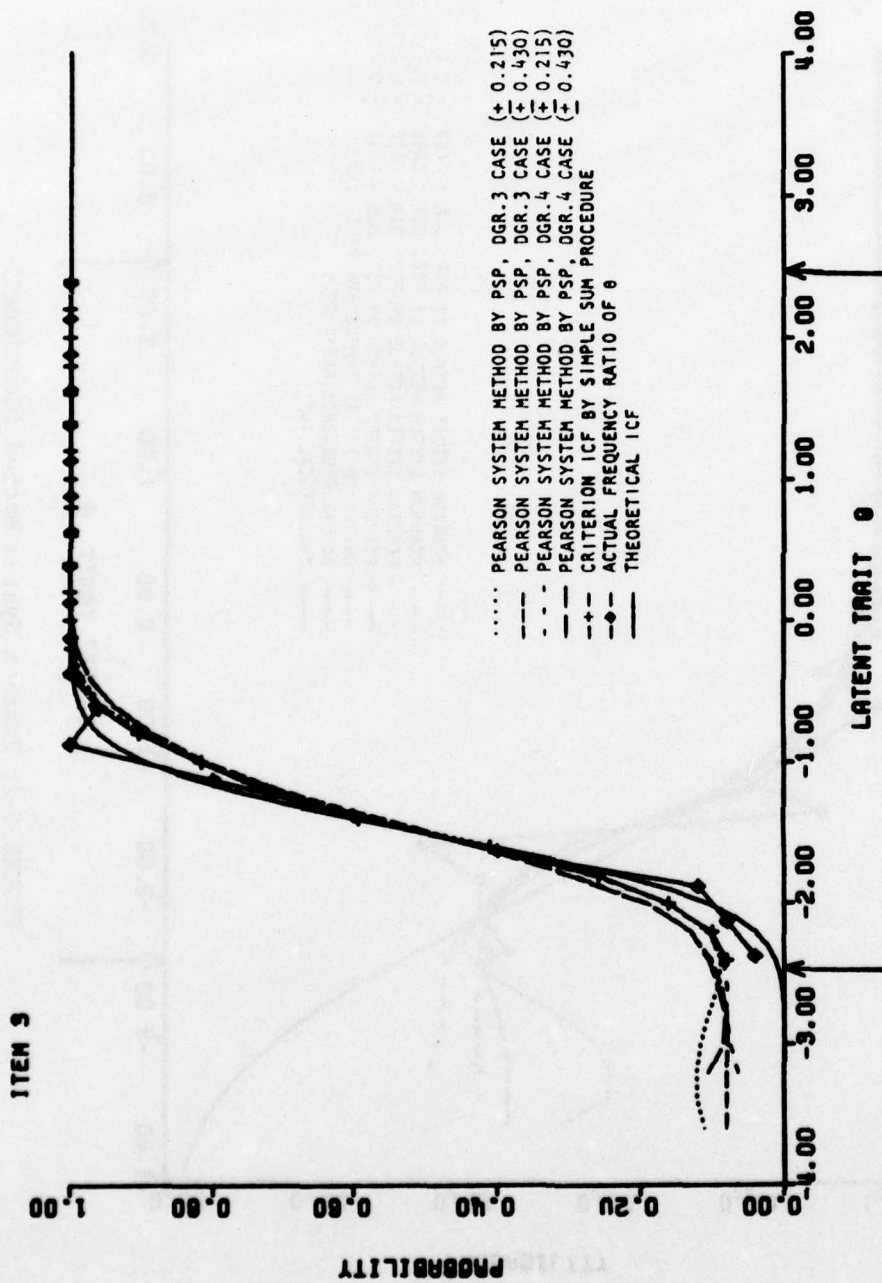


FIGURE 6-2: Pearson System Method (Continued)

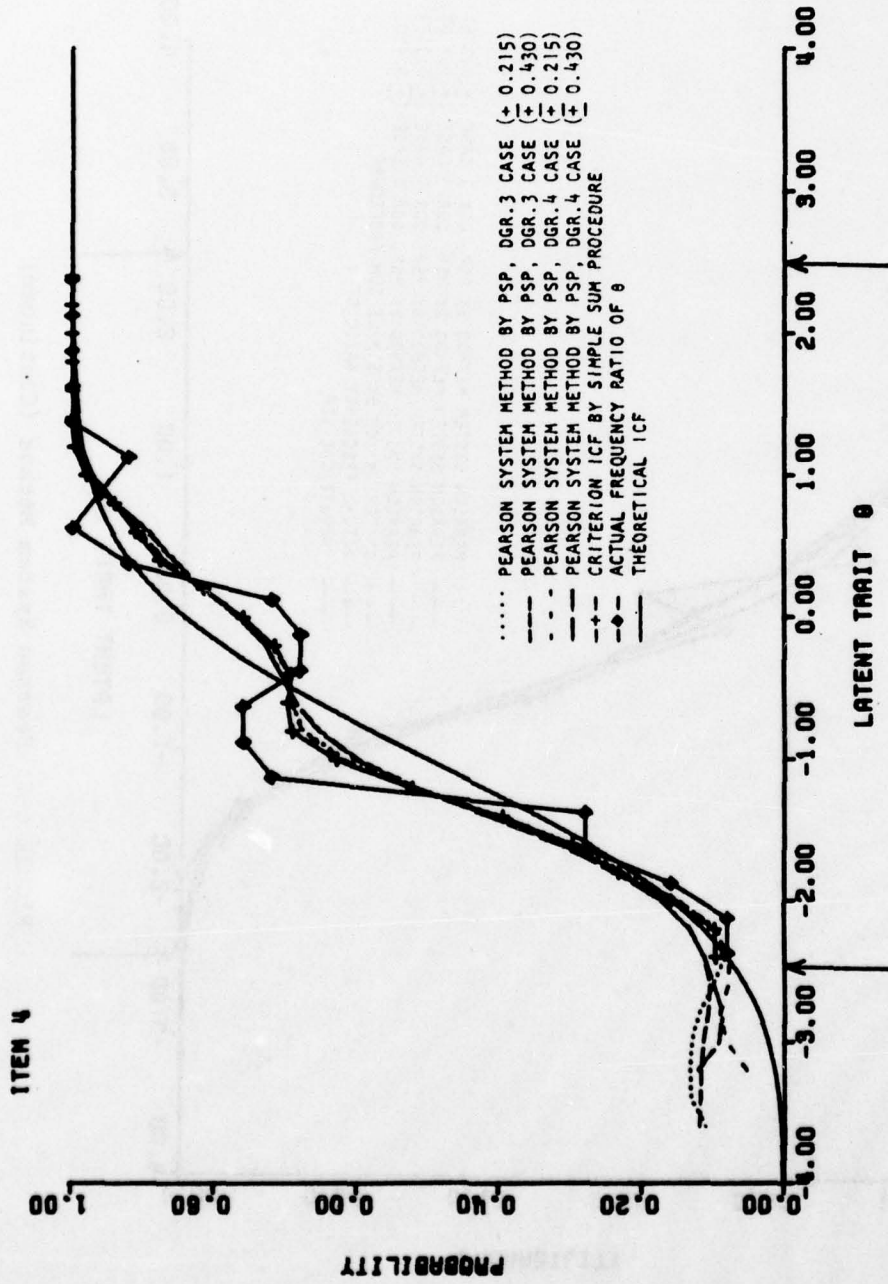


FIGURE 6-2: Pearson System Method (Continued)

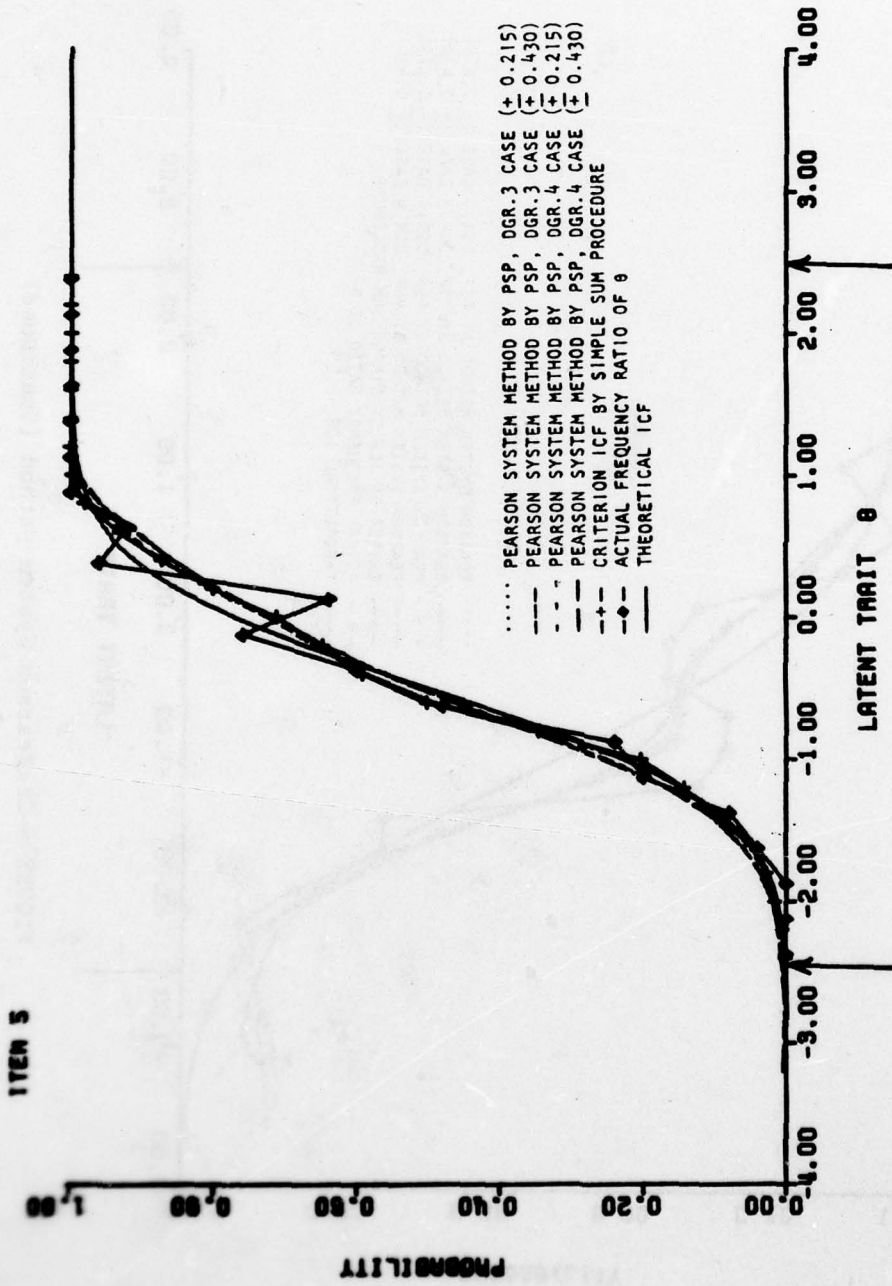


FIGURE 6-2: Pearson System Method (Continued)

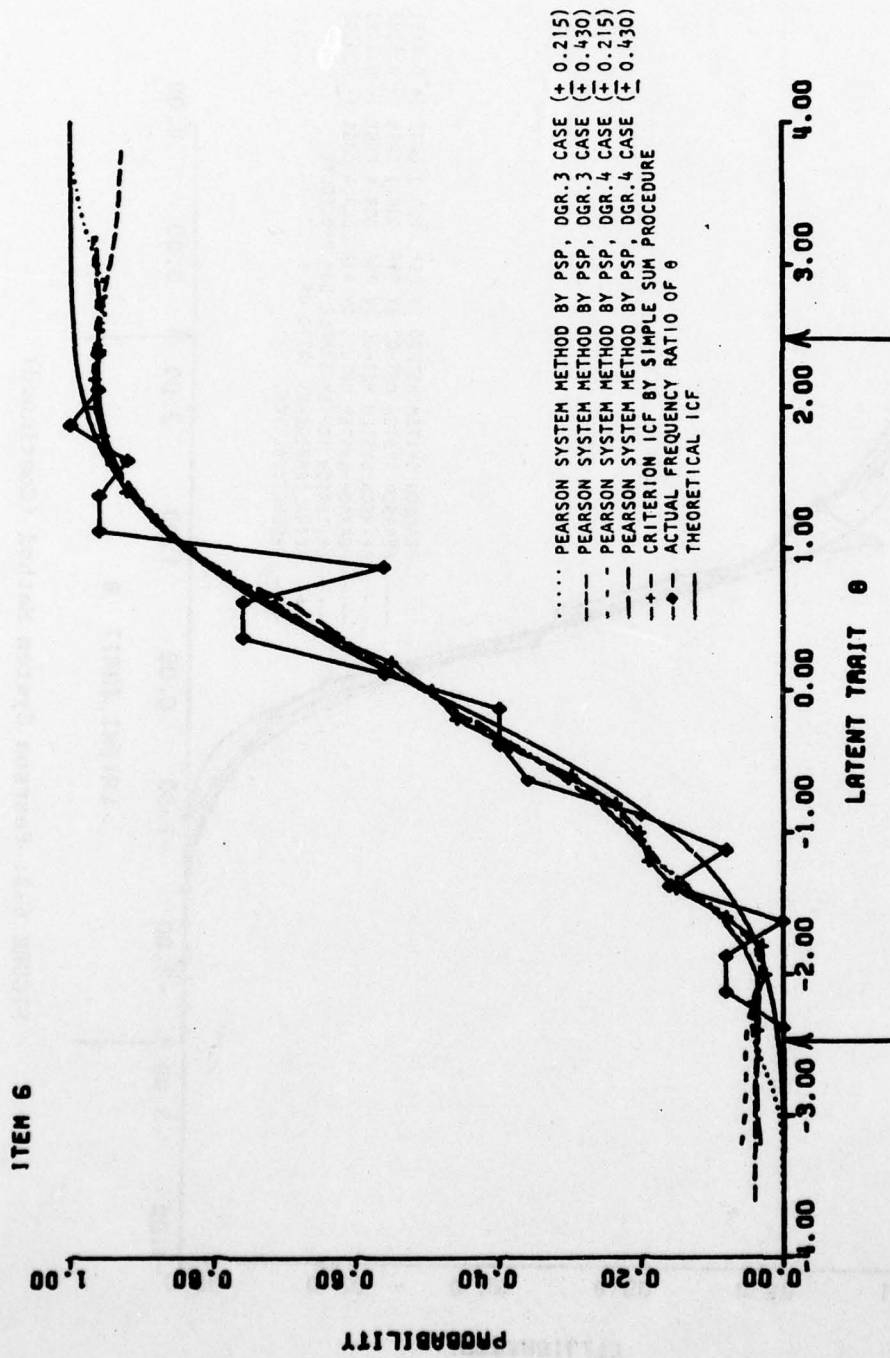


FIGURE 6-2: Pearson System Method (Continued)

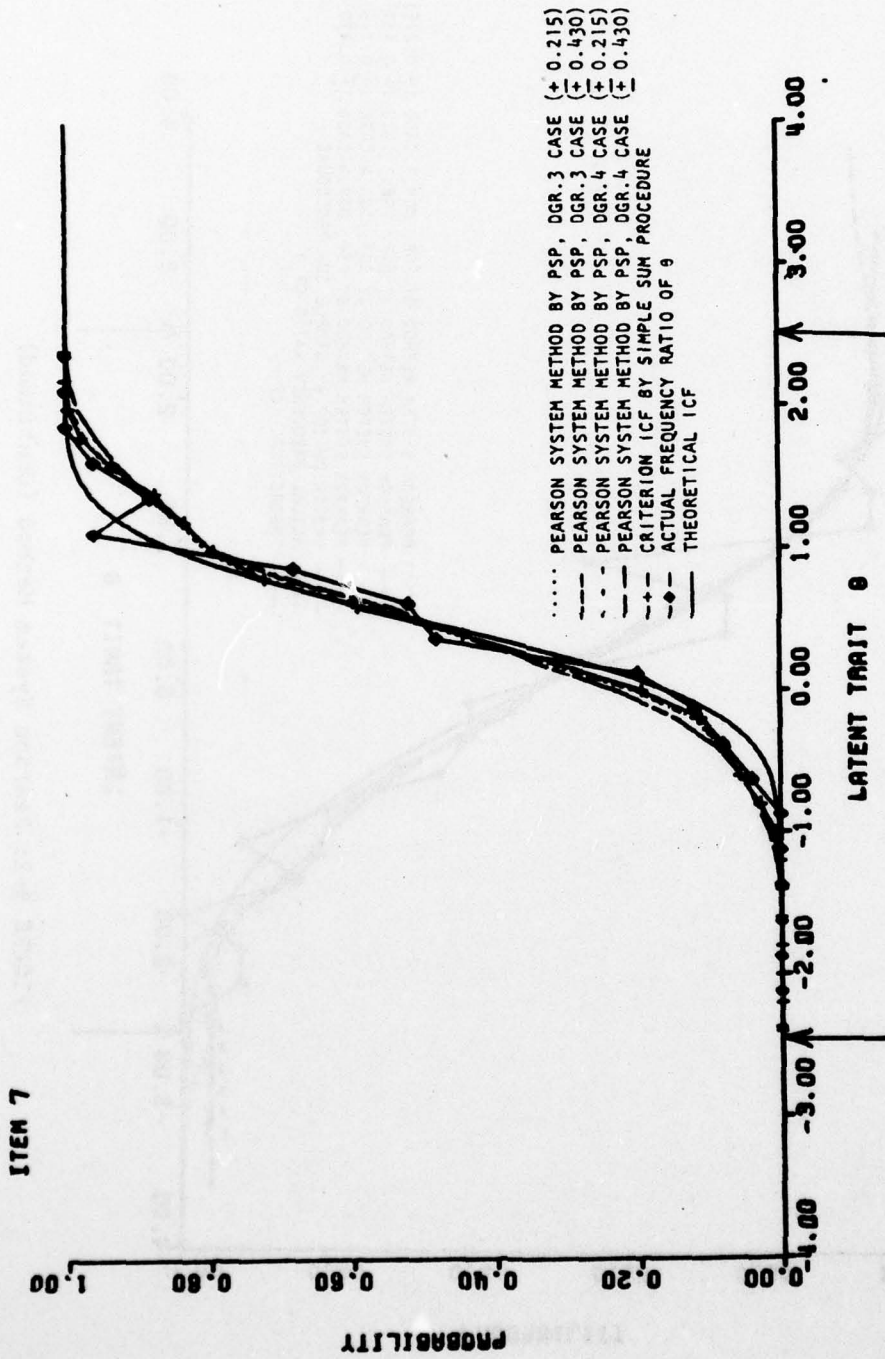


FIGURE 6-2: Pearson System Method (Continued)

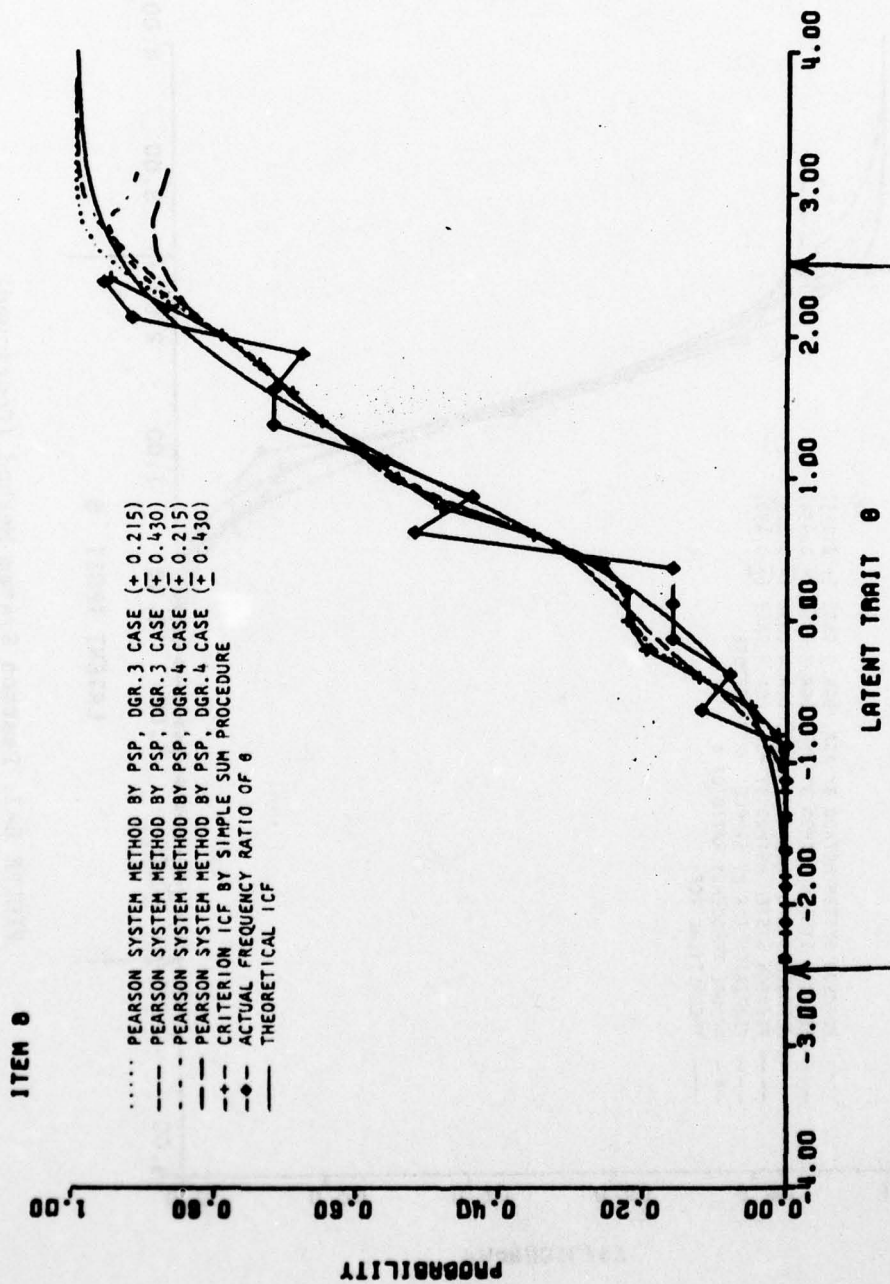


FIGURE 6-2: Pearson System Method (Continued)

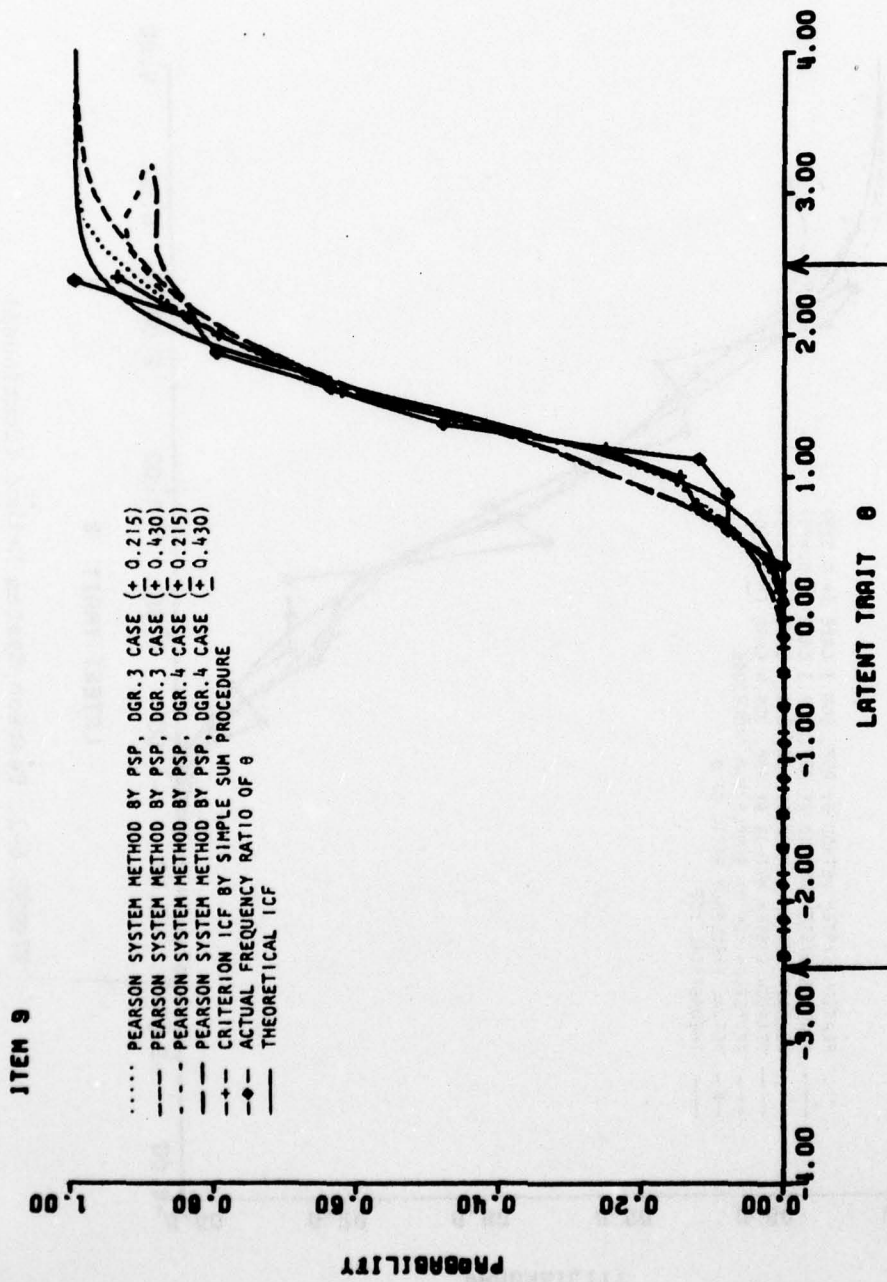


FIGURE 6-2: Pearson System Method (Continued)



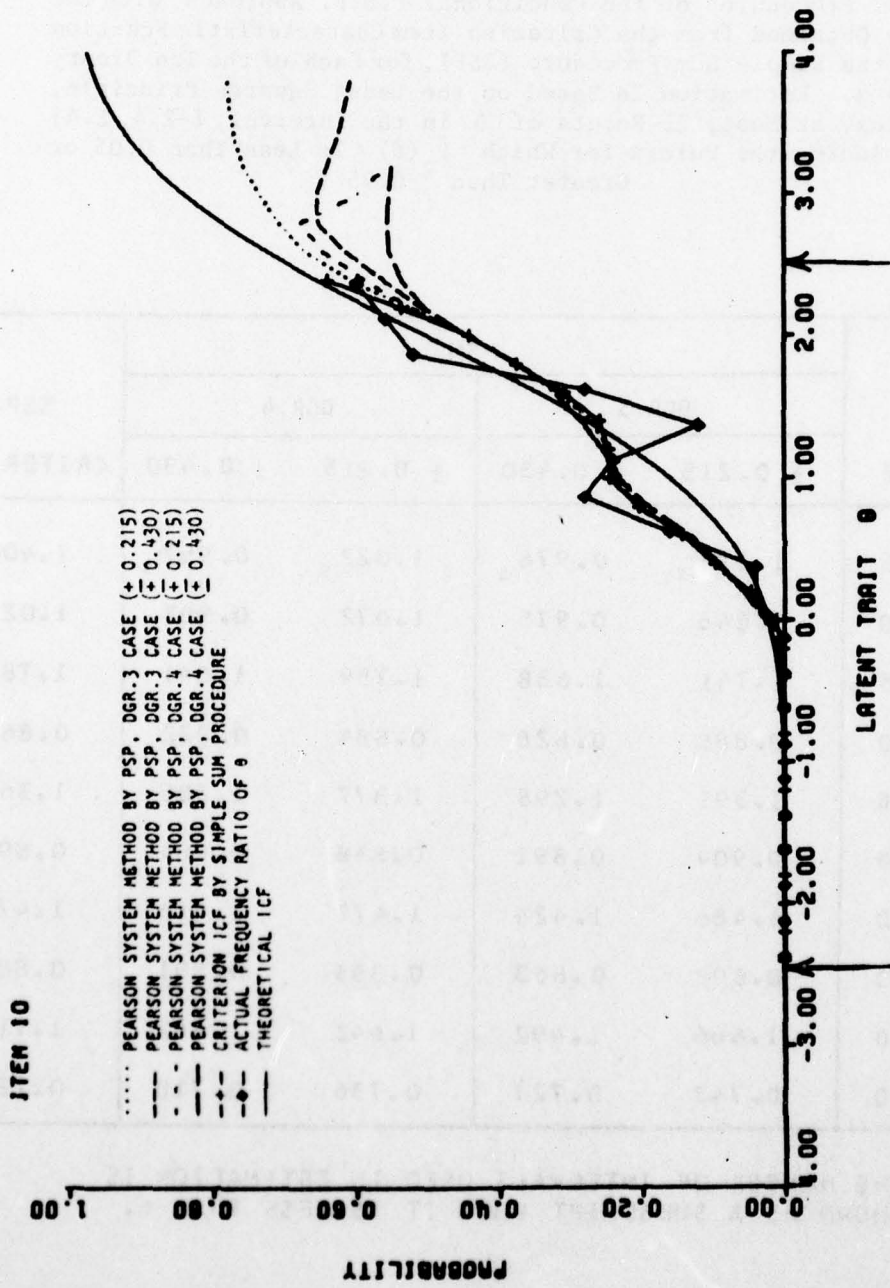


FIGURE 6-2: Pearson System Method (Continued)

TABLE 6-5

Estimated Item Discrimination Parameters,  $\hat{a}_i$ , Obtained from the Four Estimated Item Characteristic Functions by Degree 3 and 4 Cases of the Pearson System Method (PSM) of the Proportioned Sum Procedure of the Conditional P.D.F. Approach, with the One Obtained from the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP), for Each of the Ten Binary Items. Estimation Is Based on the Least Squares Principle, Using, at Most, 25 Points of  $\theta$  in the Interval,  $[-2.4, 2.4]$  Excluding the Points for Which  $\hat{P}_i(\theta)$  Is Less Than 0.05 or Greater Than  $\frac{1}{8}$  0.95 .

ITEM	TRUE	PSM				SSP CRITERION
		DGR.3		DGR.4		
		$\pm 0.215$	$\pm 0.430$	$\pm 0.215$	$\pm 0.430$	
1	1.5	1.123 <sub>5</sub>	0.976 <sub>5</sub>	1.022 <sub>5</sub>	0.957 <sub>5</sub>	1.400 <sub>5</sub>
2	1.0	1.048	0.975	1.072	0.983	1.024
3	2.5	1.741	1.638	1.759	1.641	1.788
4	1.0	0.886	0.828	0.884	0.822	0.868
5	1.5	1.391	1.298	1.377	1.285	1.368
6	1.0	0.904	0.891	0.848	0.884	0.895
7	2.0	1.486	1.424	1.477	1.416	1.473
8	1.0	0.898	0.863	0.883	0.851	0.886
9	2.0	1.666	1.492	1.642	1.482	1.716
10	1.0	0.742	0.727	0.736	0.718	0.725

THE NUMBER OF INTERVALS USED IN ESTIMATION IS SHOWN AS A SUBSCRIPT WHEN IT IS LESS THAN 6.

TABLE 6-6

Estimated Item Difficulty Parameters,  $\hat{b}_i$ , Obtained from the Four Estimated Item Characteristic Functions by Degree 3 and 4 Cases of the Pearson System Method (PSM) of the Proportioned Sum Procedure of the Conditional P.D.F. Approach, with the One Obtained from the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP), for Each of the Ten Binary Items. Estimation Is Based on the Least Squares Principle, Using, at Most, 25 Points of  $\theta$  in the Interval,  $[-2.4, 2.4]$ , Excluding the Points for Which  $\hat{P}_i(\theta)$  Is Less than 0.05 or Greater Than  $0.95$ .

ITEM	TRUE	FSM				SSP CRITERION
		DGR.3		DGR.4		
		$\pm 0.215$	$\pm 0.430$	$\pm 0.215$	$\pm 0.430$	
1	-2.5	-2.833 <sub>5</sub>	-2.933 <sub>5</sub>	-2.952 <sub>5</sub>	-2.968 <sub>5</sub>	-2.651 <sub>5</sub>
2	-2.0	-1.970	-2.015	-1.960	-2.019	-2.002
3	-1.5	-1.500	-1.517	-1.507	-1.525	-1.507
4	-1.0	-0.984	-1.023	-0.984	-1.027	-1.005
5	-0.5	-0.467	-0.464	-0.471	-0.468	-0.472
6	0.0	-0.062	-0.061	-0.088	-0.060	-0.075
7	0.5	0.517	0.517	0.522	0.522	0.527
8	1.0	0.983	0.982	0.996	0.994	0.981
9	1.5	1.519	1.509	1.533	1.521	1.502
10	2.0	2.134	2.172	2.154	2.198	2.118

THE NUMBER OF INTERVALS USED IN ESTIMATION IS SHOWN AS A SUBSCRIPT WHEN IT IS LESS THAN 6.

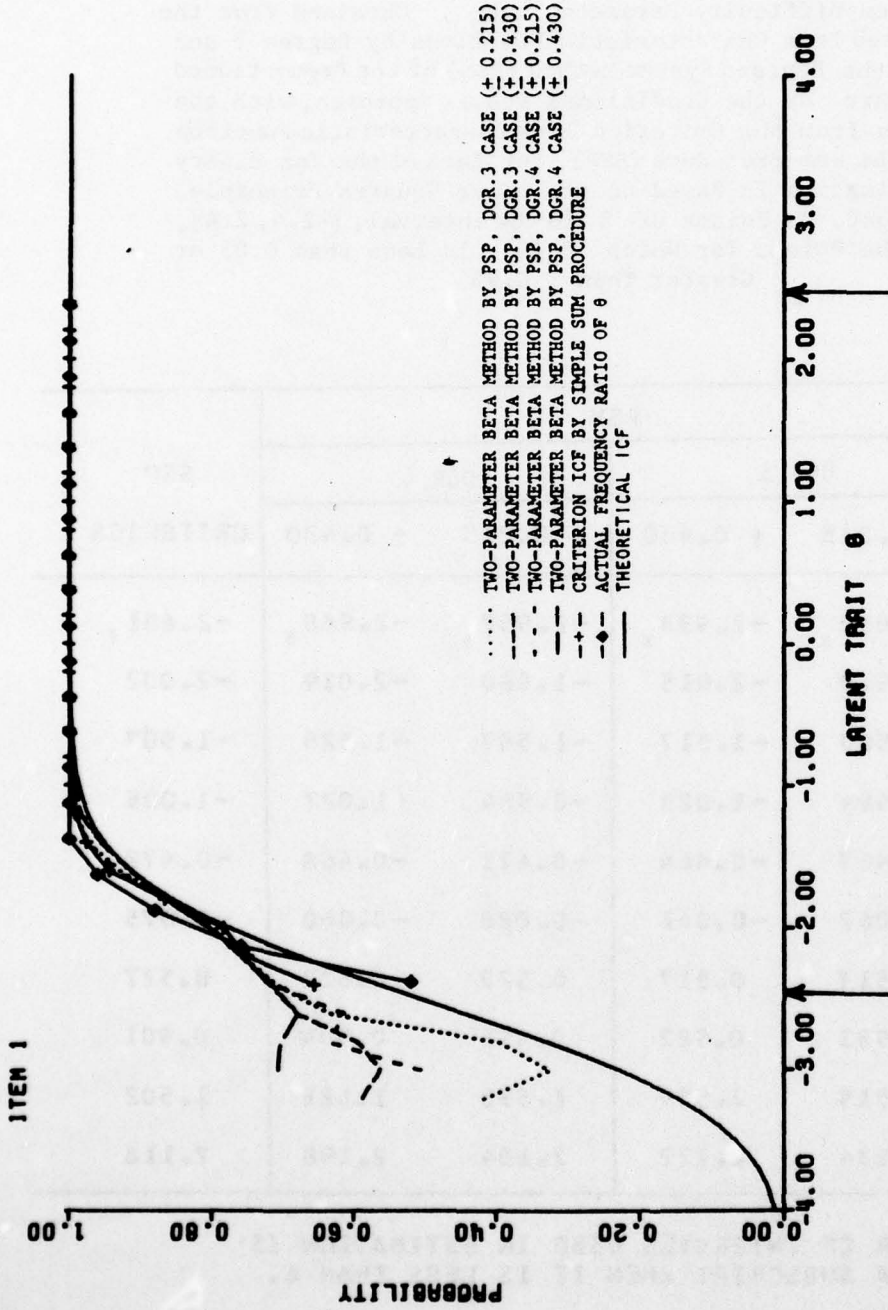


FIGURE 6-3

Comparison of the Four Estimated Item Characteristic Functions of Degree 3 and 4 Cases by the Two-Parameter Beta Method of the Proportioned Sum Procedure with the Criterion Item Characteristic Function by the Simple Sum Procedure. Actual Frequency Ratios and Theoretical Item Characteristic Function Are Also Presented for Comparison.

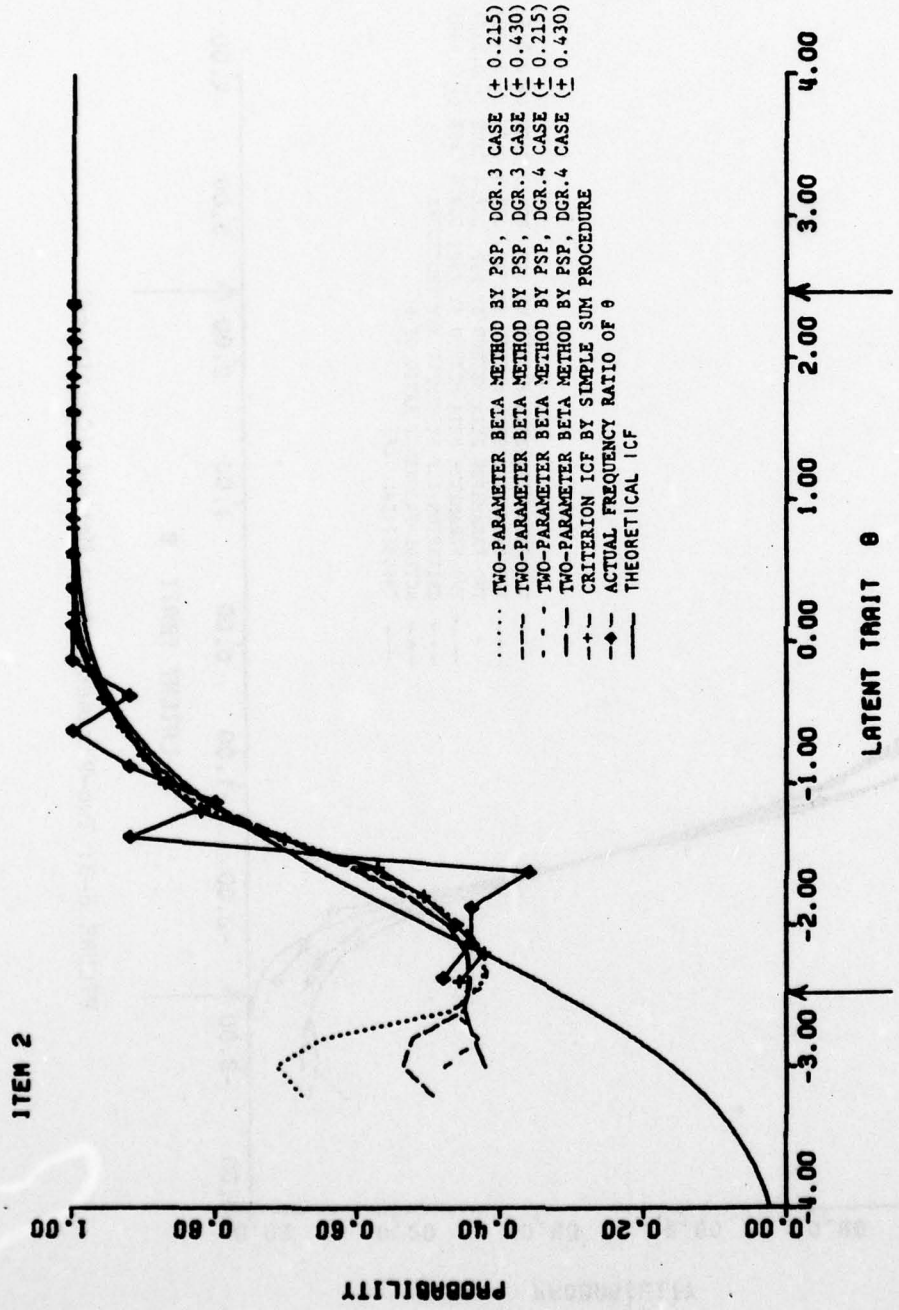


FIGURE 6-3: Two-Parameter Beta Method (Continued)

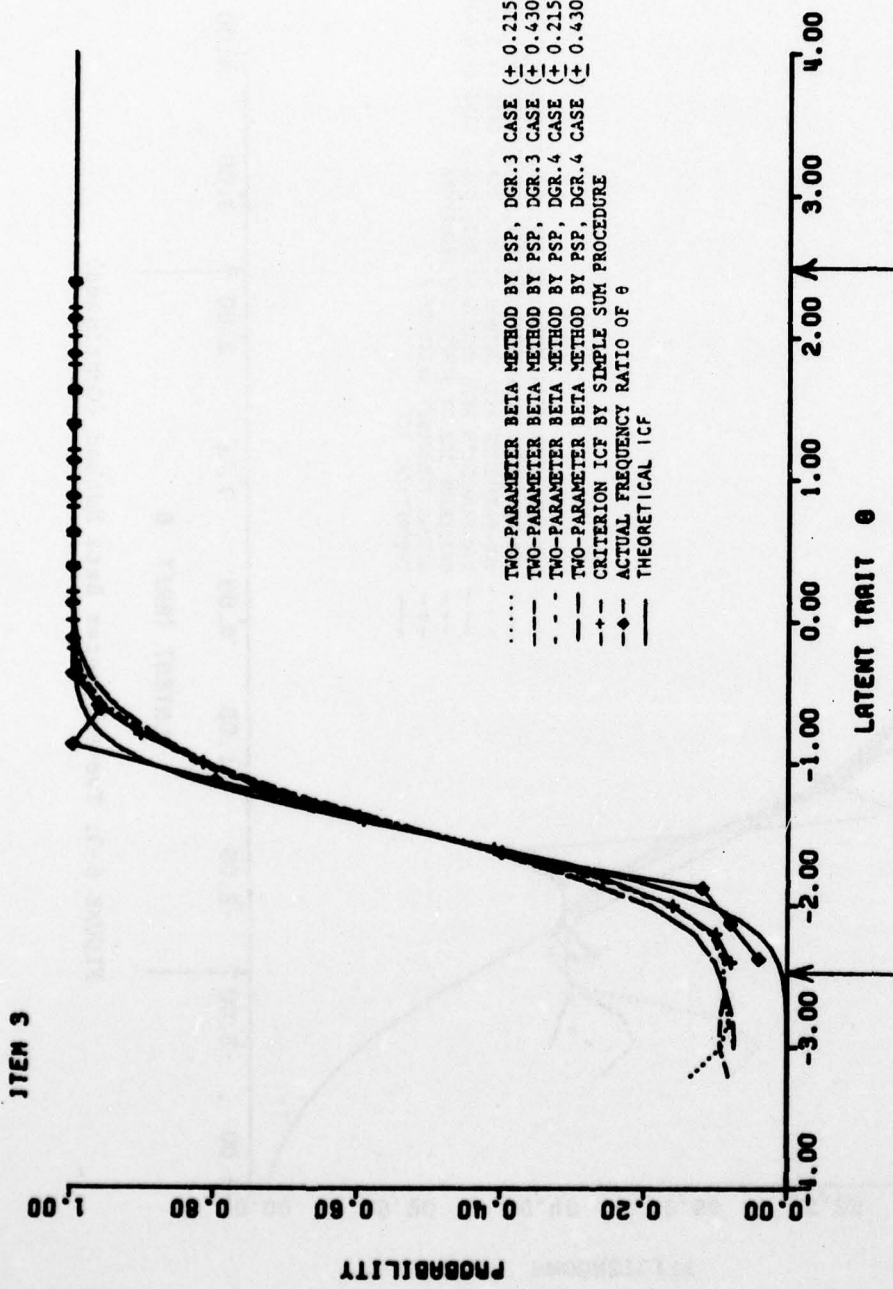


FIGURE 6-3: Two-Parameter Beta Method (Continued)

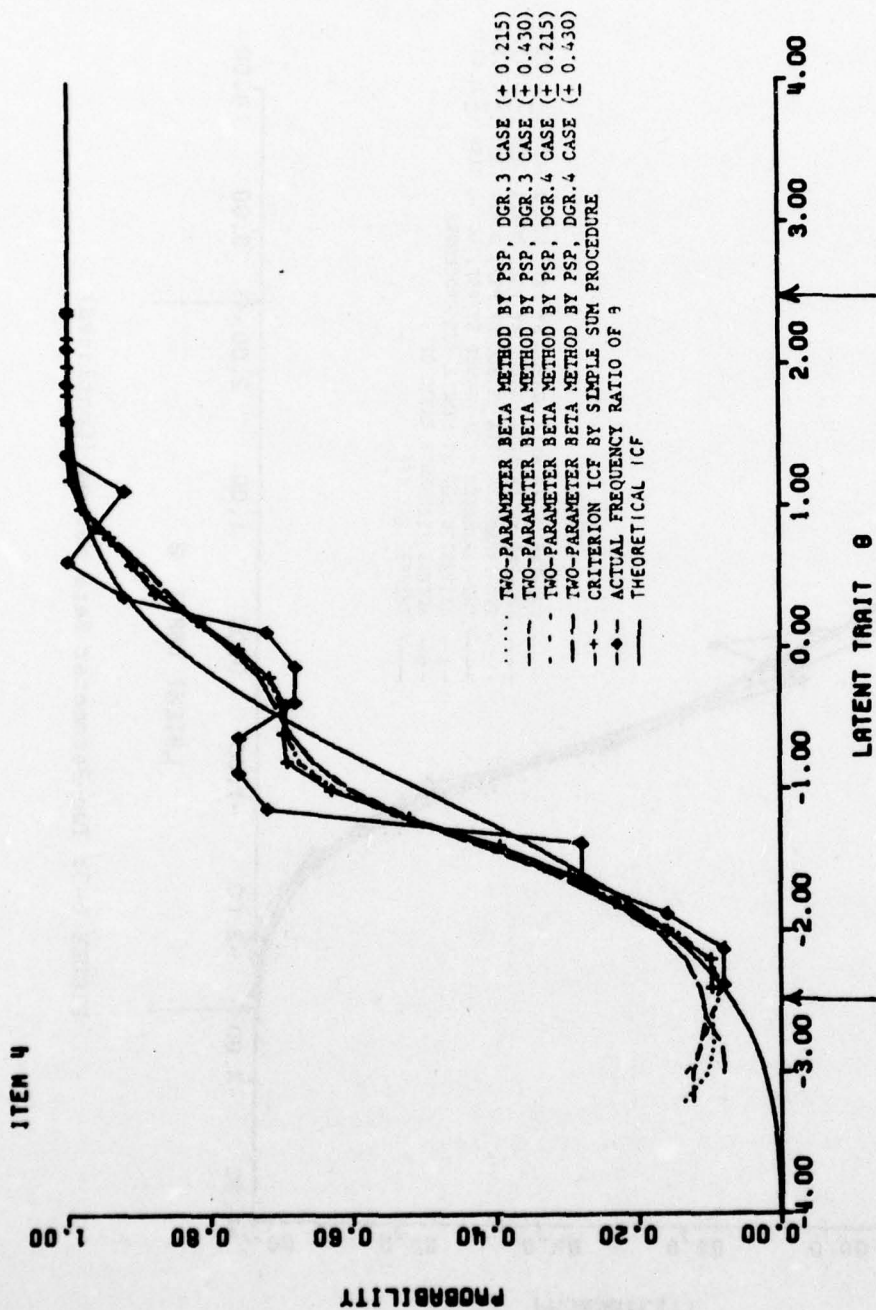


FIGURE 6-3: Two-Parameter Beta Method (Continued)

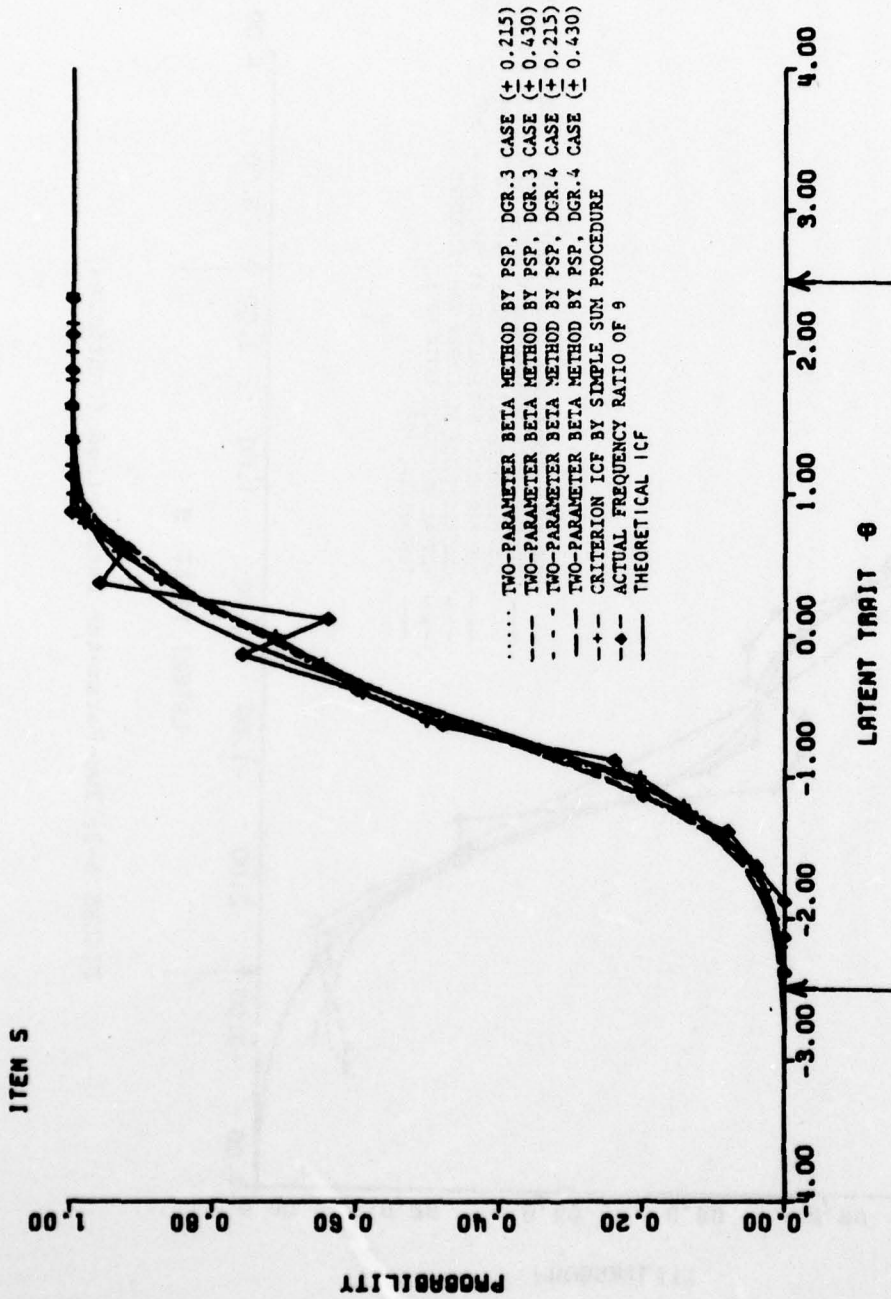


FIGURE 6-3: Two-Parameter Beta Method (Continued)



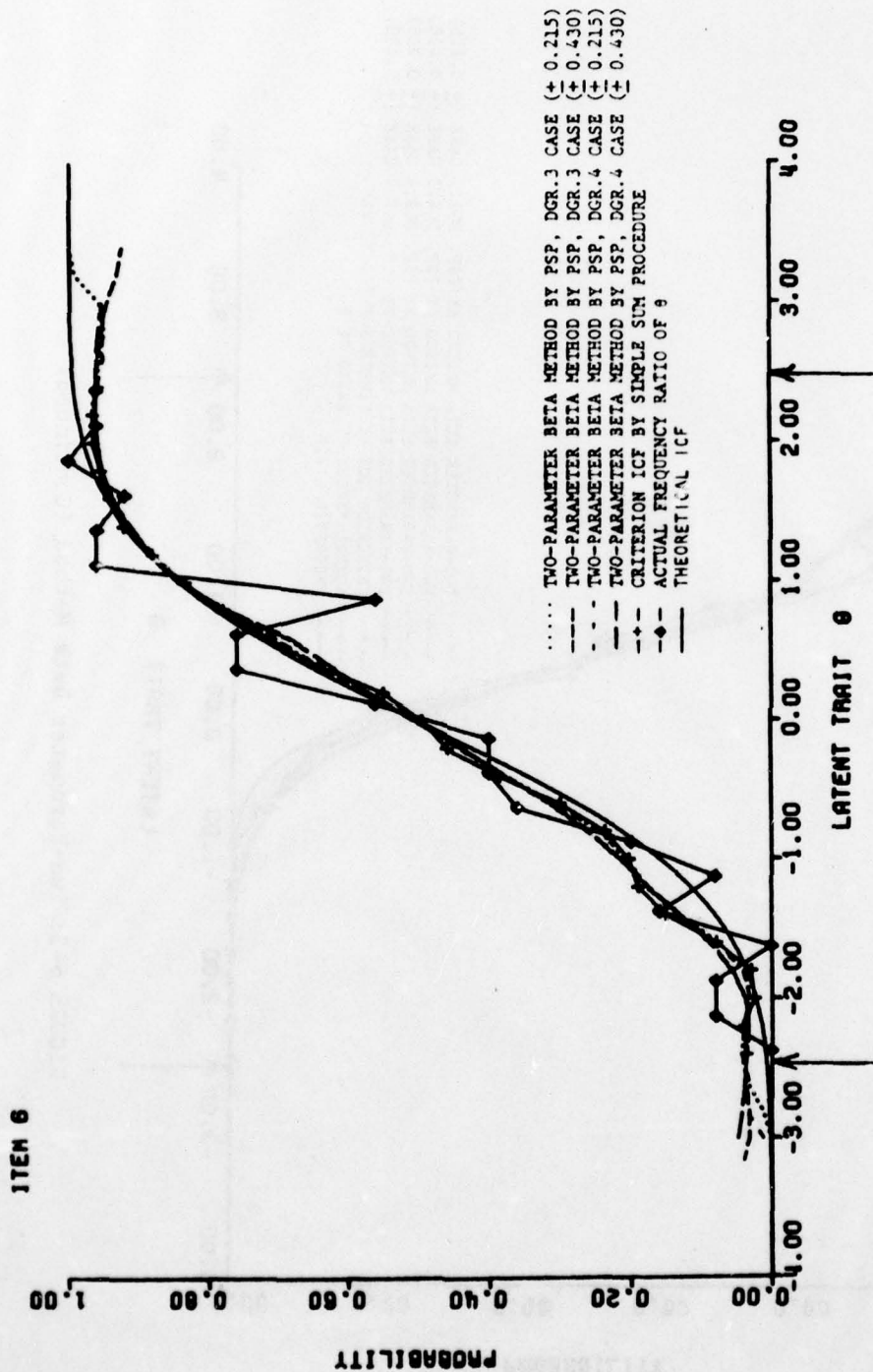


FIGURE 6-3: Two-Parameter Beta Method (Continued)

AD-A063 918

TENNESSEE UNIV KNOXVILLE DEPT OF PSYCHOLOGY

F/G 12/1

ESTIMATION OF THE OPERATING CHARACTERISTICS OF ITEM RESPONSE CA--ETC(U)

DEC 78 F SAMEJIMA

N00014-77-C-0360

UNCLASSIFIED

RR-78-5

NL

2 OF 2

AD  
A063918



END  
DATE  
FILMED  
4 -79

DDC

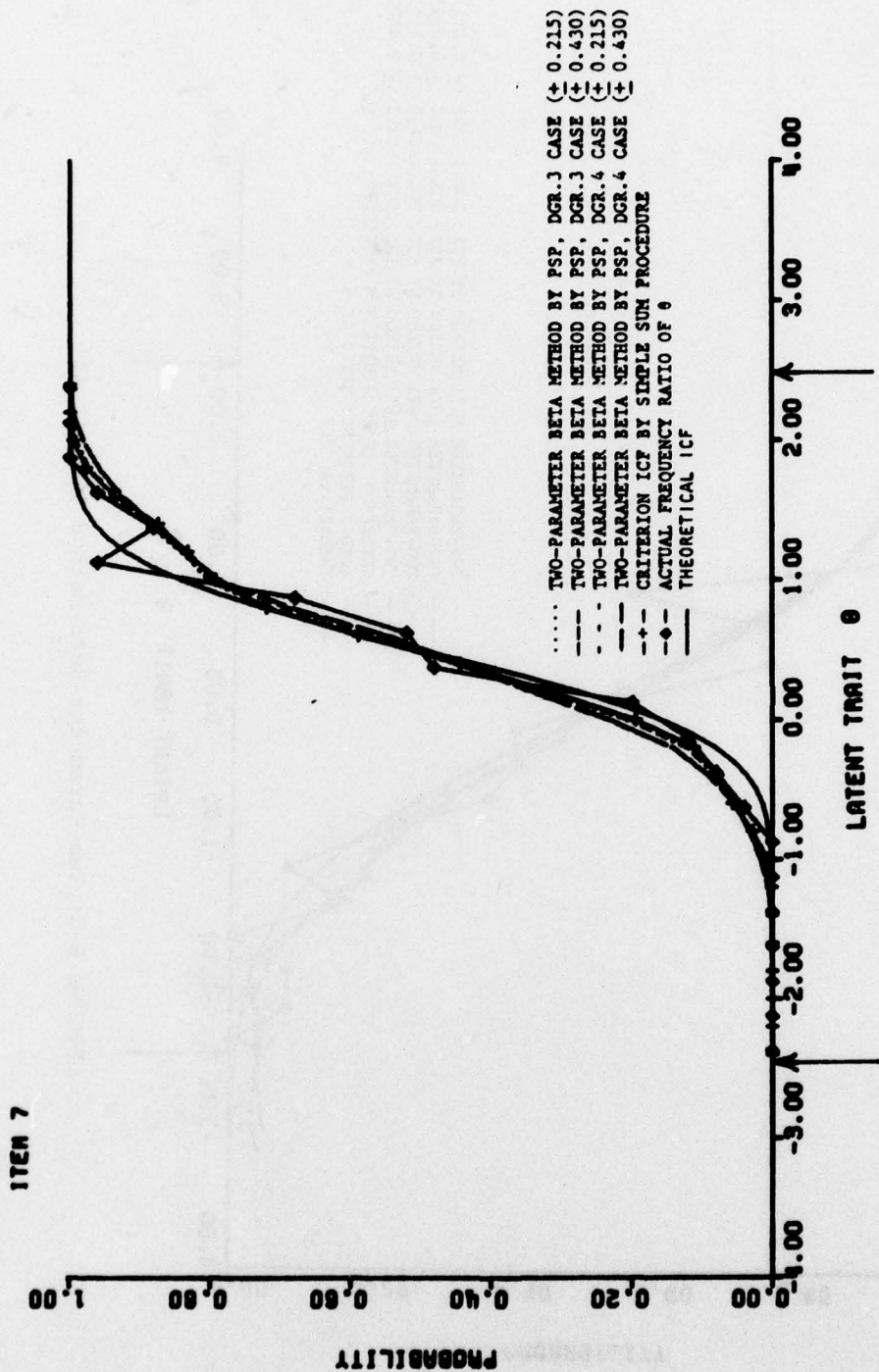


FIGURE 6-3: Two-Parameter Beta Method (Continued)

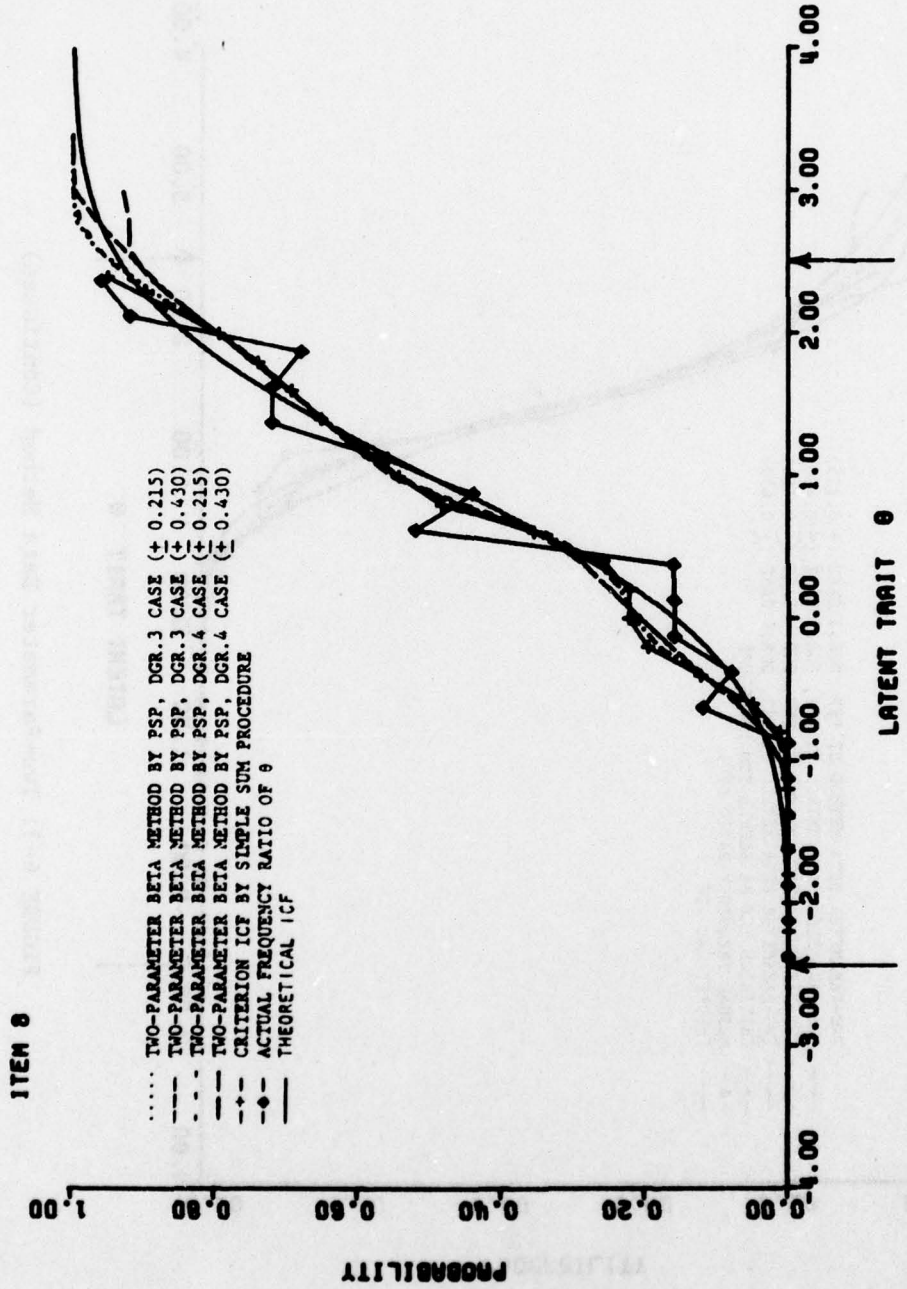


FIGURE 6-3: Two-Parameter Beta Method (Continued)

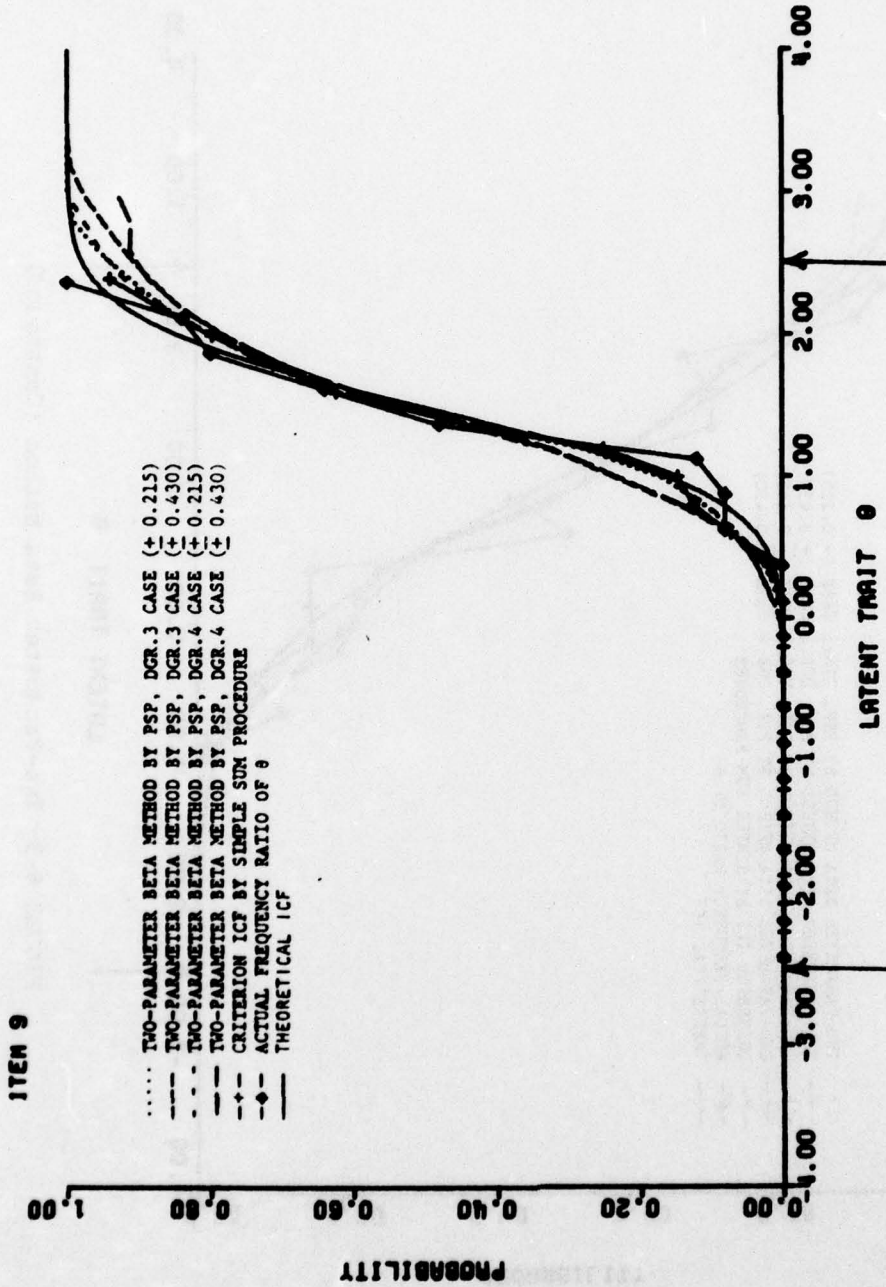


FIGURE 6-3: Two-Parameter Beta Method (Continued)

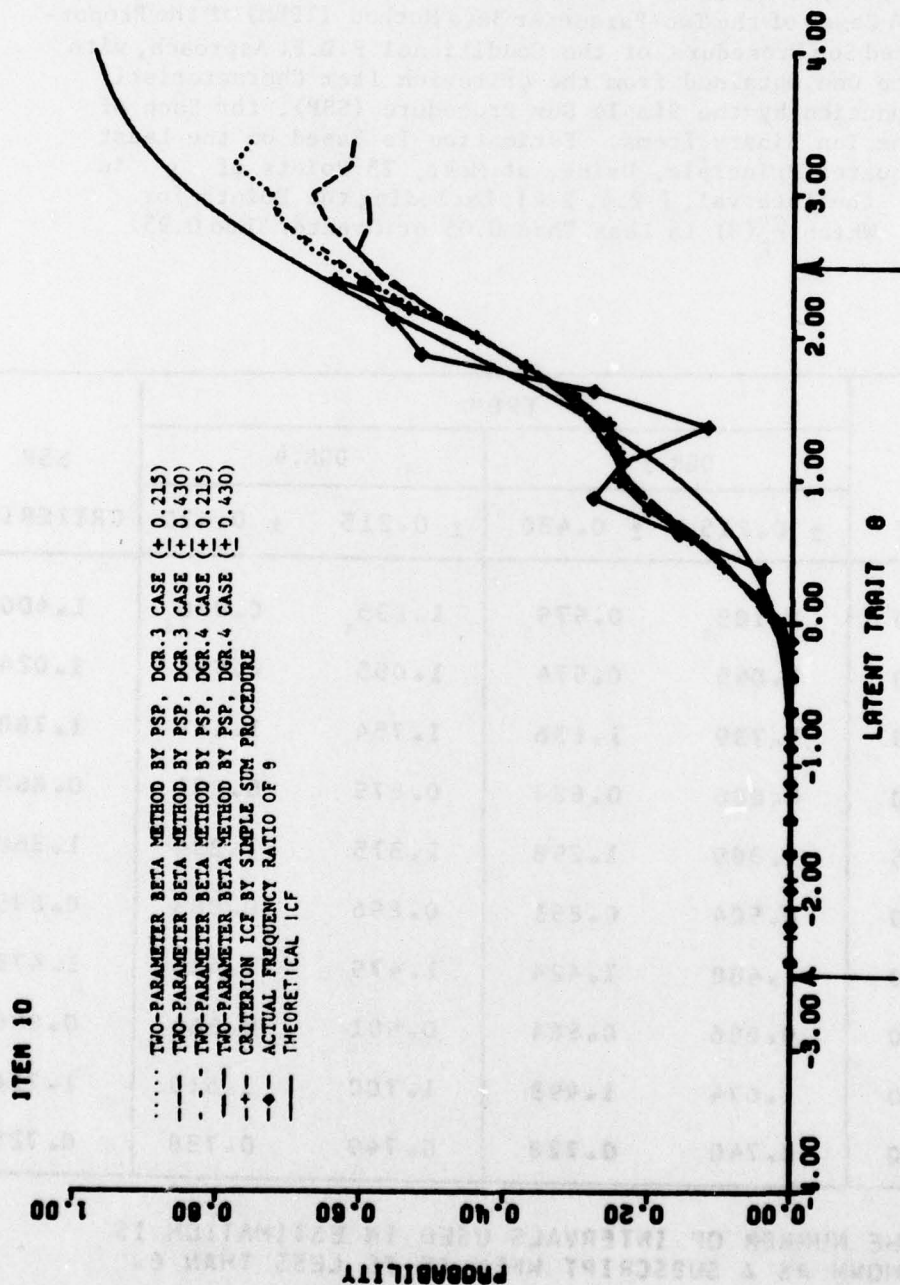


FIGURE 6-3: Two-Parameter Beta Method (Continued)

TABLE 6-7

Estimated Item Discrimination Parameters,  $\hat{a}_g$ , Obtained from the Four Estimated Item Characteristic Functions by Degree 3 and 4 Cases of the Two-Parameter Beta Method (TPBM) of the Proportioned Sum Procedure of the Conditional P.D.F. Approach, with the One Obtained from the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP), for Each of the Ten Binary Items. Estimation Is Based on the Least Squares Principle, Using, at Most, 25 Points of  $\theta$  in the Interval,  $[-2.4, 2.4]$ , Excluding the Points for Which  $\hat{P}_g(\theta)$  Is Less Than 0.05 or Greater Than 0.95.

ITEM	TRUE	TPBM				SSP CRITERION
		DGR.3		DGR.4		
		$\pm 0.215$	$\pm 0.430$	$\pm 0.215$	$\pm 0.430$	
1	1.5	1.105 <sub>5</sub>	0.979 <sub>5</sub>	1.135 <sub>5</sub>	0.980 <sub>5</sub>	1.400 <sub>5</sub>
2	1.0	1.045	0.974	1.055	0.981	1.024
3	2.5	1.739	1.636	1.754	1.644	1.788
4	1.0	0.886	0.827	0.879	0.821	0.868
5	1.5	1.389	1.298	1.375	1.284	1.368
6	1.0	0.904	0.891	0.896	0.883	0.895
7	2.0	1.488	1.424	1.479	1.415	1.473
8	1.0	0.896	0.864	0.901	0.866	0.886
9	2.0	1.674	1.493	1.700	1.510	1.716
10	1.0	0.740	0.728	0.749	0.738	0.725

THE NUMBER OF INTERVALS USED IN ESTIMATION IS SHOWN AS A SUBSCRIPT WHEN IT IS LESS THAN 6.

TABLE 6-8

Estimated Item Difficulty Parameters,  $\hat{b}_g$ , Obtained from the Four Estimated Item Characteristic Functions by Degree 3 and 4 Cases of the Two-Parameter Beta Method (TPBM) of the Proportioned Sum Procedure of the Conditional P.D.F. Approach, with the One Obtained from the Criterion Item Characteristic Function by the Simple Sum Procedure (SSP), for Each of the Ten Binary Items. Estimation Is Based on the Least Squares Principle, Using, at Most, 25 Points of  $\theta$  in the Interval,  $[-2.4, 2.4]$ , Excluding the Points for Which  $\hat{P}_g(\theta)$  Is Less Than 0.05 or Greater Than 0.95.

ITEM	TRUE	TPBM				SSP CRITERION
		DGR.3		DGR.4		
		$\pm 0.215$	$\pm 0.430$	$\pm 0.215$	$\pm 0.430$	
1	-2.5	-2.839 <sub>5</sub>	-2.931 <sub>5</sub>	-2.827 <sub>5</sub>	-2.938 <sub>5</sub>	-2.651 <sub>5</sub>
2	-2.0	-1.972	-2.015	-1.973	-2.019	-2.002
3	-1.5	-1.500	-1.516	-1.508	-1.524	-1.507
4	-1.0	-0.986	-1.023	-0.991	-1.026	-1.005
5	-0.5	-0.467	-0.463	-0.471	-0.466	-0.472
6	0.0	-0.061	-0.060	-0.062	-0.060	-0.075
7	0.5	0.518	0.518	0.523	0.522	0.527
8	1.0	0.985	0.982	0.981	0.982	0.981
9	1.5	1.520	1.509	1.523	1.514	1.502
10	2.0	2.138	2.171	2.130	2.162	2.118

THE NUMBER OF INTERVALS USED IN ESTIMATION IS SHOWN AS A SUBSCRIPT WHEN IT IS LESS THAN 6.



The mean square errors, and their square roots, of these estimated item characteristic functions from the true item characteristic functions, which are obtained through (6.1), are presented in Tables 6-1 and 6-2. Again these results are similar to those obtained by the Pearson System Method, and those from the criterion item characteristic functions by the Simple Sum Procedure and from those by the Proportioned Sum Procedure.

The estimated discrimination and difficulty parameters,  $\hat{a}_g$  and  $\hat{b}_g$ , are presented in Tables 6-7 and 6-8. They are comparable to those obtained by the Pearson System Method, and, again, there is a slight tendency that the estimates of the discrimination parameters are better for Degree 3 Case with the interval,  $\hat{\theta}_g \pm 0,215$ ,

173.5	110.7	73.1	173.5	110.7	73.1	1
100.2	71.8	47.2	100.2	71.8	47.2	2
78.1	53.5	35.1	78.1	53.5	35.1	3
56.1	37.2	24.0	56.1	37.2	24.0	4
34.0	23.1	15.0	34.0	23.1	15.0	5
12.0	8.0	5.0	12.0	8.0	5.0	6
0.0	0.0	0.0	0.0	0.0	0.0	7
0.0	0.0	0.0	0.0	0.0	0.0	8
0.0	0.0	0.0	0.0	0.0	0.0	9
0.0	0.0	0.0	0.0	0.0	0.0	10

THE WIDTH OF INTERVAL OF ESTIMATION OF  $\theta_g$  IS 0.215

## VII Discussion and Conclusion

Conditional P.D.F. Approach was expanded, and a new procedure, i.e., Proportioned Sum Procedure, was introduced, and tried on the simulated data, which have been used to test many different methods and approaches. As the result, we have obtained sufficiently good estimates of the item characteristic functions of the ten binary items. They are similar to the results obtained in previous studies, using Simple Sum Procedure and Weighted Sum Procedure, however, and we cannot say that the present method is superior to the others. Thus we ended up with expanding our repertory of methods of estimating the operating characteristics of the graded item categories.

One interesting finding is that, using as small a number of points of  $\hat{\theta}$  as 61, we have succeeded in producing criterion item characteristic functions which are practically the same as those obtained on the 500  $\hat{\theta}_s$ 's. This result is all the more interesting if we call our attention to the fact that neither of the numerator and the denominator of (3.3), which are divided by 61, is close to the density function of  $\theta$  and the shared density function of  $\theta$  by the success group, as we can see in Figures 5-1 and 5-2. Although it is too early to jump to the conclusion, if this is further tested with different types of data and proves to be successful, this will contribute to saving our time and trouble in the estimation process. In such a case, the set of a certain larger number of maximum likelihood estimates is used, mainly, for obtaining the estimates of the moments, and to specify a polynomial of a certain degree using the method of moments

to approximate the density function of the maximum likelihood,  $g(\hat{\theta})$ , for all the conditional moments of ability  $\theta$ , given  $\hat{\theta}$ , are estimated from  $g(\hat{\theta})$  and  $\sigma^2$  (cf. Appendix II).

As was the case in most of the previous studies, we have found no substantial differences between the set of results obtained by Degree 3 Case and that by Degree 4 Case, in spite of the fact that these two polynomials of degree 3 and 4 are substantially different (cf. Appendix II). It seems that the accuracy of estimation of the density function  $g(\hat{\theta})$  is not too important, although we should try not to generalize this tentative conclusion too far. Again, we find no significant differences between the results obtained by the Pearson System Method and the Two-Parameter Beta Method, and both of them proved to be very close to the criterion item characteristic functions.

REFERENCES

- [1] Elderton, W. P. and N. L. Johnson. Systems of frequency curves. Cambridge University Press, 1969.
- [2] Johnson, N. L. and S. Kotz. Continuous univariate distributions. Vol. 2. Houghton Mifflin, 1970.
- [3] Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, No. 17, 1969.
- [4] Samejima, F. A general model for free-response data. Psychometrika Monograph, No. 18, 1972.
- [5] Samejima, F. Graded response model of the latent trait theory and tailored testing. Proceedings of the First Conference on Computerized Adaptive Testing, 1975, Civil Service Commission and Office of Naval Research, 1975, pages 5-17.
- [6] Samejima, F. A use of the information function in tailored testing. Applied Psychological Measurement, 1, 1977a, pages 233-247.
- [7] Samejima, F. A method of estimating item characteristic functions using the maximum likelihood estimate of ability. Psychometrika, 42, 1977b, pages 163-191.
- [8] Samejima, F. Weakly parallel tests in latent trait theory with some criticisms of classical test theory. Psychometrika, 42, 1977c, pages 193-198.
- [9] Samejima, F. Estimation of the Operating Characteristics of item response categories I: Introduction to the Two-Parameter Beta Method. Office of Naval Research, Research Report 77-1, 1977d.
- [10] Samejima, F. Estimation of the operating characteristics of item response categories II: Further development of the Two-Parameter Beta Method. Office of Naval Research, Research Report 78-1, 1978a.
- [11] Samejima, F. Estimation of the operating characteristics of item response categories III: The Normal Approach Method and the Pearson System Method. Office of Naval Research, Research Report 78-2, 1978b.
- [12] Samejima, F. Estimation of the operating characteristics of item response categories IV: Comparison of the different methods. Office of Naval Research, Research Report 78-3, 1978c.

REFERENCES (Continued)

- [13] Samejima, F. Estimation of the Operating Characteristics of item response categories V: Weighted Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, Research Report 78-4, 1978d.

APPENDIX I

A-I Simulated Data

The simulated data used in the present study are characterized as follows.

- (1) There are 500 hypothetical examinees.
- (2) Their ability, or latent trait, distributes uniformly for the interval of  $\theta$ ,  $(-2.5, 2.5)$ . Actually, we use 100 discrete points of  $\theta$ , such as  $-2.475, -2.425, -2.375, -2.325, \dots, 2.375, 2.425$  and  $2.475$ , i.e., the midpoints of the 100 subintervals with the width of 0.05, and at each point five examinees are located.
- (3) There is a hypothetical test of 35 graded items, each of which has four item score categories, and which provides us with an approximately constant test information function, 21.63, for the interval of  $\theta$ ,  $[-3.0, 3.0]$ , following the normal ogive model of the graded response level (Samejima, 1969, 1972). The test is called the Old Test, to distinguish from the New Test, which will be described later.
- (4) Each of the 500 examinees is assumed to have taken the Old Test, and his response pattern on the 35 graded items has been calibrated by the Monte Carlo method. The score categories of each item are 0, 1, 2 and 3, and a typical response pattern looks like: (3,3,3,2,3,3,2,2,2,2,2,2,1,2,2,2,1,2,1,1,1,0,1,1,1,0,1,0,1,1,0,0,0,0,0).
- (5) From each response pattern, the maximum likelihood estimate of the examinee's ability has been obtained, using a computer program written for this purpose. In this process, out of 140 basic functions (Samejima, 1969, 1972), an appropriate set of 35 basic

functions are chosen depending upon the item scores in the response pattern, and, using the Newton-Raphson procedure, the point of  $\theta$  at which the sum total of these 35 basic functions equals zero is searched.

- (6) There is another hypothetical test of 10 binary items, each of which follows the normal ogive model of the dichotomous response level. This is called the New Test.
- (7) Each of the 500 examinees is assumed to have taken the New Test also, and his response pattern on the New Test has been calibrated by the Monte Carlo method. A typical response pattern looks like: (1,1,1,0,1,0,0,1,0,0).
- (8) The item characteristic functions of the test items of the New Test are assumed to be unknown, and they are the target of estimation. Each method of estimation is evaluated by the "closeness" of the resultant estimated item characteristic functions to the true item characteristic functions, i.e.,

$$P_g(\theta) = (2\pi)^{-1/2} \int_{-\infty}^{a_g(\theta-b_g)} e^{-\frac{t^2}{2}} dt .$$



4-11. Conditional density function of  $X$  given  $Y = y$  is given by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_Y(y)$  is the marginal density function of  $Y$ .

4-12. Conditional density function of  $Y$  given  $X = x$  is given by

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_X(x)$  is the marginal density function of  $X$ .

4-13. The conditional density function of  $X$  given  $Y = y$  is given by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_Y(y)$  is the marginal density function of  $Y$ .

4-14. The conditional density function of  $Y$  given  $X = x$  is given by

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_X(x)$  is the marginal density function of  $X$ .

APPENDIX II

4-15. The conditional density function of  $X$  given  $Y = y$  is given by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_Y(y)$  is the marginal density function of  $Y$ .

4-16. The conditional density function of  $Y$  given  $X = x$  is given by

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_X(x)$  is the marginal density function of  $X$ .

4-17. The conditional density function of  $X$  given  $Y = y$  is given by

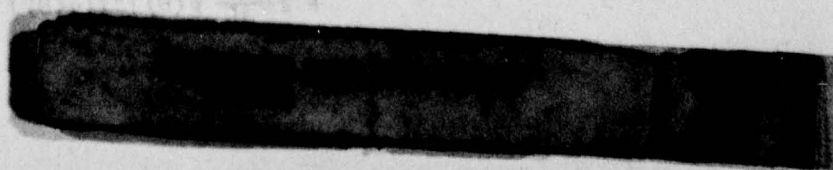
$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_Y(y)$  is the marginal density function of  $Y$ .

4-18. The conditional density function of  $Y$  given  $X = x$  is given by

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

where  $f_{XY}(x,y)$  is the joint density function of  $X$  and  $Y$  and  $f_X(x)$  is the marginal density function of  $X$ .



A-II Conditional Moments and Degree 3, 4 and 5 Cases

Let  $\phi(\theta|\hat{\theta})$  be the conditional density function of ability  $\theta$ , given its maximum likelihood estimate,  $\hat{\theta}$ . We can write

$$(A.2.1) \quad \begin{aligned} \phi(\theta|\hat{\theta}) &= \psi(\hat{\theta}|\theta) f(\theta) \left[ \int_{-\infty}^{\infty} \psi(\hat{\theta}|\theta) f(\theta) d\theta \right]^{-1} \\ &= \psi(\hat{\theta}|\theta) f(\theta) [g(\hat{\theta})]^{-1}, \end{aligned}$$

where  $\psi(\hat{\theta}|\theta)$  is the conditional density function of  $\hat{\theta}$ , given  $\theta$ ,  $f(\theta)$  is the probability density function of ability  $\theta$ , and  $g(\hat{\theta})$  is the probability density function of the maximum likelihood estimate  $\hat{\theta}$ .

By virtue of the asymptotic property of the maximum likelihood estimate and the constancy of the test information function of the Old Test, the conditional density  $\psi(\hat{\theta}|\theta)$  is approximated by the normal density  $n(\theta, \sigma^2)$ , where  $\sigma^2 = (21.63)^{-1} \doteq 0.046225$ . Because of this fact, although the density function  $\phi(\theta|\hat{\theta})$  given by (A.2.1) is not observable in the empirical situation, it is possible to estimate the conditional moments of ability  $\theta$ , given the maximum likelihood estimate  $\hat{\theta}$ , provided that the density function  $g(\hat{\theta})$  is estimated. That is to say, we can derive the following equations.

$$(A.2.2) \quad E(\theta|\hat{\theta}) \doteq \hat{\theta} + \sigma^2 \cdot \frac{d}{d\hat{\theta}} \log g(\hat{\theta}) = \hat{\theta} + \sigma^2 \left[ \frac{d}{d\hat{\theta}} g(\hat{\theta}) \right] [g(\hat{\theta})]^{-1}.$$

$$(A.2.3) \quad \begin{aligned} \text{Var.}(\theta|\hat{\theta}) &\doteq \sigma^2 \left[ 1 + \sigma^2 \frac{d^2}{d\hat{\theta}^2} \log g(\hat{\theta}) \right] \\ &= \sigma^2 \left[ 1 + \sigma^2 \left\{ \frac{d^2}{d\hat{\theta}^2} g(\hat{\theta}) \cdot g(\hat{\theta}) - \left[ \frac{d}{d\hat{\theta}} g(\hat{\theta}) \right]^2 \right\} \{g(\hat{\theta})\}^{-2} \right]. \end{aligned}$$

$$(A.2.4) \quad \begin{aligned} E\{(\theta - E(\theta|\hat{\theta}))^3|\hat{\theta}\} &\doteq -\sigma^6 \left[ \frac{d^3}{d\hat{\theta}^3} \log g(\hat{\theta}) \right] \\ &= -\sigma^6 \left[ \{g(\hat{\theta})\}^2 \cdot \frac{d^3}{d\hat{\theta}^3} g(\hat{\theta}) - 3 g(\hat{\theta}) \cdot \frac{d}{d\hat{\theta}} g(\hat{\theta}) \cdot \frac{d^2}{d\hat{\theta}^2} g(\hat{\theta}) \right. \\ &\quad \left. + 2 \left\{ \frac{d}{d\hat{\theta}} g(\hat{\theta}) \right\}^3 \right] \{g(\hat{\theta})\}^{-3}. \end{aligned}$$

$$\begin{aligned}
 \text{(A.2.5)} \quad E\{(\theta - E(\theta|\hat{\theta}))^4|\hat{\theta}\} &= \sigma^4\{3 + 6\sigma^2\{-\frac{d^2}{d\theta^2} \log g(\hat{\theta})\} + 3\sigma^4\{-\frac{d^2}{d\theta^2} \log g(\hat{\theta})\}^2 \\
 &\quad + \sigma^4\{-\frac{d^4}{d\theta^4} \log g(\hat{\theta})\}\} \\
 &= \sigma^4\{3 + 6\sigma^2\{[g(\hat{\theta}) \cdot \frac{d^2}{d\theta^2} g(\hat{\theta}) - \{\frac{d}{d\theta} g(\hat{\theta})\}^2][g(\hat{\theta})]^{-2}\} \\
 &\quad + 3\sigma^4 [g(\hat{\theta}) \cdot \frac{d^2}{d\theta^2} g(\hat{\theta}) - \{\frac{d}{d\theta} g(\hat{\theta})\}^2]^2 [g(\hat{\theta})]^{-1} \\
 &\quad + \sigma^4\{[g(\hat{\theta})]^3 \cdot \frac{d^4}{d\theta^4} g(\hat{\theta}) - 4[g(\hat{\theta})]^2 \cdot \frac{d}{d\theta} g(\hat{\theta}) \cdot \frac{d^3}{d\theta^3} g(\hat{\theta}) \\
 &\quad - 3[g(\hat{\theta})]^2 \{\frac{d^2}{d\theta^2} g(\hat{\theta})\}^2 + 12 g(\hat{\theta}) \{\frac{d}{d\theta} g(\hat{\theta})\}^2 \frac{d^2}{d\theta^2} g(\hat{\theta}) \\
 &\quad - 6\{\frac{d}{d\theta} g(\hat{\theta})\}^4 [g(\hat{\theta})]^{-4}\}.
 \end{aligned}$$

Thus it is obvious that we need an estimate of  $g(\hat{\theta})$  in order to obtain the above four conditional moments. This is done by approximating  $g(\hat{\theta})$  by a polynomial of degree 3, 4 or 5 using the method of moments (Elderton and Johnson, 1969, Johnson and Kotz, 1970). These three polynomials obtained for the five hundred maximum likelihood estimates, which have been used as our simulated data, are as follows.

$$\text{(A.2.6)} \quad g(\hat{\theta}) = 0.22416 - 0.00351\hat{\theta} - 0.01873\hat{\theta}^2 + 0.00095\hat{\theta}^3$$

$$\text{(A.2.7)} \quad g(\hat{\theta}) = 0.19620 + 0.00238\hat{\theta} + 0.01319\hat{\theta}^2 - 0.00062\hat{\theta}^3 - 0.00427\hat{\theta}^4$$

$$\text{(A.2.8)} \quad g(\hat{\theta}) = 0.19539 - 0.00638\hat{\theta} + 0.01449\hat{\theta}^2 + 0.00405\hat{\theta}^3 - 0.00449\hat{\theta}^4 \\ - 0.00048\hat{\theta}^5$$

We distinguish three different cases, depending upon the polynomial used for  $g(\hat{\theta})$ , and call them Degree 3, 4 and 5 Cases. Since the polynomial of degree 5, shown as (A.2.8), is close to that of degree 4, in many cases we deal only with Degree 3 and 4 Cases.

DISTRIBUTION LIST

Navy	Navy
1 Dr. Ed Aiken Navy Personnel R&D Center San Diego, CA 92152	1 CHAIRMAN, LEADERSHIP & LAW DEPT. DIV. OF PROFESSIONAL DEVELOPMENT U.S. NAVAL ACADEMY ANNAPOLIS, MD 21402
1 Dr. Jack R. Borsting Provost & Academic Dean U.S. Naval Postgraduate School Monterey, CA 93940	1 Dr. James McEride Code 301 Navy Personnel R&D Center San Diego, CA 92152
1 Dr. Richard Elster Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	1 DR. WILLIAM MONTAGUE LRDC UNIVERSITY OF PITTSBURGH 3939 O'HARA STREET PITTSBURGH, PA 15213
1 DR. PAT FEDERICO NAVY PERSONNEL R&D CENTER SAN DIEGO, CA 92152	1 Commanding Officer Naval Health Research Center Attn: Library San Diego, CA 92152
1 CDR John Ferguson, MSC, USN Naval Medical R&D Command (Code 44) National Naval Medical Center Bethesda, MD 20014	1 Naval Medical R&D Command Code 44 National Naval Medical Center Bethesda, MD 20014
1 Dr. Paul Foley Navy Personnel R&D Center San Diego, CA 92152	1 CAPT Paul Nelson, USN Chief, Medical Service Corps Code 7 Bureau of Medicine & Surgery U. S. Department of the Navy Washington, DC 20372
1 Dr. Steve Harris Code L522 NAMRL Pensacola FL 32508	1 DR. RICHARD J. NIEHAUS CODE 301 OFFICE OF CIVILIAN PERSONNEL NAVY DEPT WASHINGTON, DC 20390
1 CDR Robert S. Kennedy Naval Aerospace Medical and Research Lab Box 29407 New Orleans, LA 70189	1 Library Navy Personnel R&D Center San Diego, CA 92152
1 Dr. Norman J. Kerr Chief of Naval Technical Training Naval Air Station Memphis (75) Millington, TN 38054	5 Technical Director Navy Personnel R&D Center San Diego, CA 92152
1 Dr. Leonard Kroeker Navy Personnel R&D Center San Diego, CA 92152	

Navy

- 6 Commanding Officer  
Naval Research Laboratory  
Code 2627  
Washington, DC 20390
- 1 OFFICE OF CIVILIAN PERSONNEL  
(CODE 26)  
DEPT. OF THE NAVY  
WASHINGTON, DC 20390
- 1 JOHN OLSEN  
CHIEF OF NAVAL EDUCATION &  
TRAINING SUPPORT  
PENSACOLA, FL 32509
- 1 Psychologist  
ONR Eranch Office  
495 Summer Street  
Boston, MA 02210
- 1 Psychologist  
ONR Eranch Office  
536 S. Clark Street  
Chicago, IL 60605
- 1 Office of Naval Research  
Code 200  
Arlington, VA 22217
- 1 Code 436  
Office of Naval Research  
Arlington, VA 22217
- 5 Personnel & Training Research Programs  
(Code 458)  
Office of Naval Research  
Arlington, VA 22217
- 1 Psychologist  
OFFICE OF NAVAL RESEARCH BRANCH  
223 OLD MARYLEBONE ROAD  
LONDON, NW, 15TH ENGLAND
- 1 Psychologist  
ONR Branch Office  
1030 East Green Street  
Pasadena, CA 91101

Navy

- 1 Scientific Director  
Office of Naval Research  
Scientific Liaison Group/Tokyo  
American Embassy  
APO San Francisco, CA 96503
- 1 Head, Research, Development, and Studies  
(OP102X)  
Office of the Chief of Naval Operations  
Washington, DC 20370
- 1 Scientific Advisor to the Chief of  
Naval Personnel (Pers-Or)  
Naval Bureau of Personnel  
Room 4410, Arlington Annex  
Washington, DC 20370
- 1 DR. RICHARD A. POLLAK  
ACADEMIC COMPUTING CENTER  
U.S. NAVAL ACADEMY  
ANNAPOLIS, MD 21402
- 1 Mr. Arnold Rubenstein  
Naval Personnel Support Technology  
Naval Material Command (08T244)  
Room 1044, Crystal Plaza #5  
2221 Jefferson Davis Highway  
Arlington, VA 20360
- 1 A. A. SJOHOLM  
TECH. SUPPORT, CODE 201  
NAVY PERSONNEL R & D CENTER  
SAN DIEGO, CA 92152
- 1 Mr. Robert Smith  
Office of Chief of Naval Operations  
OP-987E  
Washington, DC 20350
- 1 Dr. Alfred F. Smode  
Training Analysis & Evaluation Group  
(TAEG)  
Dept. of the Navy  
Orlando, FL 32813
- 1 Dr. Richard Sorensen  
Navy Personnel R&D Center  
San Diego, CA 92152

Navy

Army

- 1 CDR Charles J. Theisen, JR. MSC, USN  
Head Human Factors Engineering Div.  
Naval Air Development Center  
Warminster, PA 18974
- 1 W. Gary Thomson  
Naval Ocean Systems Center  
Code 7132  
San Diego, CA 92152
- 1 Dr. Ronald Weitzman  
Department of Administrative Sciences  
U. S. Naval Postgraduate School  
Monterey, CA 93940
- 1 DR. MARTIN F. WISKOFF  
NAVY PERSONNEL R & D CENTER  
SAN DIEGO, CA 92152

- 1 Technical Director  
U. S. Army Research Institute for the  
Behavioral and Social Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 HQ USAREUE & 7th Army  
ODCSOPS  
USAAREUE Director of GED  
APO New York 09403
- 1 DR. RALPH CANTER  
U.S. ARMY RESEARCH INSTITUTE  
5001 EISENHOWER AVENUE  
ALEXANDRIA, VA 22333
- 1 Dr. Milton S. Katz  
Individual Training & Skill  
Evaluation Technical Area  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 Dr. Frederick Steinheiser  
U. S. Army Reserch Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 Dr. Joseph Ward  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Air Force

- 1 Air Force Human Resources Lab  
AFHRL/PED  
Brooks AFB, TX 78235
- 1 Air University Library  
AUL/LSE 76/443  
Maxwell AFB, AL 36112
- 1 Dr. Philip De Leo  
AFHRL/TT  
Lowry AFB, CO 80230
- 1 CDR. MERCER  
CNET LIAISON OFFICER  
AFHRL/FLYING TRAINING DIV.  
WILLIAMS AFB, AZ 85224
- 1 Dr. Roger Pennell  
AFHRL/TT  
Lowry AFB, CO 80230
- 1 Personnel Analysis Division  
HQ USAF/DPXXA  
Washington, DC 20330
- 1 Research Branch  
AFMPC/DPMYP  
Randolph AFB, TX 78148
- 1 Dr. Malcolm Ree  
AFHRL/PED  
Brooks AFB, TX 78235
- 1 Dr. Marty Rockway (AFHRL/TT)  
Lowry AFB  
Colorado 80230
- 1 Jack A. Thorpe, Capt, USAF  
Program Manager  
Life Sciences Directorate  
AFOSR  
Polling AFB, DC 20332
- 1 Brian K. Waters, LCOL, USAF  
Air University  
Maxwell AFB  
Montgomery, AL 36112

Marines

- 1 Director, Office of Manpower Utilization  
HQ, Marine Corps (MPU)  
BCB, Bldg. 2009  
Quantico, VA 22134
- 1 MCDEC  
Quantico Marine Corps Base  
Quantico, VA 22134
- 1 DR. A.L. SLAFKOSKY  
SCIENTIFIC ADVISOR (CODE RD-1)  
HQ, U.S. MARINE CORPS  
WASHINGTON, DC 20380

CoastGuard

Other DoD

1 MR. JOSEPH J. COWAN, CHIEF  
PSYCHOLOGICAL RESEARCH (G-P-1/62)  
U.S. COAST GUARD HQ  
WASHINGTON, DC 20590

1 Dr. Thomas Warm  
U. S. Coast Guard Institute  
P. O. Substation 18  
Oklahoma City, OK 73169

1 Dr. Stephen Andriole  
ADVANCED RESEARCH PROJECTS AGENCY  
1400 WILSON BLVD.  
ARLINGTON, VA 22209

12 Defense Documentation Center  
Cameron Station, Bldg. 5  
Alexandria, VA 22314  
Attn: TC

1 Dr. Dexter Fletcher  
ADVANCED RESEARCH PROJECTS AGENCY  
1400 WILSON BLVD.  
ARLINGTON, VA 22209

1 Dr. William Graham  
Testing Directorate  
MEPCOM  
Ft. Sheridan, IL 60037

1 Military Assistant for Training and  
Personnel Technology  
Office of the Under Secretary of Defense  
for Research & Engineering  
Room 3D129, The Pentagon  
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF  
Office of the Assistant Secretary  
of Defense (MRA&L)  
3B930 The Pentagon  
Washington, DC 20301



Civil Govt

- 1 Dr. Susan Chipman  
Basic Skills Program  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. William Gorham, Director  
Personnel R&D Center  
U.S. Civil Service Commission  
1900 E Street NW  
Washington, DC 20415
- 1 Dr. John Mays  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Arthur Melmud  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Andrew R. Molnar  
Science Education Dev.  
and Research  
National Science Foundation  
Washington, DC 20550
- 1 Dr. Lalitha P. Sanathanan  
Environmental Impact Studies Division  
Argonne National Laboratory  
9700 S. Cass Avenue  
Argonne, IL 60439
- 1 Dr. Jeffrey Schiller  
National Institute of Education  
1200 19th St. NW  
Washington, DC 20208
- 1 Dr. H. Wallace Sinaiko  
Program Director  
Manpower Research and Advisory Services  
Smithsonian Institution  
801 North Pitt Street  
Alexandria, VA 22314

Civil Govt

- 1 Dr. Thomas G. Sticht  
Basic Skills Program  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Vern W. Urry  
Personnel R&D Center  
U.S. Civil Service Commission  
1900 E Street NW  
Washington, DC 20415
- 1 Dr. Joseph L. Young, Director  
Memory & Cognitive Processes  
National Science Foundation  
Washington, DC 20550

Non Govt

- 1 1 psychological research unit  
Dept. of Defense (Army Office)  
Campbell Park Offices  
Canberra ACT 2600, Australia
- 1 Dr. Alan Baddeley  
Medical Research Council  
Applied Psychology Unit  
15 Chaucer Road  
Cambridge CB2 2EF  
ENGLAND
- 1 Dr. Nicholas A. Bond  
Dept. of Psychology  
Sacramento State College  
600 Jay Street  
Sacramento, CA 95819
- 1 Dr. David G. Eowers  
Institute for Social Research  
University of Michigan  
Ann Arbor, MI 48106
- 1 Dr. John E. Carroll  
Psychometric Lab  
Univ. of No. Carolina  
Davie Hall 013A  
Chapel Hill, NC 27514
- 1 Charles Myers Library  
Livingstone House  
Livingstone Road  
Stratford  
London E15 2LJ  
ENGLAND
- 1 Dr. Norman Cliff  
Dept. of Psychology  
Univ. of So. California  
University Park  
Los Angeles, CA 90007
- 1 Dr. William Coffman  
Iowa Testing Programs  
University of Iowa  
Iowa City, IA 52242

Non Govt

- 1 Dr. Meredith Crawford  
Department of Engineering Administration  
George Washington University  
Suite 805  
2101 L Street N. W.  
Washington, DC 20037
- 1 Dr. Hans Cronbag  
Education Research Center  
University of Leyden  
Boerhaavelaan 2  
Leyden  
The NETHERLANDS
- 1 MAJOR I. N. EVONIC  
CANADIAN FORCES PERS. APPLIED RESEARCH  
1107 AVENUE ROAD  
TORONTO, ONTARIO, CANADA
- 1 Dr. Richard L. Ferguson  
The American College Testing Program  
P.O. Box 168  
Iowa City, IA 52240
- 1 Dr. Victor Fields  
Dept. of Psychology  
Montgomery College  
Rockville, MD 20850
- 1 Dr. Edwin A. Fleishman  
Advanced Research Resources Organ.  
8555 Sixteenth Street  
Silver Spring, MD 20910
- 1 Dr. John R. Frederiksen  
Bolt Beranek & Newman  
50 Moulton Street  
Cambridge, MA 02138
- 1 DR. ROBERT GLASER  
LRDC  
UNIVERSITY OF PITTSBURGH  
3939 O'HARA STREET  
PITTSBURGH, PA 15213
- 1 Dr. Ron Hambleton  
School of Education  
University of Massachusetts  
Amherst, MA 01002

Non Govt

- 1 Library  
HumRRO/Western Division  
27857 Berwick Drive  
Carmel, CA 93921
- 1 Dr. Earl Hunt  
Dept. of Psychology  
University of Washington  
Seattle, WA 98105
- 1 DR. LAWRENCE B. JOHNSON  
LAWRENCE JOHNSON & ASSOC., INC.  
SUITE 502  
2001 S STREET NW  
WASHINGTON, DC 20009
- 1 Dr. Roger A. Kaufman  
203 Dodd Hall  
Florida State Univ.  
Tallahassee, FL 32306
- 1 Dr. Steven W. Keele  
Dept. of Psychology  
University of Oregon  
Eugene, OR 97403
- 1 LCOL. C.R.J. LAFLEUR  
PERSONNEL APPLIED RESEARCH  
NATIONAL DEFENSE HQS  
101 COLONEL BY DRIVE  
OTTAWA, CANADA K1A 0K2
- 1 Dr. Michael Levine  
Department of Psychology  
University of Illinois  
Champaign, IL 61820
- 1 Dr. Robert Linn  
College of Education  
University of Illinois  
Urbana, IL 61801
- 1 Dr. Frederick M. Lord  
Educational Testing Service  
Princeton, NJ 08540

Non Govt

- 1 Dr. Robert R. Mackie  
Human Factors Research, Inc.  
6780 Cortona Drive  
Santa Barbara Research Pk.  
Goleta, CA 93017
- 1 Dr. Scott Maxwell  
Department of Psychology  
University of Houston  
Houston, TX 77025
- 1 Dr. Sam Mayo  
Loyola University of Chicago  
Chicago, IL 60601
- 1 Richard T. Mowday  
College of Business Administration  
University of Oregon  
Eugene, OR 97403
- 1 Dr. Melvin R. Novick  
Iowa Testing Programs  
University of Iowa  
Iowa City, IA 52242
- 1 Dr. Robert Pachella  
Department of Psychology  
Human Performance Center  
330 Packard Road  
Ann Arbor, MI 48104
- 1 Dr. James A. Paulson  
Portland State University  
P.O. Box 751  
Portland, OR 97207
- 1 MR. LUIGI PETRULLO  
2431 N. EDGEWOOD STREET  
ARLINGTON, VA 22207
- 1 DR. STEVEN M. PINE  
4950 Douglas Avenue  
Golden Valley, MN 55416
- 1 DR. DIANE M. RAMSEY-KLEE  
R-K RESEARCH & SYSTEM DESIGN  
3947 RIDGEMONT DRIVE  
MALIBU, CA 90265

Non Govt

- 1 MIN. RET. M. RAUCH  
P II 4  
BUNDESMINISTERIUM DER VERTEIDIGUNG  
POSTFACH 161  
53 BONN 1, GERMANY
- 1 Dr. Mark U. Reckase  
Educational Psychology Dept.  
University of Missouri-Columbia  
12 Hill Hall  
Columbia, MO 65201
- 1 Dr. Andrew M. Rose  
American Institutes for Research  
1055 Thomas Jefferson St. NW  
Washington, DC 20007
- 1 DR. WALTER SCHNEIDER  
DEPT. OF PSYCHOLOGY  
UNIVERSITY OF ILLINOIS  
CHAMPAIGN, IL 61820
- 1 Dr. Edwin Shirkey  
Department of Psychology  
Florida Technological University  
Orlando, FL 32816
- 1 Dr. Richard Snow  
School of Education  
Stanford University  
Stanford, CA 94305
- 1 Dr. Robert Sternberg  
Dept. of Psychology  
Yale University  
Box 11A, Yale Station  
New Haven, CT 06520
- 1 DR. PATRICK SUPPES  
INSTITUTE FOR MATHEMATICAL STUDIES IN  
THE SOCIAL SCIENCES  
STANFORD UNIVERSITY  
STANFORD, CA 94305
- 1 Dr. Brad Sympson  
N660 Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455

Non Govt

- 1 Dr. Hariharan Swaminathan  
Laboratory of Psychometric and  
Evaluation Research  
School of Education  
University of Massachusetts  
Amherst, MA 01003
- 1 Dr. Kikumi Tatsuoka  
Computer Eased Education Research  
Laboratory  
252 Engineering Research Laboratory  
University of Illinois  
Urbana, IL 61801
- 1 Dr. Benton J. Underwood  
Dept. of Psychology  
Northwestern University  
Evanston, IL 60201
- 1 DR. THOMAS WALLSTEN  
PSYCHOMETRIC LABORATORY  
DAVIE HALL 013A  
UNIVERSITY OF NORTH CAROL  
CHAPEL HILL, NC 27514
- 1 Dr. John Wannous  
Department of Management  
Michigan University  
East Lansing, MI 48824
- 1 Dr. David J. Weiss  
N660 Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455
- 1 DR. SUSAN E. WHITELY  
PSYCHOLOGY DEPARTMENT  
UNIVERSITY OF KANSAS  
LAWRENCE, KANSAS 66044

**Navy**

- 1 Donald Calder  
Office of Naval Research  
325 Hinman Research Building  
Atlanta, GA 30332

**Army**

- 1 Dr. Randall M. Chambers, Ph.D.  
U.S. Army Research Institute  
for the Behavioral and Social  
Sciences  
Fort Sill Field Unit  
P.O.Box 3066  
Fort Sill, OK 73503

**Non Govt**

- 1 Dr. Ron Hambleton  
School of Education  
University of Massachusetts  
Amherst, Mass. 01002
- 1 Dr. William E. Coffman  
Iowa Testing Programs  
334 Lindquist Center  
Iowa City, IA 52242
- 1 Mr. Isaac I. Bejar  
Department of Psychology  
Elliott Hall  
75 East River Road  
Minneapolis, Minnesota 55455
- 1 Mr. George Woods  
1106 Newport Ave.  
Victoria, B.C.  
V8S 5E4 Canada
- 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 Enschede  
The Netherlands