ACT Research Report Series 91-2

Empirical Estimation of Standard Errors of Compensatory MIRT Model Parameters Obtained from the NOHARM Estimation Program

Research Report ONR91-2

Timothy R. Miller

Prepared under Contract No. N00014-89-J-1908, Contract Authority Identification No. 4421556-02, with the Cognitive Science Research Programs Office of Naval Research.

August 1991

For additional copies write:
ACT Research Report Series
P.O. Box 168
Iowa City, Iowa 52243

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b RESTRICTIVE MARKINGS

2a SECURITY CLASSIFICATION AUTHORITY

2b DECLASSIFICATION / DOWNGRADING SCHEDULE

3 DISTRIBUTION/AVAILABILITY OF REPORT
Approved for public release: distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the U.S. Govt.

4 PERFORMING ORGANIZATION REPORT NUMBER(S)
ONR 91-2

5 MONITORING ORGANIZATION REPORT NUMBER(S)

6a NAME OF PERFORMING ORGANIZATION
ACT

6b OFFICE SYMBOL (If applicable)

7a NAME OF MONITORING ORGANIZATION
COGNITIVE SCIENCE RESEARCH PROGRAMS
OFFICE OF NAVAL RESEARCH

6c ADDRESS (City, State, and ZIP Code)
P.O. Box 168
Iowa City, IA 52243

7b ADDRESS (City, State, and ZIP Code)
Code 1142CS
Arlington, VA 22217-5000

8a. NAME OF FUNDING/SPONSORING ORGANIZATION

8b OFFICE SYMBOL (If applicable)

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-89-J-1908

8c ADDRESS (City, State, and ZIP Code)

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153N	RR04204	RR0420401	4421556---02

11 TITLE (Include Security Classification)
Empirical Estimation of Standard Errors of Compensatory MIRT Model Parameters Obtained from the NOHARM Estimation Program

12 PERSONAL AUTHOR(S)
Timothy R. Miller

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM ____ TO ____	1991, August	27

16 SUPPLEMENTARY NOTATION

17 COSATI CODES

FIELD	GROUP	SUB-GROUP
05	09	

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Multidimensional Item Response Theory, Parameter Estimation

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Two studies were carried out to evaluate the quality of multidimensional item response theory (MIRT) model parameter estimates obtained from the computer program NOHARM. The purpose of the first study was to compute empirical estimates of the standard errors of the parameters. In addition, the parameter estimates were evaluated for bias and the effects of using different starting values and anchor items. The second study was included to compare the performance of NOHARM with the findings of an earlier simulation study which evaluated other MIRT estimation programs. Results were generally good, with fairly small standard errors for most parameter estimates and little indication of bias. Although the estimation procedure appeared to be robust under different starting values, the specific choice of items used to anchor the solution appears to have important effects on the magnitude of the estimated standard errors. The comparison of NOHARM with other

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL
Dr. Charles Davis

22b TELEPHONE (Include Area Code)
(703) 696-4046

22c OFFICE SYMBOL
ONR 1142CS

DD Form 1473, JUN 86 Previous editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

S/N 0102-LF-014-6603

19. (cont.)

programs was very favorable and supports the use of NOHARM for practical MIRT applications.

Abstract

Two studies were carried out to evaluate the quality of multidimensional item response theory (MIRT) model parameter estimates obtained from the computer program NOHARM. The purpose of the first study was to compute empirical estimates of the standard errors of the parameters. In addition, the parameter estimates were evaluated for bias and the effects of using different starting values and anchor items. The second study was included to compare the performance of NOHARM with the findings of an earlier simulation study which evaluated other MIRT estimation programs. Results were generally good, with fairly small standard errors for most parameter estimates and little indication of bias. Although the estimation procedure appeared to be robust under different starting values, the specific choice of items used to anchor the solution appears to have important effects on the magnitude of the estimated standard errors. The comparison of NOHARM with other programs was very favorable and supports the use of NOHARM for practical MIRT applications.

Empirical Estimation of Standard Errors of Compensatory MIRT Model Parameters Obtained from the NOHARM Estimation Program

Introduction

The practical utility of multidimensional item response theory (MIRT) depends upon the ability to obtain reasonably accurate parameter estimates. Several estimation programs are currently available, including MIRTE (Carlson, 1987) and MULTIDIM (McKinley, 1987) which were developed specifically as MIRT programs, TESTFACT (Wilson, Wood and Gibbons, 1984) which is a full information item factor analysis program that can be used to obtain MIRT parameter estimates, and NOHARM (Fraser, 1986) a general program for fitting unidimensional and multidimensional normal ogive models by a least squares procedure. An earlier simulation study (Ackerman, 1988) compared MIRTE, MULTIDIM and TESTFACT along several criteria and found MULTIDIM and TESTFACT to be far superior to MIRTE, with TESTFACT performing the best overall under the conditions of that study.

In this study, NOHARM is evaluated for its accuracy and usefulness as a MIRT program. The main question is whether the estimates provided by NOHARM are sufficiently accurate for practical applications. Since NOHARM employs a least squares procedure, standard errors are not directly available and must be established empirically. The purpose of this study is to estimate, through approximation of the sampling distribution by repeated sampling, the standard errors of the parameter estimates provided by NOHARM.

In addition to estimating standard errors, this research will evaluate the estimates for bias and the effects of using different starting values and different anchor items to fix the solution. Finally, the performance of NOHARM is compared with the other programs mentioned above. The assessments of standard errors, bias, and robustness will involve analyses of real datasets. The comparison with other programs will be accomplished through a simulation identical to that used by Ackerman (1988).

The NOHARM Model and Procedures

NOHARM (Normal Ogive Harmonic Analysis Robust Method) is a program for fitting unidimensional and multidimensional normal ogive item response models. The generalized multidimensional normal ogive model is given as

$$P(y_{ij}=1|\boldsymbol{\theta}_j)=c_i+(1-c_i)\Phi[d_i+\boldsymbol{a}_i\boldsymbol{\theta}_j], \tag{1}$$

where $P(x_{ij}=1|a_i, d_i, \theta_j)$ is the probability in an m-dimensional space of a correct response to item i by person j, $\mathbf{a}_i$ is an m-dimensional vector of item discrimination parameters, d_i is a scalar parameter related to item difficulty, $\boldsymbol{\theta}_j$ is an m-dimensional vector of latent abilities, c_i is a pseudo-guessing parameter, and Φ is the normal distribution function.

The model is fit by an ordinary least squares procedure which seeks to minimize the squared differences between the sample and estimated bivariate proportions correct. A four term polynomial series is used to approximate the model given by equation (1), and the estimated bivariate proportions correct are derived from this approximation, allowing the minimization with respect to the model parameters d, **a**, and $\boldsymbol{\Sigma}_\theta$. The vector c is not estimated but is treated as fixed. The function to be minimized is a least squares function and is minimized using a conjugate gradients minimization algorithm.

To run the program, the vector c must be supplied by the user. This can be a null vector, in which case a multidimensional extension of the two-parameter model is invoked, a vector of *a priori* values supplied by the user, or a vector of estimates obtained from some other program such as BILOG (1989). The user may specify either an exploratory or confirmatory analysis. In either case, starting values for the parameters to be estimated may be supplied by NOHARM or the user. The default starting values are .5 for the a-parameters and .1 for any off-diagonal elements of the $\boldsymbol{\Sigma}_\theta$ correlation matrix that may be estimated in a confirmatory analysis. In general, the solution is anchored by fixing items to load only on certain dimensions. If the analysis is two dimensional, a single item will be fixed to load only on the first dimension. For a three dimensional analysis, a second item is fixed to load only on the first two dimensions, and so on. If the analysis is exploratory the pattern matrix is set such that the first m-1 items are fixed in this manner. In a confirmatory analysis the user may specify which items are used to anchor the solution. Also, in a confirmatory analysis, the user may allow for correlated thetas while in the exploratory mode the analysis is orthogonal. For further details on running NOHARM the reader is referred to Fraser (1986).

The program estimates the d-parameters and a-parameters, and, when appropriate, the off-diagonal elements of $\boldsymbol{\Sigma}_\theta$. Other output includes the residual covariances of the items and the root mean square of these values. The program also provides the common factor model parameterization of the normal ogive model parameters, and, when the analyses are exploratory, provides Varimax and Promax rotations of the pattern matrix.

In addition to the parameters of the multidimensional normal ogive, this study will compute and evaluate indices proposed by Reckase (1985, 1986) for multidimensional item difficulty (MDIFF) and multidimensional item discrimination (MDISC). MDIFF consists of a set of statistics that describes item difficulty as the direction from the origin in the multidimensional space in which the item provides the most information and the signed distance in that direction to the most informative point on the item response surface. For a given item, the direction cosines of MDIFF are given by

$$\cos\alpha_{ik}=\frac{a_{ik}}{\left(\sum_{k=1}^{m}a_{ik}^{2}\right)^{\frac{1}{2}}}. \tag{2}$$

where the a_{ik} are elements of the vector $\mathbf{a}_i$ given in equation 1. The distance component of MDIFF is given by

$$D_i=\frac{-d_i}{\left(\sum_{k=1}^{m}a_{ik}^{2}\right)^{\frac{1}{2}}}. \tag{3}$$

where d_i is the item difficulty index given in equation 1.

MDISC indicates item discrimination in the MDIFF direction and is given as,

$$MDISC = \left(\sum_{k=1}^{m} a_{ik}^{2} \right)^{\frac{1}{2}}. \tag{4}$$

To summarize, the parameters of interest in this study were:

1. **a** - the (i x m) matrix of NOHARM estimated item discriminations
2. **d** - the (i x 1) vector of NOHARM estimated item difficulties
3. **MDISC** - the (i x 1) vector of multidimensional item discriminations
4. **α** - the (i x m) matrix of angles obtained from the **cos α** components of MDIFF
5. **D** - the (i x 1) vector of distance components of MDIFF

Two separate studies are reported. The first involves real data and was designed to establish empirical estimates of standard errors, assess bias, and evaluate the effects of using different starting values and anchor items. The second study consisted of a simulation intended to compare NOHARM with other estimation programs. Following the design of the Ackerman (1988) study, the focus was on the ability to reproduce data using NOHARM estimated item parameters.

Method

Real Data Analyses

Data. The data used in this study were obtained from a 1987 national administration of a form of the P-ACT$^+$ mathematics test. This test is given primarily to high school sophomores and consists of 40 multiple-choice items measuring achievement in the content areas of pre-algebra, algebra, plane geometry and coordinate geometry. A "population" sample of 30,000 cases was selected at random from a total administration sample of approximately 140,000 examinees. Ten replication samples of n=2000 each were then selected at random and with replacement from the population sample.

Analyses. Earlier factor analyses of several PACT datasets had suggested three factors, interpreted as a geometry factor, an algebraic symbol manipulation factor, and a word problems factor. A preliminary NOHARM analysis of the 30000 case sample was carried out in three dimensions to confirm this structure and to assess how well this model would fit the "population" data, an important pre-requisite for the subsequent analyses. Results indicated a very good fit, with a root mean squared residual (RMSR) product moment of .003. Therefore, product moment matrices for each of the 10 samples were also fit by a three-dimensional model. Estimates of the c_j-parameters were obtained from a unidimensional analysis using BILOG (1989) and were input as fixed values for the NOHARM analyses. Initially, default settings were employed, so that the first two items were used to anchor the solution (see earlier discussion), starting values were .5 for the **a** estimates, and the solutions were orthogonal. Additional analyses were carried out to assess the effects of using different starting values and different anchor items. For questions related to starting values, three additional analyses were carried out on the population sample using starting values of .3, .8 and 1.5. To assess the effects of using different anchor items, the ten replication samples were re-run using two different sets of two anchor items.

As stated earlier, the main interest in this study was in obtaining empirical estimates of the standard errors of the parameters. This was accomplished by computing the standard deviations of the parameter estimates for the 10 replications. This was done for both the NOHARM model parameter estimates as well as the MIRT statistics. In addition, an estimate of bias was computed for each parameter as the average of the difference between each of the ten estimates of that parameter and the "population" value. For the follow-up studies pertaining to starting values, the d and **a** estimates were averaged over items and these averages were compared across the different analyses. Also, correlations were obtained for each set of 40 parameter estimates across the different starting value conditions. For the analyses involving different anchor items, the main concern was whether the arbitrary use of the first m-1 items as anchors would lead to unnecessarily high standard errors. Therefore, for these analyses the standard errors were re-computed for the different configurations and compared with those obtained under the default conditions.

Analysis of Simulated Data

Data. Data for the simulation were generated from a multidimensional two-parameter logistic (M2PL) model using bivariate normal theta distributions and item parameters from an earlier study (Ackerman, 1988). These parameters, given in Table 1, were selected to provide uniform information over the ability continuum. Fifty items and two dimensions were used in the simulation. Two data sets of n=2000 were generated, one with $r_{\theta 1 \theta 2}=0.0$ and the other with $r_{\theta 1 \theta 2}=0.5$.

Insert Table 1 about here

Analyses. The purpose of the simulation study was to investigate how well input data could be reproduced using NOHARM estimated item parameters. NOHARM was used to obtain two dimensional solutions for each of the datasets. Default settings were employed for both analyses, with the *c*-parameters fixed to zero to create a multidimensional extension of the 2-parameter model. In order to compare the results of this study with those of the earlier study, estimates of ability were needed. Since NOHARM does not provide such estimates, a program was written to compute expected *a posterior* (EAP) means for each examinee. The choice to use EAP scores was made to provide the most direct comparison with TESTFACT.

For each person and item, a standardized residual was computed as

$$RES_{ij}=\frac{y_{ij}-p_{ij}}{\sqrt{p_{ij}(1-p_{ij})}} \tag{5}$$

where y_{ij} is a 0/1 score on item i for person j, and p_{ij} is the expected probability of a correct response on item i for person j computed from equation 1. The focus of the evaluation was on the moments of the distribution of the residuals for each item and on the average of the means and standard deviations of these values over items. The mean residuals (both for individual items and overall) will serve primarily to provide a check on the accuracy of the estimation procedure and should be very near zero if the program is functioning properly and providing unbiased estimates. However, assessment of bias alone is not sufficient to address the practical utility of the procedure, since a procedure may be unbiased but have such high variance that it is practically useless. A better indication of the overall quality of the procedure will be provided by the standard deviations of the fitted residuals.

Results

Real Data Analyses

Tables 2 and 3 contain the means, average biases and standard deviations (empirical standard errors) for the NOHARM and MIRT parameter estimates, respectively. The last row in each table gives the means of these values over items. From Table 2 it can be seen that the overall average of the empirical standard errors for d is .15 and ranges from .12 to .15 for the a's. For the MIRT statistics, the average standard errors are .17 for MDISC, .09 for D, and range from 5.76 degrees to 7.04 degrees for the α's. Inspection of the standard errors at the item level indicates that most of the parameters were reasonably well estimated. There were however some notable exceptions. For example, the estimates of d, a_1, and MDISC for item 1 were extremely unstable, indicating a possible problem in using that item to anchor the first axis. There was also a tendency for the d and MDISC estimates to be less stable for the more difficult items (indicated by large negative values for d_i). On the other hand, D, the distance component of MDIFF seems to have been generally well estimated. For the a_{ik}, there appears to be a tendency for the estimation to become less stable in the second and third dimensions. For the α_{ik} this occurred only for the third dimension.

Overall, there seems to be little important bias occurring. As with the standard errors, some exceptions can be found at the individual item level. Note in particular that d, a_1 and MDISC for Item 1 were apparently quite far off the value obtained in the analysis of the large sample, again suggesting a possible problem in using this item to anchor the solutions.

Insert Tables 2 & 3 about here

Additional Analyses: Starting Values and Anchor Items

The follow-up analyses were intended to address two questions: (1) would it be possible to reduce the standard errors of the estimates by a better choice of anchor items and (2), how sensitive is the analysis to the choice of starting values for the **a**- and **d**-parameters?

There were two reasons for the concern over the choice of anchor items. First, in many tests, including the PACT+, the items are ordered by difficulty so that the first items are easier and generally less discriminating. The question was whether the use of items with relatively low discriminations as anchor items would lead to less stable solutions and poorer estimates overall than might be obtained by using items with better discrimination. The second concern stemmed from the fact that in solutions involving $m > 2$ dimensions, the first m-1 items are chosen arbitrarily by NOHARM as the anchor items. Alternatively, it would seem advantageous to use items to anchor different dimensions that were somehow known to measure different dimensions.

To address these questions the analyses were re-run on the ten replication samples using two different sets of anchor items. The first set was chosen on purely statistical grounds: two items (items 18 and 24) were chosen that were found to have average values of difficulty (d) and multidimensional discrimination (MDISC) in the default analyses. The other set of items was chosen on substantive grounds: the results of a previous factor analysis were used to identify two items (items 3 and 32) that loaded on fairly distinct dimensions. As in the previous study, empirical standard errors were computed as the standard deviations of the parameter estimates over the ten replications.

Tables 4 and 5 contain the average of the empirical standard errors over items for the original analyses using NOHARM defaults and the two additional sets of analyses. Contrary to expectations, the use of different anchor items not only failed to improve the standard errors but actually caused them to increase, in some cases substantially.

Although the standard errors of item 1 were reduced to some extent, the standard errors of one of the new anchor items increased. For example, in the 18/24 analysis, the standard error of a_1 for item 1 was .34, down considerably from its value of .60 in the default analysis. However the standard errors of a_1 for item 18 in the 18/24 analysis inflated from .12 to .82. Similar results were obtained for the other parameters of item 18 in this analysis and for item 32 in the 3/32 analysis. Thus it seems that the problem is not so much which items are fixed but rather the method itself which leads to larger standard errors for the fixed items. Nevertheless, it is not altogether clear why selecting items on substantive grounds led to increased standard errors overall. Further research is needed to clarify these findings.

Insert Tables 4 & 5 about here

The results of the analyses run under different starting values are summarized in Tables 6 and 7. Recall that three additional analyses were carried out on the population sample of n=30000 using starting values of .3, .8 and 1.5. Table 6 gives the means and standard deviations of the NOHARM parameter estimates for these analyses along with those from the default analyses. The correlations between the estimates for each of the starting value conditions are given in Table 7.

The results given in Table 6 indicate that varying the starting values had some impact, although the effects are not large and are somewhat inconsistent. Increasing the starting values led to a decrease in the levels of parameter estimates, with the exception of a_1 under starting values of 1.5. There was also a tendency for the variability of the estimates to decrease with larger starting values, although again the trends were not consistent. Moreover, since the standard deviations reported in Table 6 are not estimates of standard errors, it is difficult to make valuative judgements regarding increased or decreased variability.

The correlations reported in Table 7 reveal a relationship between the degree of correspondence between the a_1 estimates obtained from different starting values and the closeness of those starting values. In general, the greater the disparity between starting values, the lower the correspondence between estimates. This trend was not observed for the d estimates.

Insert Tables 6 & 7 about here

Analyses of Simulated Data

Tables 8 and 9 contain the summary statistics of the residual analyses for the $r_{\theta_1\theta_2}=0.0$ data (Dataset 1) and the $r_{\theta_1\theta_2}=0.5$ data (Dataset 2), respectively. The results indicate that NOHARM performed well in terms of being able to reproduce the data with little or no bias on average. At the item level, the mean residuals were less than .01 in absolute value for 42 of 50 items in Dataset 1 and 38 of 50 items in Dataset 2. The overall mean residual was .001 for Dataset 1 and .000 for Dataset 2. While it is apparent that some extreme values occurred, the magnitudes of the standard deviations of the residuals suggest that the estimated probabilities of correct response were reasonably well behaved. For comparative purposes, Table 10 presents the overall mean and standard deviation of the residuals obtained form the NOHARM analyses along with those obtained for the other estimation programs evaluated in the Ackerman (1988) study. It is apparent that NOHARM and TESTFACT were equally effective in reproducing the data as reflected by the lack of average bias in the residuals. Both programs also appear to be roughly equivalent in terms of the variance of the residuals.

Insert Tables 8, 9 & 10 about here

Summary and Conclusions

The parameter estimates provided by NOHARM, along with MIRT item statistics computed from those estimates, were evaluated in terms of their estimated standard errors, bias relative to population values, and robustness under different starting configurations. In addition, a simulation was carried out to permit comparisons with an earlier study that evaluated and compared several other estimation programs.

For most of the items the estimated standard errors of the parameter estimates seemed to be reasonably small, and there was little indication of important bias in the estimation. Overall, D, the distance component of MDIFF was the most stable parameter, while the a_3 and α_3 estimates were the least stable. Also, the estimation procedure used by NOHARM seems fairly robust to different starting values. Somewhat surprisingly, attempts to improve the standard errors by using different anchor items were unsuccessful. It is not clear why the arbitrary use of the first m-1 items as anchors of an m-dimensional solution led to lower standard errors than did the use of items selected on statistical or substantive grounds. It does, however, appear that regardless of which items are chosen as anchors, the parameters for at least one of them will be poorly estimated. Further research is needed to clarify these findings.

Although it was necessary in the simulation study to employ an external program to obtain the needed ability estimates from the NOHARM analysis, the results nevertheless indicated that both the marginal maximum likelihood algorithm used by TESTFACT and the least squares algorithm used by NOHARM were equally effective at reproducing data under well-fitting model conditions. Together the findings of this study support the use of NOHARM in practical MIRT applications.

References

Ackerman, T. A. (1988, May). *Comparison of multidimensional IRT estimation procedures using benchmark data.* Paper presented at the ONR Contractor's meeting, Iowa City, IA.

Carlson, J. E. (1987). *Multidimensional item response theory estimation* [Computer program] (Research Report ONR97-2). The American College Testing Program.

Fraser, C. (1986). *NOHARM: An IBM PC Computer Program for Fitting Both Unidimensional and Multidimensional Normal Ogive Models of Latent Trait Theory* [Computer Program]. Center for Behavioral Studies, The University of New England, Armidale, New South Wales, Australia.

McKinley, R. L. (1987) *MULTIDIM User's Guide* [Computer program manual]. Educational Testing Service, Princeton N.J.

Mislevy, R. J., & Bock, R. D. (1989). *BILOG: Item analysis and test scoring with binary logistic models* [Computer Program]. Scientific Software, Inc, Moorseville, IN.

Reckase, M. D. (1985). The difficulty of items that measure more than one ability. *Applied Psychological Measurement*, *9*, 401-412.

Reckase, M. D. (1986, April). *The discriminating power of items that measure more than one dimension.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Wilson, D., Wood, R., & Gibbons, R. (1984). *TESTFACT: Test scoring, item statistics, and item factor analysis* [Computer program]. Scientific Software, Inc.: Mooresville, IN.

Table 1

Uniform Information Item Set

Item No.	a_1	a_2	D	d	MDISC	α
1	1.351	0.270	-2.499	3.442	1.377	11.311
2	0.653	1.136	0.008	-0.011	1.311	60.095
3	1.365	0.027	-0.791	1.080	1.366	1.151
4	0.298	1.450	2.482	-3.675	1.481	78.386
5	1.391	1.171	2.495	-4.536	1.818	40.089
6	1.828	0.000	0.470	-0.860	1.828	0.001
7	1.796	0.011	-0.985	1.769	1.796	0.365
8	1.474	0.017	2.000	-2.948	1.474	0.644
9	0.012	1.422	-1.500	-0.823	1.422	89.526
10	0.153	1.336	2.491	-3.351	1.345	83.464
11	1.326	0.286	2.072	-2.810	1.356	12.151
12	1.678	0.222	-0.096	0.163	1.693	7.541
13	1.424	0.001	-2.498	3.557	1.424	0.042
14	0.117	1.808	0.869	-1.574	1.811	86.289
15	0.176	1.294	-0.441	0.576	1.306	82.249
16	1.414	0.040	-2.223	3.145	1.415	1.612
17	1.350	0.000	2.390	-3.227	1.350	0.000
18	0.236	1.743	-2.039	3.586	1.759	82.276
19	1.109	0.839	-0.240	0.333	1.390	37.114
20	0.000	1.438	1.306	-1.879	1.438	89.999
21	0.011	1.522	1.747	-2.660	1.522	89.576
22	1.399	0.063	1.939	-2.717	1.401	2.578
23	0.351	1.376	-0.251	0.356	1.420	75.694
24	0.000	1.568	1.358	-2.129	1.568	89.990
25	0.093	1.377	2.334	-3.290	1.380	86.131

(Table continues)

Item No.	a_1	a_2	D	d	MDISC	α
26	0.206	1.481	-1.500	-1.151	1.495	82.077
27	1.545	0.430	0.894	-1.434	1.604	15.551
28	0.404	1.338	-2.363	3.302	1.397	73.199
29	0.811	1.522	-0.934	1.611	1.725	61.944
30	1.459	0.133	2.047	-3.000	1.465	5.192
31	0.606	2.123	-2.221	4.903	2.208	74.064
32	1.375	0.002	2.000	-2.750	1.375	0.081
33	0.093	1.640	-1.975	3.244	1.642	86.739
34	0.158	1.504	2.500	-3.781	1.512	83.998
35	0.000	1.343	2.336	-3.137	1.343	90.000
36	1.451	0.288	-0.217	0.320	1.480	11.241
37	1.893	0.117	-2.428	4.604	1.896	3.546
38	0.026	1.385	-1.168	1.617	1.385	88.909
39	0.395	1.351	0.055	-0.077	1.408	73.712
40	2.168	0.006	-0.712	1.544	2.168	0.150
41	0.057	1.355	1.565	-2.122	1.356	87.603
42	0.685	1.276	-0.861	1.246	1.448	61.772
43	0.064	1.471	2.492	-3.669	1.472	87.495
44	1.273	0.815	2.488	-3.759	1.511	32.622
45	0.439	1.413	-1.407	2.082	1.479	72.727
46	1.451	0.266	0.981	-1.448	1.475	10.391
47	0.077	1.425	-0.341	0.486	1.427	86.894
48	1.318	0.036	-2.393	3.154	1.318	1.560
49	1.409	0.000	-2.500	3.522	1.409	0.009
50	1.402	0.000	0.401	-0.563	1.402	0.000

Table 2

Means, Average Bias and Empirical Standard Errors of NOHARM Parameter Estimates

	d			a_1			a_2			a_3		
Item	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD
1	4.38	-.52	.73	3.27	[illegible]	[illegible]	.00	.00	.00	.00	.00	.00
2	1.33	.02	.03	.52	.03	.08	.50	.01	.09	.00	.00	.00
3	.77	.02	.04	.49	[illegible]	.05	.19	-.03	.09	.15	.08	.10
4	1.09	.06	.06	[illegible]	-.08	.14	.49	-.03	.09	.45	.00	.06
5	.51	.01	.06	.43	-.01	.08	.10	.00	.07	.28	-.01	.11
6	.21	.03	.12	.62	-.08	.08	.47	-.04	.08	.41	.05	.07
7	.84	.01	.05	.75	-.06	.07	.32	-.01	.08	.39	.01	.07
8	1.31	-.03	.09	1.16	-.10	.14	.09	.05	.09	.29	-.01	.11
9	.85	-.01	.06	.81	-.15	.10	.46	-.04	.11	.36	.05	.11
10	.67	.02	.06	.97	-.10	.07	.35	-.02	.11	.58	-.04	.06
11	1.31	-.01	.05	.87	-.09	.07	.17	.00	.11	.29	.01	.08
12	1.07	-.04	06	.85	.05	.09	.46	.04	.05	.55	-.05	.09
13	.67	.03	.09	1.30	-.07	.12	.09	.01	.13	.89	-.16	.24
14	-.39	.03	.15	.73	-.17	.15	.70	-.08	.17	.55	.09	.17
15	-.52	-.15	.22	.59	-.03	.14	.69	.06	.15	.53	.18	.19
16	-.48	-.16	.31	.25	-.04	.16	.74	-.02	.26	.22	.13	.32
17	.99	.01	.09	1.43	-.04	.15	.16	-.03	.17	.94	-.17	.33
18	-.32	.02	[illegible]	.65	-.05	.12	.60	-.07	.15	.50	.03	.08
19	-.07	.00	.09	.50	-.06	.06	.36	-.03	.06	.31	.05	.07
20	-.16	.04	.09	.84	-.12	.12	.80	-.10	.13	.60	.11	.13
21	.16	-.09	.06	.79	.17	.14	.95	.18	.22	.43	-.09	.21
22	-.11	-.10	.10	.60	.11	.08	.95	.21	.11	.34	.01	.15
23	-.23	-.15	.16	.38	.05	.14	.90	.06	.26	.33	.11	.16
24	.20	.02	.06	.79	.01	.06	.27	-.02	.07	.58	.01	.08
25	-1.12	.17	.12	.62	-.12	.12	.33	-.07	.06	.68	-.04	.13
26	-1.01	.18	.22	.70	-.15	.14	.60	-.16	.17	.77	-.04	.30
27	-.51	-.03	.08	.50	-.03	.07	.38	-.01	.09	.47	.12	.11
28	.55	.01	.11	1.22	.36	.27	1.55	.41	.39	.72	-.10	.29
29	-.19	-.03	.08	.73	.00	.09	.60	-.02	.11	.69	.07	.13
30	-.41	.03	.07	.48	-.04	.06	.47	-.05	.08	.50	.01	.07
31	-.60	-.09	.12	.58	.06	.16	1.19	.11	.25	.59	.04	.14
32	-1.01	.00	.23	.76	-.01	.08	.75	-.07	.18	.97	.06	.28
33	-.58	.02	.11	.39	-.01	.09	.35	-.02	.08	.60	-.03	.10
34	-1.07	-.17	.22	.20	.11	.05	.42	.08	.12	.59	.04	.17
35	-.85	.02	.06	.63	-.08	.06	.83	-.05	.10	.88	.06	.08
36	-.42	-.07	.05	.34	.05	.08	.49	.01	.08	.46	-.04	.14
37	-1.35	.13	.30	.18	-.08	.10	.70	-.11	.24	.87	-.02	.25
38	-1.45	-.24	.63	-.01	.02	.13	.66	-.01	.30	1.13	.24	.47

(Table continues)

Item	d			a_1			a_2			a_3		
	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD
39	-2.60	.14	.56	.27	-.16	.11	.83	-.20	.30	1.35	.12	.34
40	-.69	-.01	.07	.27	.01	.11	.37	-.05	.14	.96	.00	.15
Overall Mean	.00	-.02	.15	.67	-.04	.12	.53	-.01	.14	.57	.02	.15

Table 3

Means, Average Bias and Empirical Standard Errors of MIRT Parameter Estimates

	MDISC			D			α_1			α_2			α_3		
It.	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD
1	3.27	-.51	.60	-1.34	-.05	.03	.00	.00	.00	90.00	.00	.00	90.00	.00	.00
2	.72	.02	.09	-1.86	.05	.22	43.68	-.85	6.57	46.33	.85	6.57	90.00	.00	.00
3	.56	-.07	.06	-1.38	-.23	.12	28.63	6.55	6.83	70.00	.13	8.77	73.76	-11.49	11.10
4	1.11	-.08	.14	-.99	-.12	.13	36.73	1.41	3.74	64.07	-.15	4.23	66.20	-1.90	3.01
5	.54	-.03	.04	-.95	-.07	.14	35.29	-.89	12.32	79.20	-.67	7.94	58.37	-.28	13.48
6	.88	-.05	.10	-.24	-.04	.15	45.66	4.34	3.18	57.74	.75	3.40	61.88	-5.94	4.52
7	.90	-.06	.06	-.93	-.07	.07	34.08	1.63	3.92	69.48	-.51	4.94	64.62	-2.02	4.95
8	1.21	-.10	.13	-1.09	-.07	.06	15.50	.72	5.89	85.55	-2.83	4.84	75.90	-.31	5.74
9	1.01	-.13	.11	-.85	-.10	11	35.96	5.44	5.94	63.13	-1.19	5.69	68.89	-6.59	6.40
10	1.19	-.11	.06	-.57	-.08	.07	35.44	.81	2.71	72.66	-.76	5.19	60.66	-.87	3.97
11	.94	-.09	.05	-1.40	-.14	.08	22.65	1.38	4.41	79.35	-1.17	6.79	71.73	-2.33	6.09
12	1.11	.03	.06	-.96	.06	.04	40.24	-2.15	4.93	65.63	-1.27	3.72	60.51	3.41	5.05
13	1.59	-.16	.16	-.42	-.07	.05	34.69	-3.85	6.89	86.60	-.44	4.96	56.10	3.36	8.04
14	1.17	-.12	.18	.32	.02	.10	50.81	6.92	[illegible]	53.63	.26	5.24	61.27	-8.60	9.55
15	1.07	.11	.19	.47	.10	.13	56.46	5.32	6.73	49.48	.64	5.07	59.67	-6.87	9.80
16	.88	-.05	.26	.51	.27	.26	71.19	3.98	11.46	32.63	-2.80	10.84	73.85	-9.10	21.01
17	1.74	-.15	.24	-.57	-.05	.05	33.82	-4.59	7.42	84.48	.83	5.79	57.65	3.57	9.34
18	1.02	-.07	.14	.30	.01	.12	50.66	.48	5.31	54.20	2.42	7.40	60.27	-3.72	5.36
19	.69	-.04	.05	.10	.01	.13	43.42	3.65	4.83	59.04	1.29	4.79	63.60	-6.25	6.41
20	1.31	-.09	.14	.12	-.02	.06	50.25	3.74	4.55	52.55	2.37	2.96	62.54	-7.23	6.58
21	1.33	.14	.15	-.12	.07	.04	53.36	-3.82	6.84	45.03	-1.46	7.58	70.17	6.41	11.37
22	1.19	.22	.11	.09	.06	.08	59.48	.29	4.30	36.49	-2.40	3.71	73.40	2.33	7.59
23	1.05	.08	.26	.20	.14	.12	67.71	.40	9.03	34.41	1.15	5.60	71.36	-4.19	7.97
24	1.02	.01	.08	-.19	-.02	.06	39.25	-.37	3.54	75.55	1.22	3.94	55.23	-.50	3.58
25	.99	-.14	.12	1.14	-.02	.05	50.50	3.29	7.40	70.40	2.22	4.18	46.41	-4.96	6.52
26	1.22	-.21	.28	.83	-.01	.07	54.25	2.73	5.64	60.04	4.10	7.45	51.21	-6.94	10.88
27	.80	.04	.07	.64	-.00	.08	50.52	5.29	5.92	61.71	2.55	6.52	53.29	-8.04	8.76
28	2.13	.46	.37	-.26	.04	.05	54.94	-2.64	5.38	43.84	-2.83	5.79	69.13	7.10	10.15
29	1.18	.02	.12	.16	.03	.06	51.19	1.03	5.34	59.31	1.70	3.74	54.34	-2.97	5.72
30	.84	-.04	.09	.49	-.01	.06	55.31	1.45	4.09	55.97	1.82	3.89	53.22	-3.35	3.95
31	1.47	.11	.23	.41	.03	.05	66.05	.21	7.13	35.90	-1.27	4.32	65.92	.43	4.31
32	1.46	-.01	.26	.69	.01	.06	57.22	1.15	6.79	59.06	2.91	4.10	48.82	-4.05	6.68
33	.80	-.04	.09	.71	.02	.07	60.58	-.64	6.85	64.04	.40	6.21	41.91	-.52	5.08
34	.75	.10	.15	1.42	.03	.14	74.15	-4.91	5.18	57.04	-2.11	5.24	37.96	4.50	4.60
35	1.37	-.03	.09	.62	.00	.03	62.36	3.49	3.26	52.64	1.65	3.20	49.99	-4.52	3.89
36	.76	-.01	.07	.56	.01	.09	63.20	-4.22	7.02	49.46	-.56	8.87	53.34	3.31	10.29
37	1.15	-.11	.29	1.18	-.01	.09	80.00	4.77	6.92	52.48	3.07	8.66	40.23	-5.28	7.33

(Table continues)

	MDISC			D			α_1			α_2			α_3		
It.	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD	Mean	Bias	SD
38	1.33	.18	.51	1.07	.05	.14	89.24	.23	5.32	60.04	4.73	10.54	30.57	-5.34	10.18
39	1.63	-.03	.38	1.61	-.06	.11	80.46	5.89	3.40	59.19	7.83	8.14	32.88	-9.57	7.78
40	1.02	-.03	.16	.69	.02	.07	73.72	-.01	7.03	68.70	2.52	5.96	28.39	-3.11	4.68
Overall Mean	1.16	-.03	.17	.01	.00	.09	50.22	1.19	5.76	60.33	.63	5.67	59.13	-2.56	7.04

Table 4

Average Standard Errors of NOHARM Parameter Estimates Using Different Anchor Items

Anchor Items	d	a_1	a_2	a_3
Default 1/2	.150	.117	.138	.151
18/24	.170	.211	.237	.221
3/32	.169	.165	.211	.329

Table 5

Average Standard Errors of MIRT Parameter Estimates Using Different Anchor Items

Anchor Items	MDISC	D	α_1	α_2	α_3
Default 1/2	.168	.090	5.759	5.668	7.042
18/24	.204	.094	12.757	12.664	9.429
3/32	.213	.093	9.142	8.899	14.913

Table 6

Means and SD's of NOHARM Parameter Estimates Using Different Starting Values

	d		a_1		a_2		a_3	
Starting Value	Mean	SD	Mean	SD	Mean	SD	Mean	SD
.3	.022	1.178	.680	.567	.532	.372	.580	.299
.5*	-.003	1.082	.671	.471	.528	.372	.574	.300
.8	-.014	1.026	.674	.419	.516	.357	.562	.286
1.5	-.011	1.059	.715	.468	.508	.375	.544	.277

*Default

Table 7

Correlations Between NOHARM Parameter Estimates Obtained Under Different Starting Values

d

	Starting Value			
	.3	.5*	.8	1.5
.3	1.000			
.5	.994	1.000		
.8	.984	.998	1.000	
1.5	.988	.999	.999	1.000

a_1

	Starting Value			
	.3	.5*	.8	1.5
.3	1.000			
.5	.987	1.000		
.8	.957	.991	1.000	
1.5	.931	.964	.979	1.000

a_2

	Starting Value			
	.3	.5*	.8	1.5
.3	1.000			
.5	1.000	1.000		
.8	.986	.990	1.000	
1.5	.894	.907	.950	1.000

a_3

	Starting Value			
	.3	.5*	.8	1.5
.3	1.000			
.5	1.000	1.000		
.8	.987	.989	1.000	
1.5	.893	.901	.949	1.000

* Default

Table 8

Residual Analysis of NOHARM Calibration: Dataset 1

Item	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
1	-.025	1.227	-13.364	266.058	-30.113	.974
2	-.003	.955	-.051	.010	-4.774	4.075
3	.002	.924	-1.253	1.491	-4.365	2.108
4	.009	1.000	6.714	59.421	-.998	14.077
5	-.001	.833	8.651	97.126	-.980	13.751
6	-.007	.888	1.190	6.380	-3.270	8.833
7	.003	.867	-1.824	4.545	-6.030	1.623
8	.043	1.111	7.311	96.068	-1.213	21.981
9	.008	.880	-.720	.693	-3.850	2.938
10	.016	1.089	9.411	139.657	-.999	22.700
11	.001	.936	4.241	23.170	-1.081	9.543
12	-.006	.876	-.384	1.656	-5.677	3.379
13	.006	.926	-7.230	71.752	-12.605	.815
14	.005	.900	2.245	11.883	-3.423	9.445
15	-.008	.957	-.910	2.037	-6.226	3.068
16	-.023	1.085	-3.167	8.239	-4.992	.423
17	-.013	.863	4.356	22.445	-.812	8.642
18	.008	.824	-5.585	40.828	-9.548	1.063
19	-.002	.940	-.500	.403	-5.000	3.657
20	.011	.958	4.259	42.344	-1.870	15.423
21	.005	.907	4.035	23.603	-1.571	11.014
22	-.007	.903	6.889	103.085	-1.206	18.922
23	-.006	.944	-.502	.300	-4.450	3.838
24	.010	.961	4.531	49.269	-2.093	16.547
25	.002	.892	5.583	46.205	-1.307	13.413
26	.002	.899	1.215	2.090	-3.758	5.243
27	.000	.903	1.780	6.262	-3.947	6.787
28	-.008	.990	-6.375	57.331	-14.420	1.001
29	.002	.859	-1.657	4.437	-6.011	2.223
30	-.004	.963	7.[illegible]93	100.806	-.882	19.694
31	.009	.684	-8.167	86.056	-11.135	.919
32	-.003	.882	4.069	22.253	-1.472	8.408
33	.005	.854	-5.146	45.846	-13.637	1.074
34	.000	.921	8.062	94.821	-1.013	16.198
35	.001	.941	5.774	58.480	-1.284	16.202
36	.002	.907	-.211	-.345	-3.825	2.592
37	.008	.936	-9.337	117.878	-17.041	.683
38	-.005	.946	-2.135	6.614	-6.827	1.942
39	-.002	.934	-.035	.211	-4.452	4.015
40	.004	.782	-1.904	9.505	-7.624	2.740
41	.002	.928	2.694	9.026	-1.565	8.373
42	-.007	.960	-1.910	6.575	-8.490	2.861

(Table continues)

Item	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
43	.004	.921	6.014	47.640	-1.133	12.848
44	-.002	.887	6.989	65.535	-.853	13.188
45	-.001	.916	-2.784	9.584	-7.780	1.528
46	-.002	.912	1.741	5.055	-2.234	7.569
47	-.002	.933	-.567	.330	-4.465	3.625
48	.005	.980	-3.258	8.623	-3.645	.312
49	.019	.882	-4.253	18.627	-7.050	.601
50	.000	.925	.688	.491	-2.656	4.673

Table 9

Residual Analysis of NOHARM Calibration: Dataset 2

Item	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
1	-.024	1.024	-9.016	122.787	-20.199	1.199
2	.001	.949	-.069	.096	-4.318	4.171
3	-.001	.948	-1.147	1.957	-5.985	3.654
4	.016	1.168	17.936	518.789	-1.208	37.178
5	.009	.803	8.951	106.537	-1.229	13.507
6	.000	.894	.794	3.729	-6.085	6.176
7	-.003	.878	-2.800	19.436	-11.303	2.263
8	.009	1.003	5.991	58.410	-1.229	15.742
9	.004	.916	-.900	1.227	-5.733	2.944
10	.002	.874	5.205	33.354	-1.047	9.188
11	.008	.972	5.066	37.213	-1.558	11.717
12	-.009	.984	-1.774	21.359	-12.502	5.711
13	-.001	.856	-6.398	55.508	-11.793	1.009
14	.000	.874	1.664	5.563	-3.591	6.305
15	-.007	.966	-1.625	12.849	-11.905	5.218
16	-.007	.978	-9.631	170.501	-23.114	1.120
17	.016	1.098	8.449	104.167	-.987	18.496
18	-.008	.897	-10.751	208.599	-22.364	1.259
19	.000	.920	-.462	2.470	-6.105	3.897
20	.003	.933	3.354	25.208	-3.333	13.232
21	.007	.934	4.446	30.402	-1.374	11.422
22	.013	1.016	7.122	97.734	-1.563	20.389
23	-.005	.937	-.690	2.631	-6.084	4.491
24	.005	.966	3.087	19.278	-1.408	12.789
25	.013	1.032	7.092	76.751	-1.367	16.962
26	-.006	.896	.869	1.385	-4.621	3.807
27	.004	.926	2.107	10.561	-2.872	9.876
28	.003	.877	-4.174	20.275	-8.092	1.381
29	.003	.938	-2.404	17.081	-9.742	5.554
30	.000	.912	4.430	25.699	-1.067	10.057
31	-.069	2.278	-30.475	7.540	-82.177	1.185
32	.002	.933	4.232	27.227	-1.359	11.649
33	-.025	1.137	-9.691	137.001	-20.151	1.042
34	.019	1.096	19.192	544.247	-1.290	35.087
35	.008	.976	5.432	40.855	-.999	13.574
36	-.005	.937	-.286	2.318	-5.871	5.792
37	-.006	.922	-21.036	649.789	-30.987	.993
38	-.001	.911	-2.178	8.699	-9.412	1.670
39	.000	.926	.093	1.565	-5.019	4.471
40	-.006	.857	-2.549	17.539	-9.660	3.041
41	.016	1.167	9.724	184.576	-1.735	27.255
42	.000	.934	-1.458	4.051	-6.994	3.566

(Table continues)

Item	Mean	SD	Skewness	Kurtosis	Minimum	Maximum
43	.017	1.050	9.556	143.616	-1.059	22.679
44	.011	.956	7.001	73.232	-1.268	15.481
45	-.004	.911	-2.900	13.006	-8.299	2.296
46	.004	.920	2.200	9.510	-2.710	8.604
47	.006	.931	-.443	1.192	-5.463	3.912
48	-.013	1.030	-6.099	49.314	-13.754	.949
49	-.001	.947	-6.323	53.701	-12.397	.832
50	.003	.917	.735	.912	-2.857	5.429

Table 10

Means and Standard Deviations of Standardized Residuals for Different Estimation Programs

	Program							
	MIRTE		TESTFACT		MULTIDIM		NOHARM	
Dataset	Mean	SD	Mean	SD	Mean	SD	Mean	SD
$\rho=0.0$	.251	1.452	.001	.893	-.026	1.321	.001	.966
$\rho=0.5$	.253	1.312	.000	1.154	-.024	1.217	.000	.982

Distribution List

Dr. Terry Ackerman
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. James Algina
1403 Norman Hall
University of Florida
Gainesville, FL 32605

Dr. Nancy Allen
Educational Testing Service
Princeton, NJ 08541

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

Dr. Gregory Anrig
Educational Testing Service
Princeton, NJ 08541

Dr. Ronald Armstrong
Rutgers University
Graduate School of Management
Newark, NJ 07102

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

Dr. Laura L. Barnes
College of Education
University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Dr. Isaac Bejar
Law School Admissions
Services
P.O. Box 40
Newtown, PA 18940-0040

Dr. Anne Beland
Educational Testing Service
Princeton, NJ 08541

Dr. Ira Bernstein
Department of Psychology
University of Texas
P.O. Box 19528
Arlington, TX 76019-0528

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

Dr. Bruce Bloxom
Defense Manpower Data Center
99 Pacific St.
Suite 155A
Monterey, CA 93943-3231

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

Dr. Gwyneth Boodoo
Educational Testing Service
Princeton, NJ 08541

Dr. Robert Breaux
Code 252
Naval Training Systems Center
Orlando, FL 32826-3224

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

Dr. David V. Budescu
Department of Psychology
University of Haifa
Mount Carmel, Haifa 31999
ISREAL

Dr. Gregory Candell
CTB/McGraw-Hill
2500 Garden Road
Monterey, CA 93940

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

Dr. John M. Carroll
IBM Watson Research Center
User Interface Institute, H1-B52
P.O. Box 704
Yorktown Heights, NY 10598

Dr. Robert M. Carroll
Chief of Naval Operations
OP-01B2
Washington, DC 20350

Dr. W. Chambers
Technology Manager, Code 2B
Naval Training Systems Center
12350 Research Parkway
Orlando, FL 32826-3224

Mr. Hua Hua Chang
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Raymond E. Christal
UES LAMP Science Advisor
AFHRL/MOEL
Brooks AFB, TX 78235

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
Los Angeles, CA 90089-1061

Director, Manpower Program
Center for Naval Analyses
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
4401 Ford Avenue
Alexandria, VA 22302-0268

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

Dr. Hans F. Crombag
Faculty of Law
University of Limburg
P.O. Box 616
Maastricht
The NETHERLANDS 6200 MD

Ms. Carolyn R. Crone
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Timothy Davey
American College Testing Program
P.O. Box 168
Iowa City, IA 52243

Dr. C. M. Dayton
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Ralph J. DeAyala
Measurement, Statistics
and Evaluation
Benjamin Bldg., Rm. 4112
University of Maryland
College Park, MD 20742

Dr. Lou DiBello
CERL
University of Illinois
103 South Mathews Avenue
Urbana, IL 61801

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

Dr. Neil Dorans
Educational Testing Service
Princeton, NJ 08541

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

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
(2 Copies)

Dr. Stephen Dunbar
224B Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

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

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

Dr. George Engelhard, Jr.
Division of Educational Studies
Emory University
210 Fishburne Bldg.
Atlanta, GA 30322

ERIC Facility-Acquisitions
2440 Research Blvd, Suite 550
Rockville, MD 20850-3238

Dr. Benjamin A. Fairbank
Operational Technologies Corp.
5825 Callaghan, Suite 225
San Antonio, TX 78228

Dr. Marshall J. Farr, Consultant
Cognitive & Instructional Sciences
2520 North Vernon Street
Arlington, VA 22207

Dr. P-A. Federico
Code 51
NPRDC
San Diego, CA 92152-6800

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

Dr. Richard L. Ferguson
American College Testing
P.O. Box 168
Iowa City, IA 52243

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

Dr. Myron Fischl
U.S. Army Headquarters
DAPE-MRR
The Pentagon
Washington, DC 20310-0300

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

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

Dr. Alfred R. Fregly
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Alice Gerb
Educational Testing Service
Princeton, NJ 08541

Dr. Robert D. Gibbons
Illinois State Psychiatric Inst.
Rm 529W
1601 W. Taylor Street
Chicago, IL 60612

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

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

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

Dr. Karen Gold
Educational Testing Service
Princeton, NJ 08541

Dr. Timothy Goldsmith
Department of Psychology
University of New Mexico
Albuquerque, NM 87131

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

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

Michael Habon
DORNIER GMBH
P.O. Box 1420
D-7990 Friedrichshafen 1
WEST GERMANY

Prof. Edward Haertel
School of Education
Stanford University
Stanford, CA 94305

Dr. Ronald K. Hambleton
University of Massachusetts
Laboratory of Psychometric
and Evaluative Research
Hills South, Room 152
Amherst, MA 01003

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

Dr. Grant Henning
Mail Stop 18-P
Educational Testing Service
Princeton, NJ 08541

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 63
San Diego, CA 92152-6800

Dr. Thomas M. Hirsch
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Paul W. Holland
Educational Testing Service, 21-T
Rosedale Road
Princeton, NJ 08541

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 92010

Ms. Julia S. Hough
Cambridge University Press
40 West 20th Street
New York, NY 10011

Dr. William Howell
Chief Scientist
AFHRL/CA
Brooks AFB, TX 78235-5601

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Steven Hunka
3-104 Educ. N.
University of Alberta
Edmonton, Alberta
CANADA T6G 2G5

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

Dr. Martin J. Ippel
Postbus 9555
2300 RB Leiden
THE NETHERLANDS

Dr. Robert Jannarone
Elec. and Computer Eng. Dept.
University of South Carolina
Columbia, SC 29208

Dr. Kumar Joag-dev
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright Street
Champaign, IL 61820

Dr. Peder Johnson
Department of Psychology
University of New Mexico
Albuquerque, NM 87131

Dr. Douglas H. Jones
1280 Woodfern Court
Toms River, NJ 08753

Dr. Brian Junker
Carnegie-Mellon University
Department of Statistics
Schenley Park
Pittsburgh, PA 15213

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Milton S. Katz
European Science Coordination
Office
U.S. Army Research Institute
Box 65
FPO New York 09510-1500

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

Mr. Hae-Rim Kim
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Jwa-keun Kim
Department of Psychology
Middle Tennessee State
University
P.O. Box 522
Murfreesboro, TN 37132

Dr. Sung-Ho Kim
Educational Testing Service
Princeton, NJ 08541

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

Dr. William Koch
Box 7246, Meas. and Eval. Ctr.
University of Texas-Austin
Austin, TX 78703

Dr. Richard J. Koubek
School of Civil Engineering
Grissom Hall
Purdue University
West Lafayette, IN 47907

Dr. Leonard Kroeker
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Dr. Jerry Lehnus
Defense Manpower Data Center
Suite 400
1600 Wilson Blvd
Rosslyn, VA 22209

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

Dr. Richard Lesh
Educational Testing Service
Princeton, NJ 08541

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

Dr. Charles Lewis
Educational Testing Service
Princeton, NJ 08541-0001

Ms. Hsin-hung Li
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Mr. Rodney Lim
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert L. Linn
Campus Box 249
University of Colorado
Boulder, CO 80309-0249

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

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

Dr. Richard Luecht
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. George B. Macready
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen J. Martin
Office of Chief of Naval
Operations (OP 13 F)
Navy Annex, Room 2832
Washington, DC 20350

Dr. Shin-ichi Mayekawa
The National Center for University
Entrance Examinations
2-19-23 KOMABA, MEGURO-KU
Tokyo 153
JAPAN

Dr. James R. McBride
HumRRO
6430 Elmhurst Drive
San Diego, CA 92120

Dr. Clarence C. McCormick
HQ, USMEPCOM/MEPCT
2500 Green Bay Road
North Chicago, IL 60064

Mr. Christopher McCusker
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert McKinley
Educational Testing Service
Princeton, NJ 08541

Mr. Alan Mead
c/o Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Timothy Miller
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

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

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

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Ratna Nandakumar
Educational Studies
Willard Hall, Room 213E
University of Delaware
Newark, DE 19716

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

Librarian
Naval Center for Applied Research
in Artificial Intelligence
Naval Research Laboratory
Code 5510
Washington, DC 20375-5000

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

Dr. James B. Olsen
WICAT Systems
1875 South State Street
Orem, UT 84058

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

Dr. Judith Orasanu
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

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

Dr. Peter J. Pashley
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

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

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

Dept. of Administrative Sciences
Code 54
Naval Postgraduate School
Monterey, CA 93943-5026

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

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

Mr. Steve Reiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. W. A. Rizzo
Head, Human Factors Division
Naval Training Systems Center
Code 26
12350 Research Parkway
Orlando, FL 32826-3224

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

Mr. Louis Roussos
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

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

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
310B Austin Peay Bldg.
Knoxville, TN 37916-0900

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

Mr. Kenneth Sarno
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Janice Scheuneman
Educational Testing Service
Princeton, NJ 08541

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

Dr. Mary Schratz
4100 Parkside
Carlsbad, CA 92008

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

Mr. Robert Semmes
N218 Elliott Hall
Department of Psychology
University of Minnesota
Minneapolis, MN 55455

Dr. Robin Shealy
Illinois State Water Survey
Room 149
2204 Griffith Dr.
Champaign, IL 61820

Ms. Kathleen Sheehan
Educational Testing Service
Princeton, NJ 08541

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

Dr. Randall Shumaker
Naval Research Laboratory
Code 5510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

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

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

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

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

Dr. William Stout
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

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

Mr. Brad Sympson
Navy Personnel R&D Center
Code-62
San Diego, CA 92152-6800

Dr. John Tangney
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Maurice Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

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

Mr. Thomas J. Thomas
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

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

Mr. Sherman Tsien
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

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

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

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 440
St. Paul, MN 55114

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

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Michael T. Waller
University of Wisconsin-Milwaukee
Educational Psychology Department
Box 413
Milwaukee, WI 53201

Dr. Ming-Mei Wang
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Thomas A. Warm
FAA Academy AAC934D
P.O. Box 25082
Oklahoma City, OK 73125

Dr. Brian Waters
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Ronald A. Weitzman
Box 146
Carmel, CA 93921

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Douglas Wetzel
Code 51
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90089-1061

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

Dr. David Wiley
School of Education
Northwestern University
Evanston, IL 60201

Dr. Charles Wilkins
Navy Personnel R&D Center
Code 13
San Diego, CA 92152

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

Dr. Mark Wilson
School of Education
University of California
Berkeley, CA 94720

Dr. Hilda Wing
Federal Aviation Administration
800 Independence Ave, SW
Washington, DC 20591

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

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

Dr. Wallace Wulfeck, III
Science Advisor
NAVOP 01SA/PERS 00R
Washington, DC 20350

Dr. Kentaro Yamamoto
02-T
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Ms. Duanli Yan
Educational Testing Service
Princeton, NJ 08541

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

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, N.W.
Washington, DC 20550

Mr. Anthony R. Zara
National Council of State
Boards of Nursing, Inc.
625 North Michigan Avenue
Suite 1544
Chicago, IL 60611