Work on item response theory was extended to two areas not extensively researched previously, including models for: (1) test items that require more than one ability for a correct response (MIRT); and (2) interaction between modules of instruction that have a hierarchical relationship (HST). In order to develop the MIRT and HST models, the author evaluated characteristics of the models, developed estimation procedures for the parameter of the models, and evaluated the models on their ability to describe real test data. A summary of the research is presented here, and references are made to papers and technical reports containing more detailed descriptions of the research efforts. (Author/LMO)
Final Report

Models for Multidimensional Tests and Hierarchically Structured Training Materials

Mark D. Reckase

Research Report ONR85-1
May 1985

Prepared under Contract No. N00014-81-K0817 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.
Models for multidimensional tests and hierarchically structured training materials.

Work on item response theory was extended to include two areas that had not been extensively researched previously. They include models for test items that require more than one ability for a correct response and models for the interaction between modules of instruction that have a hierarchical relationship. For both of these types of models, estimation procedures were developed for model parameters and extensive work was done to determine the appropriate interpretation of the parameter values. This report is a summary of work performed on these models over a three year period.
Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Development and Evaluation of MIRT Models</td>
<td>2</td>
</tr>
<tr>
<td>Analysis of the General Rasch Model</td>
<td>6</td>
</tr>
<tr>
<td>Interpretation of the Model Parameters</td>
<td>11</td>
</tr>
<tr>
<td>Summary and Conclusions</td>
<td>13</td>
</tr>
<tr>
<td>Models for Performance on Hierarchically Structured Training Materials</td>
<td>14</td>
</tr>
<tr>
<td>The Module Characteristic Curve Model</td>
<td>16</td>
</tr>
<tr>
<td>Summary and Conclusions</td>
<td>18</td>
</tr>
<tr>
<td>References</td>
<td>20</td>
</tr>
</tbody>
</table>
Since the 1950's, there has been increasing interest in psychological and educational measurement that is based upon probalistic models of the interaction between a person and a test item. These model-based procedures demonstrate how strong assumptions can be used to gain increased control over the measurement process. For example, using item response theory (IRT), the precision of measurement at every point along an ability scale can be determined. Also, items can be selected from a pool to form a test with any desired level of precision at any point on the score scale.

The strong assumptions needed for these model-based procedures are basically that the probabilistic model that has been selected accurately reflects the test data, and that local independence holds for the model. This latter assumption means that the response to one item does not affect the response to another item, and that the response by one person does not affect the response by another person.

Most of the current models assume that the measuring instrument measures only a single trait (Rasch, 1960; Lord, 1952; Birnbaum, 1968). For many tests, this assumption is at least approximated, and for other tests, it is unlikely to be met at all. Most of the current models also are limited to describing a person's response to a single item. In some cases this limitation may make it difficult to solve some measurement problems.

The purpose of the research done on this contract was to extend the types of models available for model-based measurement. Two types of extensions were considered. The first was an extension of item response theory models to the
case where the measurement device was not assumed to be measuring a single dimension. These models were labelled multidimensional item response theory (MIRT) models.

The second type of extension was to cases where sets of related items were considered as a unit. These related sets of items were assumed to be measuring educational constructs that could be arranged into a hierarchy that facilitated learning. These models could be used to determine the interrelationship between the constructs in the hierarchy and the level that must be reached on each construct before a person should be moved on to the next higher level of the hierarchy. Models for tests used with hierarchically arranged instructional units were labelled models for hierarchically structured tests (HST).

The approach taken to develop and evaluate the MIRT and HST models was to first logically evaluate the characteristics of potential models, then to develop estimation procedures for the parameter of the models, and finally to evaluate the models on their ability to describe real test data. These steps were performed separately for a wide class of models of each type. The results of the research will now be described for each type of model, with the analysis of the MIRT models being presented first. Only a summary of the outcome of the research will be presented here, but references will be made to papers and technical reports that contain the details of the research efforts.

The Development and Evaluation of MIRT Models

The class of possible multidimensional, probabilistic models of the interaction between a person and a test item is essentially infinite in
size. Any expression that maps a vector of abilities into a probability could be considered as a MIRT model.

Therefore, the first step in the research effort was to limit the possible models to a manageable subset. This was done by reviewing the literature to determine what MIRT models had been proposed. The review identified three general classes of models that had been suggested for use with multidimensional data.

The first of the classes of models considered were extensions of the general model proposed by Rasch (1961). This model, in its most general form, is given by

\[
p(x_{ij} | \theta_j, \sigma_i) = \frac{1}{\gamma(\theta_j, \sigma_i)} e^{[\phi(x_{ij})' \theta_j + \psi(x_{ij})' \sigma_i + \theta_j' x(x_{ij}) \sigma_i + \rho(x_{ij})]}
\]

where \( p(x_{ij} | \theta_j, \sigma_i) \) is the probability of response \( x_{ij} \) given the values of vector parameters \( \theta_j \) and \( \sigma_i \); \( \theta_j \) is a vector of parameters that describes the characteristics of person \( j \); \( \sigma_i \) is a vector of parameters that describes item \( i \); \( \gamma(\theta_j, \sigma_i) \) is a normalizing function defined by

\[
\gamma(\theta_j, \sigma_i) = \frac{1}{x_{ij}} e^{[\phi(x_{ij})' \theta_j + \psi(x_{ij})' \sigma_i + \theta_j' x(x_{ij}) \sigma_i + \rho(x_{ij})]}
\]

that ensures that the sum of the probabilities of the responses to this item is equal to 1.0; \( \phi(x_{ij}) \) is a vector of scoring weights that indicates the value to be given to each response to the items when considering the estimation of the ability parameters; \( \psi(x_{ij}) \) is a vector of scoring weights that indicates the value to be given to each response to the item when considering the estimation of item parameters; \( x(x_{ij}) \) is a matrix of scoring weights that indicates the value to be given to different products of the
elements of $\delta_j$ and $\sigma_i$; and $p(x_{ij})$ is a constant that is used to set the origin of the linear function defined by the exponent. This equation defines a very general class of models that specifies the dimensionality of the complete latent space by a linear function in the exponent of the logistic model form. Note that this model allows one ability to compensate for another in the metric of $\theta_j$. That is, a high value of $\theta_{j1}$ can compensate for a low value of $\theta_{jn}$ in the linear function of $\theta_j$ defined by

$$\psi_1(x_{ij})\theta_{j1} + \psi_2(x_{ij})\theta_{j2} + \cdots + \psi_m(x_{ij})\theta_{jm}$$

The same type of linear compensation is present for the item parameters.

The second class of models considered was proposed by Mulaik (1972). This class of models is of the form

$$P(x_{ij} | \theta_j, \sigma_i) = \frac{\prod_{k=1}^{m} e^{(\theta_{jk} + \sigma_{ik})x_{ij}}}{1 + \prod_{k=1}^{m} e^{(\theta_{jk} + \sigma_{ik})}}$$

where $x_{ij} = 0, 1$; $m$ is the number of dimensions; and all of the other terms have been defined previously. This model specifies the dimensionality of the complete latent space as a sum of exponential terms. Ability and item parameters can also compensate for each other in this model, but the compensation occurs on an exponential scale. An interesting point to note is that if each exponent is zero in this model, the probability of a correct response is $m/(m+1)$. Thus, as the number of dimensions, $m$, increases, the
probability of a correct response increases unless all of the person and item parameters are rescaled. For the model presented in Equation 1, the probability is always .5 when the exponent is zero.

The third class of models that was considered was proposed by Sympson (1978) and in a slightly different form by Whitely (1980). This class of models is of the general form given by

$$P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + (1-c_i) \prod_{k=1}^{m} \frac{e^{a_{ik}(\theta_{jk} - b_{ik})}}{1 + e^{a_{ik}(\theta_{jk} - b_{ik})}}$$

(5)

where $a_i$ is a vector of discrimination parameters, $b_i$ is a vector of difficulty parameters, $c_i$ is the lower asymptote of the probability function, and all of the other terms have been defined previously. This class of models determines the probability of a response based on abilities in a multidimensional space as the product of a series of probability like terms. These terms are, in effect, the probability of the response to the item if the item only required the one dimension. The overall probability is the product of the probabilities on each dimension. If the exponent is zero on each dimension, the probability will be $c_i + (1 - c_i) (0.5)^m$. Thus, the probability of a correct response will be reduced as each additional dimension is included, unless the parameters are rescaled for each level of dimensionality.

Since the models given in Equations 4 and 5 both require a rescaling of the ability scales with each change in dimensionality, and because both of these models present some very difficult problems in parameter estimation, they were removed from initial consideration and the model presented in Equation 1 became the focus of research effort.
Analysis of the General Rasch Model

The model presented in Equation 1 defines a very rich class of special cases. By selectively setting the weight functions to zero, many different possible models can be derived, each of which have different properties. Each of these special cases was studied both through a mathematical analysis of the equation for each model and through a statistical analysis of simulated data generated using each model. The results of these analyses were reported in a technical report and in a series of papers presented at professional meetings. The full references to the report and the papers are given below.


The results of these analyses showed that two special cases of the general Rasch were capable of modeling realistic multidimensional item response data. The first case uses only the \( \Theta_j^i \chi(x_{ij})\sigma_i \) and \( \Psi(x_{ij})'\sigma_i \) terms.
of the general model. The weights for the other terms were set to zero. The model for this case is given by

$$P(x_{ij} | \theta_j, \sigma_i) = \frac{1}{\gamma(\theta_j, \sigma_i)} e^{\sum_{k=1}^{m} a_{ik} \theta_{jk} + \sum_{k=1}^{m} \sigma_{i,m+k}}$$

(6)

where the symbols have been defined earlier. This form of the model can be written in the more familiar form given by

$$P(x_{ij} | \theta_j, \sigma_i, \Omega_i) = \frac{e^{\sum_{k=1}^{m} a_{ik} \theta_{jk} + \Omega_i}}{1 + e^{\sum_{k=1}^{m} a_{ik} \theta_{jk} + \Omega_i}}$$

(7)

where $a_{ik} = \sigma_{ik}, \Omega_i = -\frac{1}{2} \sum_{k=1}^{m} a_{ik} \sigma_{ik} \Omega_i, \Omega_i = 1 + e^{\sum_{k=1}^{m} a_{ik} \theta_{jk} + \Omega_i} = \gamma(\theta_j, \sigma_i)$

and $a_{ik}$ and $b_{ik}$ can be interpreted as the $a$- and $b$-parameters from unidimensional IRT models. Equation 7 can also be thought of as a multidimensional extension of the two-parameter logistic model; therefore, it has been labelled the M2PL model.

The second special case of the general Rasch model that was found to model multidimensional item response data uses only the $\phi(x_{ij})'\theta_j$ and $\psi(x_{ij})'\sigma_i$ terms from the general model. This model is of the form

$$P(x_{ij} | \theta_j, \sigma_i) = \frac{1}{\gamma(\theta_j, \sigma_i)} e^{\phi(x_{ij})'\theta_j + \psi(x_{ij})'\sigma_i}$$

(8)
where all of the terms have been defined previously. This model has been labelled the "cluster model" because in order for it to model multidimensional data, $x_{ij}$ must be the response string for a cluster of items rather than the response to a single item. If the item cluster contains two dichotomously scored items, the possible $x_{ij}$ responses would be 0,0; 0,1; 1,0; and 1,1. For each of these responses, a different weight function would be available for the $\theta$- and $\sigma$-vectors.

Although the cluster model was very promising, it had one difficulty that made it less attractive. In order to use the model, items had to be clustered, and no rigorous means for doing the clustering has been developed. Therefore, research efforts concentrated on the M2PL model.

Estimation of Model Parameters

In order for a model to be useful, it must be possible to estimate the parameters of the model. Once the M2PL model was selected as the model for further research efforts, work was begun on developing procedures for estimating the model parameters. Two different approaches were taken to solve the estimation problem: (a) unconditional maximum likelihood, and (b) conditional maximum likelihood. Once computer programs were developed for these two approaches, they were validated using both simulated test data generated from the M2PL model, and real test data that were selected because of their multivariate properties. The estimation procedures and results of the program validation studies were presented in the publications and papers listed below.


The results of these studies showed that both the unconditional and conditional maximum likelihood procedures could be used to estimate the item and ability parameters of the M2PL model, but that the unconditional maximum likelihood procedure required somewhat less computer time. However, both procedures require fairly extensive computer facilities, and as the number of dimensions in the model increased, the computer time required became prohibitive. It was clear that improved estimation procedures were needed if the M2PL model was to be widely used.

The validation of the estimation procedures yielded uniformly good results when simulated test results were used. However, when real test data were analyzed, the results were inconsistent. Some studies gave readily interpretable results that were in many ways similar to factor analytic results. In other studies anomalies appeared, such as highly negatively correlated ability estimates that suggested that added constraints were needed to control the estimation process.

In order to study the estimation process in more detail, the M2PL procedure was used to analyze simulated test data that had been produced using a multivariate ability distribution that had varying degrees of correlation between the abilities. The results of the study were presented in the following report.

The study showed that the dimensionality of both the items and the examinee population was important in interpreting the results of an M2PL analysis. If each item were a relatively pure measure of an ability, the procedure obtained good estimates of the ability parameters, even when they were correlated. But, as the correlation between ability estimates increased, there was some deterioration of the accuracy of the estimates. When each item measured more than one ability, the effect of correlated abilities was more extreme. As the correlation between abilities increased, the M2PL solution tended to collapse to a single dimension. The results seemed to imply the need for procedures for oblique rotations to improve the recovery of the ability dimensions.

Interpretation of the Model Parameters

When a MIRT model is used, estimates can be obtained for the ability and the item parameters. The ability parameter estimates can be interpreted in a fairly straightforward manner as the amount of ability a person has on each dimension. The item parameter estimates, however, do not have the same intuitive meaning. Therefore, a major part of this project dealt with determining the MIRT model analogs to the unidimensional IRT item parameters and the measures of quality, such as item and test information. The results of the work in this area were presented in the following documents.


Initial work in this area concentrated on deriving a direct generalization of the interpretations of the difficulty and discrimination parameters and item and test information from the unidimensional item response theory models to the MIRT models. Since the difficulty of an item was defined for the unidimensional models as the point on the ability scale corresponding to the point of inflection of the item characteristic curve, multidimensional difficulty was conceptually thought of as the point of inflection of the multidimensional item response surface (IRS). An analysis of this approach quickly made two important points evident. First, for an IRT there is not a single point of inflection, but rather a locus of points of inflection. Depending upon the MIRT model and the dimensionality being considered, this locus of points of inflection could be a straight line, a curve, a hyperplane, or a hypersurface. The complexity of the locus of points of inflection made its practical application difficult.
The second point that became evident was that the locus of points of inflection could change with the direction taken relative to the surface in the multidimensional space. This is a direct consequence of the fact that the slope at a point on the IRS is different in different directions. The direction in the space is one way of indicating the composite of abilities that is of interest.

In order to take these two points into account, a definition of multidimensional difficulty was derived that was based upon a vector conceptualization. The multidimensional difficulty of an item was defined as the direction from the origin of the multidimensional space to the point of steepest slope and the distance from the origin to the point of steepest slope. Discrimination of an item was related to the slope in the difficulty direction at the point of the steepest slope. Information was also given a directional interpretation. For a group centered at the origin of the space, an item is most informative in the difficulty direction. The item information can also be determined in any other direction, but the maximum information will be less than in the direction indicated by the multidimensional difficulty.

The definitions of multidimensional difficulty, discrimination, and information are general enough that they apply to any MIRT model that is monotonically increasing in probability with an increase in any ability dimension. The definition also includes the unidimensional definitions as special cases.

Summary and Conclusions

This portion of the research project accomplished several important tasks in the development of MIRT. A number of models were analyzed and the
multidimensional extension of the two-parameter logistic model was selected as a promising model for future work. Estimation procedures were developed for this model and the results were validated using simulated and real test data. A theoretical foundation was laid for an interpretation of the item parameters of the MIRT models, and definitions of multidimensional item difficulty, discrimination, and information were developed. At this point, a sufficient framework has been developed to make multidimensional item response theory a viable technique.

Although substantial advances have been made in the area of MIRT, even more work is left to be done. The current estimation programs require excessive amounts of computer time when more than two or three dimensions are specified for a model. Work needs to be done to make estimation of the parameter more efficient. Procedures are needed to determine the appropriate number of dimensions for a set of test data, and procedures for indicating the fit of the models to the data are needed. A related question is whether the M2PL model is an accurate representation of the interaction between a person and an item. This model implies that one ability can compensate for another. Perhaps a model of this type is not appropriate. These and other questions will be addressed in future work.

Models for Performance on Hierarchically Structured Training Materials

Programs of instruction are often composed of many short, homogenous instructional units that have been arranged according to the logical interrelationships of the content. In many cases, short tests are given to determine each student's level of competence on a unit of instruction, and the
scores on the tests are used to route the students through the units of instruction. The purpose of this component of the project was to evaluate an IRT-type model that had potential for assisting in determining the interrelationships between the instructional units and in determining the decision points that should be used with each unit test to minimize routing error. The model treats each unit, or module, of instruction as a complex item and hypothesizes a particular mathematical form for the interrelationship between performance on one module and the probability of successfully passing the next module in the instructional program.

The first step in the evaluation of this model for performance in instructional programs was to review the literature in the area called "learning hierarchies" to determine what procedures were currently being used to evaluate the interrelationships between units of instruction and to set passing scores on the unit tests. The information obtained from the review would serve as a basis for comparison for the results obtained from the proposed model. The review of the literature was presented in the following report.


The review of the literature indicated that there were two general types of procedures that had been used to indicate the relationships between instructional units; those based on coefficients of dependence, and those based on a more complete description of the relationships between units of instruction, usually a mathematical model. The procedures based on
coefficients of dependence were found to provide insufficient information for validating the sequence of instructional units, or for setting passing scores. The procedures based on mathematical models were found to have more potential, but the currently available procedures did not seem to meet the needs of instructional programs. There seemed to be a clear need for a procedure that could be used to arrange units of instruction into a hierarchy based upon the prerequisite knowledge required by each unit of instruction, and that could be used to set passing scores for each unit that would improve the efficiency and accuracy of the routing process. The model proposed and evaluated during this research effort was designed to perform these functions.

The Module Characteristic Curve Model

The basic idea behind the proposed model for the interrelationship between modules of instruction is that if two modules form a learning hierarchy, performance on the higher level instructional module is dependent upon prerequisite knowledge obtained from the lower level module of instruction. Thus, if sufficient knowledge has not been gained on the lower level module, a high level of performance cannot be exhibited on the higher level module of instruction. This implies that success on the higher module is related to the level of performance on the lower module.

The probabilistic model that was hypothesized to describe the relationship between hierarchically related instructional modules is given by

$$p_j(\theta_{ik}) = c_j + (1 - c_j - e_j) \frac{e^{D_{aj}(\theta_{ik} - b_j)}}{1 + e^{D_{aj}(\theta_{ik} - b_j)}}$$  (9)
where \( P_j(\theta_{ik}) \) is the probability of passing module \( j \) given level of performance \( \theta_{ik} \) of examinee \( i \) on prerequisite module \( k \), \( c_j \) is the probability of passing module \( j \) if the examinee has not acquired any knowledge in module \( k \), \( e_j \) is the probability of passing module \( j \) if the examinee has mastered module \( k \), \( D = 1.7 \), \( a_j \) is a parameter related to the strength of the relationship between the two modules, and \( b_j \) is the difficulty of the passing score used on module \( j \). This model predicts the probability that an examinee will pass module \( j \) based on his/her performance on module \( k \).

In order to use this model, estimates of achievement are first obtained on module \( k \). This can either be done by analyzing the module \( k \) test using an IRT model, or by converting the raw scores on module \( k \) to z-scores. These achievement measures are then used as known values and the model parameters are estimated using a maximum likelihood estimation procedure.

A very low \( a \)-parameter estimate is an indication that the two modules are not very highly related. A high \( a \)-value indicates that knowledge on module \( k \) is very important for module \( j \). A high estimate for the \( c \)-parameter indicates that examinees can perform well on module \( j \) even without mastering module \( k \). A low \( c \)-value indicates that an examinee cannot perform well on module \( j \) unless knowledge has been acquired on module \( k \).

Estimates of the \( e \)-parameter indicate the maximum probability of passing the \( j \) module given that the examinee has mastered module \( k \). Low values indicate that module \( k \) contains only a small portion of the information needed to pass module \( j \). High values indicate that module \( k \) includes most of the information needed to pass module \( j \).

The \( b \)-parameter estimates indicate the point on the module \( k \) scale that best distinguishes between persons who pass or fail module \( j \). This point will change with changes in the passing score on module \( j \). The point on the module
k scale specified by the b-parameter is the suggested decision point on module k for routing to module j if misclassification errors in either direction are considered equally serious.

In order to evaluate this model, it was applied to both simulated and real test data to determine whether the estimation procedures worked properly, and whether it realistically represented actual test results. The outcome of these studies were presented in the following documents.


The studies showed that the parameters of the model could be accurately estimated and that for one set of real test data, the model gave very reasonable results. There was some indications, however, that the upper and lower asymptote parameters might not be needed. It may be possible to simplify the model to a two-parameter logistic form.

**Summary and Conclusions**

A model for the relationship between modules of instruction that are hierarchically related was proposed and evaluated using both simulated and real test data. The results of the studies showed that the model parameters could be accurately estimated and that the model was a good representation of
real test data that should be hierarchically related. However, the upper and lower asymptotes did not appear to be needed for the particular real data set that was analyzed. Further studies need to be done to determine whether this is a general finding applicable to all hierarchically arranged modules, or whether it only applies to this case. If the c- and e-parameters are not needed, the model can be simplified to a two-parameter logistic model.

One problem with the use of the model became evident with the analysis of the real test data. In order to accurately estimate the parameters of the model, examinees must be routed to the higher level unit of instruction even when they have not performed well on the lower level unit. This is poor educational practice and, in many cases, this data collection procedure cannot be followed. This makes it difficult to obtain data for use in estimating the parameters of the model. It may be that the model will have to be modified to accommodate the routing procedures that are currently being used in modularized instructional programs.
References


Distribution List

Personnel Analysis Division
AF/MPXA
5C360, The Pentagon
Washington, DC 20330

Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

Air Force Office of Scientific Research
Life Sciences Directorate
Bolling Air Force Base
Washington, DC 20322

Dr. Robert Ahlers
Code N711
Human Factors Laboratory
NAVTACEQUIPCEN
Orlando, FL 32813

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

Technical Director
Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Special Assistant for Projects
OASN(M&RA)
5D800, The Pentagon
Washington, DC 20350

Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

Dr. Patricia Baggett
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

CDR Robert J. Biersner, USN
Naval Biodynamics Laboratory
P. O. Box 29407
New Orleans, LA 70189

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

Dr. Werner Birke
Personalstammamt der Bundeswehr
D-5000 Koeln 90
WEST GERMANY

Code N711
Attn: Arthur S. Blaiwes
Naval Training Equipment Center
Orlando, FL 32813

Dr. Robert Breaux
Code N-095R
NAVTACEQUIPCEN
Orlando, FL 32813

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

Dr. Patricia A. Butler
NIE Mail Stop 1806
1200 19th St., NW
Washington, DC 20208
American College Testing Program/Reckase

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

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

Dr. Robert Carroll
NAVOP 01B7
Washington, DC 20370

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

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

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

Scientific Advisor to the DCNO (MPT)
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Chief of Naval Education and Training
Liaison Office
AFHRL
Operations Training Division
Williams AFB, AZ 85224

Assistant Chief of Staff
Research, Development, Test, and Evaluation
Naval Education and Training Command (N-5)
NAS Pensacola, FL 32508

Office of the Chief of Naval Operations
Research Development & Studies Branch
NAVOP 01B7
Washington, DC 20350

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

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

CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940

CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217-5000

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

Dr. Dattprasad Divgi
Syracuse University
Department of Psychology
Syracuse, NY 13210

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

Dr. Fritz Drasgow
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820
19 April 1985

American College Testing Program/Beckase

Program Manager for Manpower, Personnel, and Training
NAVMAT 0722
Arlington, VA 22217-5000

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

Dr. William E. Nordbrock
FMC-ADCO Box 25
APO, NY 09710

Dr. Melvin R. Novick
356 Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

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

Library
Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. Harry F. O'Neil, Jr.
Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

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

Mathematics Group
Office of Naval Research
Code 744MA
800 North Quincy Street
Arlington, VA 22217-5000

Office of Naval Research
Code 442PT
800 N. Quincy Street
Arlington, VA 22217-5000
(5 Copies)

Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217-5000

Commanding Officer
Army Research Institute
ATTN: PERI-BR (Dr. J. Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

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

Dr. Randolph Park
AFHRL/MOAN
Brooks AFB, TX 78235

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

Dr. Roger Pennell
Air Force Human Resources Laboratory
Lowry AFB, CO 80230

Administrative Sciences Department
Naval Postgraduate School
Monterey, CA 93940
Martha Stocking  
Educational Testing Service  
Princeton, NJ 08541

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

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

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

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

Mr. Brad Sympson  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. John Tangney  
AFOSH/NL  
Bolling AFB, DC 20332

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

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

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

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

Dr. Robert Tsutakawa  
Department of Statistics  
University of Missouri  
Columbia, MO 65201

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

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

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

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

Dr. Howard Wainer  
Division of Psychological Studies  
Educational Testing Service  
Princeton, NJ 08540

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

Mr. Thomas A. Warm  
Coast Guard Institute  
P. O. Substation 18  
Oklahoma City, OK 73169

Dr. Brian Waters  
HumRRO  
300 North Washington  
Alexandria, VA 22314

Dr. Edward Wegman  
Office of Naval Research  
Code 411  
800 North Quincy Street  
Arlington, VA 22217-5000