A review of the literature was made of the various procedures for sequencing instructional units in individualized instruction programs. Procedures for validating sequences of instructional units, or learning hierarchies, were found to fall into two general categories. One category included procedures based on coefficients of dependence, while the other category contained procedures based on more complete descriptions of the relationships between units of instruction, usually a mathematical model. Procedures based on coefficients were found to provide insufficient information for optimal sequencing of units. Procedures based on mathematical models were also found to be inadequate, but the model-based approach was found to have potential for dealing with the problem of sequencing. Recommended future research emphasizes improving model-based procedures so that they provide more information useful for determining the optimal order of instructional units. (Author/PN)
The Validation of Learning Hierarchies

Mark D. Reckase
and
Robert L. McKinley

Research Report ONR 82-2
September 1982

The American College Testing Program
Resident Programs Department
Iowa City, IA 52243

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.
<table>
<thead>
<tr>
<th>REPORT NUMBER</th>
<th>ONR82-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE (and Subtitle)</td>
<td>The Validation of Learning Hierarchies</td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Mark D. Reckase and Robert L. McKinley</td>
</tr>
<tr>
<td>PERFORMING ORGANIZATION NAME AND ADDRESS</td>
<td>American College Testing Program, P.O. Box 168, Iowa City, IA 52243</td>
</tr>
<tr>
<td>CONTRACT OR GRANT NUMBER(a)</td>
<td>N00014-81-K0817</td>
</tr>
<tr>
<td>PERIOD COVERED</td>
<td>Technical Report</td>
</tr>
<tr>
<td>ABBREVIATION</td>
<td>Approval for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.</td>
</tr>
<tr>
<td>KEYWORDS</td>
<td>Learning hierarchies, Individualized instruction, Validation of hierarchies</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>A review was made of the various procedures for sequencing instructional units in individualized instruction programs and recommendations were made concerning the direction future research should follow. Procedures for sequencing units of instruction were found to fall into two general categories. One category included procedures based on coefficients of dependence, while the other category contained procedures based on more complete descriptions of the relationships between units of instruction, usually a mathematical model. Procedures based on coefficients were...</td>
</tr>
</tbody>
</table>
found to provide insufficient information for optimal sequencing of units. Procedures based on mathematical models were also found to be inadequate, but the model-based approach was found to have potential for dealing with the problem of sequencing. Recommended future research emphasizes improving model-based procedures so that they provide more information useful for determining the optimal order of instructional units.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>The Work of Gagné.</td>
<td>1</td>
</tr>
<tr>
<td>Procedures for Validating Hierarchies</td>
<td>2</td>
</tr>
<tr>
<td><strong>Coefficient-Based Procedures.</strong></td>
<td>3</td>
</tr>
<tr>
<td>The PPT Statistic</td>
<td>3</td>
</tr>
<tr>
<td>Variations on the PPT Statistic</td>
<td>3</td>
</tr>
<tr>
<td>Non-PPT Based Statistics</td>
<td>5</td>
</tr>
<tr>
<td>Summary of Coefficient Procedures</td>
<td>6</td>
</tr>
<tr>
<td><strong>Model-Based Procedures.</strong></td>
<td>7</td>
</tr>
<tr>
<td>Proctor Model</td>
<td>7</td>
</tr>
<tr>
<td>White and Clark Model</td>
<td>8</td>
</tr>
<tr>
<td>Dayton and Macready Model</td>
<td>10</td>
</tr>
<tr>
<td>Loglinear Models</td>
<td>12</td>
</tr>
<tr>
<td>Summary of Model-Based Procedures</td>
<td>13</td>
</tr>
<tr>
<td><strong>Evaluation of Procedures for Validating Hierarchies</strong></td>
<td>13</td>
</tr>
<tr>
<td>A Program for Constructing and Validating Hierarchies</td>
<td>13</td>
</tr>
<tr>
<td>Selection Criteria</td>
<td>14</td>
</tr>
<tr>
<td>Evaluation of the Coefficient Approach</td>
<td>15</td>
</tr>
<tr>
<td>Evaluation of the Model Approach</td>
<td>15</td>
</tr>
<tr>
<td><strong>Summary and Recommendations</strong></td>
<td>16</td>
</tr>
<tr>
<td>Summary</td>
<td>16</td>
</tr>
<tr>
<td>Recommendations</td>
<td>16</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>17</td>
</tr>
</tbody>
</table>
The Validation of Learning Hierarchies

One of the most rapidly growing areas in the field of education today is the area of individualized instruction—instruction in which content, organization, or pacing is modified for each individual. Programs of individualized instruction have been implemented in such areas as electronics (Peiper, Swegey, and Valverde, 1970), engineering (Schure, 1965), mathematics (Bushnell, 1966), psychology (Kulik, 1972), and vocational and technical education in the military (Impelliterri and Finch, 1971). These programs employ several different types of educational technology, including personalized systems of instruction (PSI), computer-assisted instruction (CAI), individually prescribed instruction (IPI), and programmed instruction (PI).

Although individualized instruction comes in many forms, all of these forms share the same basic design. All are essentially sequences of instructional units through which subjects are routed by means of a series of tests. The way in which these units are sequenced is one of the more crucial components of any individualized instruction program. The units should be arranged in such a fashion that prerequisite material is covered first, followed by more advanced material. How this ordering is determined has been the topic of considerable research. The purpose of this report is to evaluate the various approaches to sequencing instructional units that have been reported in the literature, and to make recommendations on the direction future research should follow.

Most of the procedures that have been proposed for sequencing instructional units have been based on the work of Gagné (1962), and have involved the use of various techniques for validating learning hierarchies constructed using his task analysis methodology. Procedures for validating hierarchies have generally taken one of two approaches. One approach has been to simply compute a coefficient that measures the strength of the hierarchical relationships, and on the basis of the coefficient accept or reject the hierarchy. The other approach has been to more completely describe the relationships among the instructional units, usually by means of a mathematical model, and to decide on the basis of the description whether the relationships are of the desired nature. Both approaches will be discussed, as will be a number of examples of each approach. First, however, a discussion of Gagné's work on learning hierarchies will be presented.

The Work of Gagné

In a series of articles dating from 1961, Gagné, and his coworkers investigated the learning of mathematical skills. A hypothesis originating from this research was that a class of tasks necessary to attain a learning outcome could be sorted into a hierarchy of sets of tasks, each set being a group of tasks at the same level in the hierarchy. The structure of the hierarchy would be such that there could be positive transfer of learning from one learning set to a higher level learning set until the final learning outcome was achieved. Gagné (1962) proposed that one could establish a learning hierarchy by determining the skills an individual would have to possess in order to attain the learning outcome. One would then make the same determination for each of those skills. This process would be continued...
until one developed a hierarchy containing the most basic skills as the
learning set at the lowest level. If the established hierarchy were correct,
then, positive transfer from each lower learning set to the next higher
learning set would be promoted (Gagné, 1962).

The successful establishment of a learning hierarchy results in a vertical structure of learning sets, with any high level learning set having one or more immediately prerequisite learning sets. Figure 1 presents an example of a derived learning hierarchy. In this hierarchy, the terminal task has three prerequisite tasks, which in turn have six second level prerequisite tasks. Mastery of the lower tasks is required before a final task can be performed.

Gagné (1968) identified two characteristics necessary for the successful establishment of a learning hierarchy. One such characteristic is that of sequencing: a learner who is able to perform successfully a higher level set of tasks should also be able to perform all lower level sets of tasks in the hierarchy. The other characteristic is transfer: attainment of a lower level learning outcome should increase the probability of successful attainment of a higher level learning set. The existence of these two characteristics establishes that there is an ordered relationship among the learning sets within the hierarchy. If a learning hierarchy has been successfully established, it should be possible to validate the hierarchy by demonstrating the presence of these characteristics. Most of the procedures for validating hierarchies, including Gagné's own procedure, are based on this premise. Those procedures will now be discussed.

Procedures for Validating Hierarchies

As was previously mentioned, there have been two basic approaches to validating hierarchies constructed using task analysis procedures, one based on coefficients and one based on more complete descriptions of the relationships among sets of learning tasks. These two approaches will now be discussed in more detail, and several examples of each will be presented. After the approaches and examples have been discussed, an evaluation of the approaches will be presented.
Coefficient-Based Procedures

The basic objective of the coefficient approach for validating learning hierarchies is to compute the value of a coefficient that will measure the strength of the ordered relationships among learning sets or tasks. A number of coefficients have been proposed for this purpose, most of which are variations on the proportion of positive transfer (PPT) statistic proposed by Gagne and Paradise (1961). The PPT statistic will be described first, and then the ways in which it has been modified will be discussed. Afterward, several coefficients not based on the PPT statistic will be discussed.

The PPT Statistic The PPT statistic is based on the four pass/fail combinations possible for two learning tasks. If Task 2 follows (is dependent on) Task 1 and $N_{00}$ is the number of subjects failing on both tasks, $N_{10}$ is the number of subjects succeeding on Task 1 but failing on Task 2, and $N_{11}$ is the number of subjects succeeding on both tasks, then the PPT statistic is given by

$$PPT = \frac{N_{00} + N_{11}}{N_{00} + N_{10} + N_{11}}$$

This statistic is used to measure the level of positive transfer from Task 1 to Task 2. If the level of positive transfer is high, then fewer subjects will succeed on Task 1 without succeeding on Task 2 than would have if the level of transfer had been low. Thus, the greater the positive transfer, the smaller is $N_{10}$. As $N_{10}$ decreases, the PPT statistic increases. When $N_{10}$ is 0, the PPT statistic has a value of 1.0, which is the maximum value it can take on.

The PPT statistic appears to be a reasonable indicator of the level of positive transfer between two tasks. However, positive transfer does not necessarily indicate that there is a hierarchical relationship between the two tasks. It is possible that the positive transfer from Task 2 to Task 1 is as great as the positive transfer from Task 1 to Task 2. The PPT statistic does not indicate the direction of the positive transfer. In fact, if the tasks are reversed so that Task 2 becomes the first task and Task 1 becomes the second task, an equal or higher value may be obtained for the PPT statistic as was obtained before the reversal, but in this situation the statistic indicates positive transfer from the subsequent task to the precedent task. Because of this limitation on the interpretation of the PPT statistic, a number of alternative statistics, most of which are variations of the PPT statistic, have been proposed.

Variations on the PPT Statistic The Commission on Science Education of the American Association for the Advancement of Science (Walbesser, 1968) proposed a procedure for validating hierarchies that employs three statistics rather than one. These three statistics are the consistency, the adequacy, and the completeness ratios. These ratios use the same four pass/fail combinations that were used for the PPT statistic.
The consistency ratio (CR) is given by

\[ CR = \frac{N_{11}}{N_{11} + N_{10}} \]

where the terms are as defined for the PPT statistic. As can be seen, the CR statistic is the same as the PPT statistic, except that the \( N_{00} \) term is omitted. The \( N_{00} \) term is omitted because, while it is consistent with the hypothesized hierarchical relationship, it is not an indication of positive transfer (Walbesser and Eisenberg, 1972).

Consistency is a necessary but not sufficient condition for a valid hierarchy, according to Walbesser (1968). In addition, adequacy and completeness must be considered. Adequacy refers to how often the learner has achieved a behavior after the relevant subordinate behavior has been attained. The adequacy of a hierarchy is measured using the adequacy ratio (AR), which is given by

\[ AR = \frac{N_{11}}{N_{11} + N_{01}} \]

Completeness is an indication of the number of examinees that have reached the terminal behavior relative to those that have made no progress. High consistency and adequacy ratios are misleading if only small numbers of examinees acquire the terminal behavior and at least some subordinate behaviors. Thus, a large value for \( N_{00} \) would be evidence of incomplete instruction. The completeness ratio (COR) is given by

\[ COR = \frac{N_{11}}{N_{11} + N_{00}} \]

A procedure for validating hierarchies proposed by Walbesser and Eisenberg (1971) employs the consistency, adequacy, and completeness ratios, and in addition uses two other coefficients. These coefficients are the inverse consistency ratio (ICR) and the inverse adequacy ratio (IAR). While consistency indicates that the acquisition of the terminal behavior implies the acquisition of subordinate behaviors, the inverse consistency ratio measures the extent to which no acquisition of the terminal behavior implies nonacquisition of subordinate behaviors. The inverse consistency ratio is given by

\[ ICR = \frac{N_{00}}{N_{00} + N_{01}} \]

The adequacy ratio measures the extent to which the acquisition of all subordinate behaviors implies acquisition of the terminal behavior, while the inverse adequacy ratio indicates the degree to which nonacquisition of the subordinate behavior implies nonacquisition of the terminal behavior. The inverse adequacy ratio is given by

\[ IAR = \frac{N_{00}}{N_{00} + N_{10}} \]
Non-PPT Based Statistics

Not all of the coefficients that have been proposed for validating hierarchies have been variations on the PPT statistic. For instance, Capie and Jones (1971) used the phi coefficient for validating hierarchies. The phi coefficient is based on the same four pass/fail combinations that were used for the PPT-like statistics. However, rather than computing a ratio of the numbers of subjects in specified groups, a product-moment correlation coefficient is computed.

The Guttman coefficient of reproducibility has also been used to validate hierarchies (Hofman, 1977; Resnick and Wang, 1969). An assumption made when this statistic is used is that the learning tasks are ordered according to their difficulty. If an individual gives five correct responses on 10 items, it is assumed that the individual responded correctly to the five easiest items. This statistic, then, is a measure of the relationship between the response pattern of an individual and the number of correct responses by that individual. If the response pattern is not accurately predicted by the number of correct responses, a perfect Guttman scale is not present. The proportion of responses that follow the predicted pattern is called the coefficient of reproducibility. To the extent that some respondents achieve success on some of the more difficult tasks but fail on less difficult tasks, the coefficient of reproducibility is reduced in magnitude. For a pair of tasks in a hierarchy, the extent to which some respondents achieve success on the superordinate task but fail the subordinate task, the coefficient is reduced. The coefficient of reproducibility for the hierarchy is the average coefficient of reproducibility for all of the pairs of superordinate and subordinate tasks in the hierarchy (Hofman, 1977).

Another statistic that has been used for validating hierarchies is the proportion of disconfirmatory response patterns (Airasian and Bart, 1975). This procedure is based on ordering theory, which provides a basis for determining logical relationships among tasks (Bart and Krus, 1973). If it is assumed that Task 1 is prerequisite to Task 2, then the response pattern (0,1), which indicates success on Task 2 but failure on Task 1, is considered disconfirmatory. The response patterns (0,0), (1,0), and (1,1) are considered to be confirmatory. The proportion of disconfirmatory response patterns (PD) is given by

$$ PD = \frac{N_{01}}{N_{00} + N_{10} + N_{01} + N_{11}} $$

One other statistic that has been used for validating hierarchies is the conditional item difficulty index (Airasian, 1971). The conditional difficulty of an item is computed using only the subjects having response patterns that are predicted from the hierarchy. For instance, for a three-task hierarchy the only response patterns expected are:

(a) 000,
(b) 100,
(c) 110,
and (d) 111.
even though there are eight patterns possible. If \( n_0, n_1, n_2, \) and \( n_3 \) are the number of subjects with response patterns (a), (b), (c), and (d) above, respectively, then only \( n_0 + n_1 + n_2 + n_3 \) subjects will be considered when computing a conditional item difficulty. The conditional difficulty for an item is given by the number of subjects having expected response patterns in which the item was correctly answered, divided by the number of subjects having expected response patterns in which all preceding items are correct. Thus for the first task in the three-task hierarchy the conditional difficulty \((CD)\) is given by

\[
CD = \frac{n_1 + n_2 + n_3}{n_0 + n_1 + n_2 + n_3}
\]

For the second task the conditional difficulty is given by

\[
CD = \frac{n_2 + n_3}{n_1 + n_2 + n_3}
\]

and for the third task the conditional difficulty is given by

\[
CD = \frac{n_3}{n_2 + n_3}
\]

The conditional item difficulty indices computed for a set of tasks help to determine the validity of a hierarchy by indicating the extent to which failure on earlier tasks is predictive of failure on later tasks. For example, if the conditional difficulty of an item is very low given that all preceding items have been correctly answered, the completeness of the sequence is questionable (Airasian, 1971).

Summary of Coefficient Procedures The basic objective of the coefficient approach to validating hierarchies is to compute a coefficient that will indicate the strength of the hierarchical relationships among learning tasks. Many different coefficients have been proposed for this purpose, including: the proportion of positive transfer statistic; the consistency, adequacy, completeness, inverse consistency, and inverse adequacy ratios; the Guttman coefficient of reproducibility; the phi coefficient; the proportion of disconfirmatory response patterns; and the conditional item difficulty index. Table 1 summarizes the coefficients that have been discussed.

One reason why so many coefficients have been proposed is that each coefficient tends to be associated with only one aspect of the complex relationship between tasks in a hierarchy. This has led some researchers to propose the use of multiple coefficients for the validation of hierarchies (Walbesser, 1968; Walbesser and Eisenberg, 1972). Other researchers have proposed that more complete descriptions of the relationship between learning tasks is needed than can be provided by coefficients alone. These researchers have tended to use mathematical models to describe the relationships among learning tasks. This approach to validating learning hierarchies will be discussed next.
Model-Based Procedures

Model-based procedures specify a mathematical model to describe the relationship between performances on different tasks in a hierarchy. This is commonly done using a mathematical expression describing the probability of success on a task given the performance (acquisition vs. nonacquisition) on lower level tasks. The adequacy of the model for describing the relationship between performances on different tasks can be tested statistically. If the performance on a task predicted from the model does not differ significantly from observed data, then the data support the hypothesized structure of the hierarchy and the form of the mathematical model. If the model does not fit the data, the failure may be due to errors in the hypothesized structure of the hierarchy, errors in the form of the mathematical model, or both.

Proctor Model

One of the first mathematical models proposed for validating learning hierarchies was described by Proctor (1970). This model is a probabilistic formulation of Guttman scaling. The items in the proposed hierarchy are assumed to form a Guttman scale. Each response pattern consistent with the Guttman scale is associated with a true ability level. It is assumed that patterns that are inconsistent with the scale are also associated with one of the true ability levels, but that they contain error. The probability that an observed response pattern is associated with a given true ability level is the probability of finding an examinee with that true ability level multiplied by the probability that an examinee with that ability level would give the observed response pattern. This second probability decreases as the difference between the observed pattern and the pure pattern associated with that level of ability increases. The probability of the occurrence of an observed response pattern is the sum of the probabilities of the pattern over each of the true ability levels.

As an example, suppose that every subject in a population belongs to one of the several ability levels, each of which has associated with it a true Guttman type response pattern. For a three item scale there are four true Guttman patterns—(000), (100), (110), and (111). Every subject, then, belongs to one of the four ability levels associated with these patterns. The probability of an observed response pattern for this scale is the sum of four terms—the probability of finding a subject whose true ability level was associated with the (000) pattern and who responded with the observed pattern, the probability of finding a subject who should have responded with the (100) pattern but who responded with the observed pattern, and so on until all four true patterns have been considered. In its mathematical form, the probability of the observed pattern for the three item scale is given by

$$P(x) = \frac{\theta_1}{1 - \alpha} (1 - \alpha)^{4-n_1} + \frac{\theta_2}{1 - \alpha} (1 - \alpha)^{4-n_2} + \frac{\theta_3}{1 - \alpha} (1 - \alpha)^{4-n_3} + \frac{\theta_4}{1 - \alpha} (1 - \alpha)^{4-n_4}, \quad (1)$$

where $x$ is the observed pattern; $\theta_1$, $\theta_2$, $\theta_3$, and $\theta_4$ are the proportions of the population having each true ability level; the $\alpha$ term is the probability of a subject responding to an item in a way inconsistent with the true pattern associated with the subject's true ability; and the $n_1$, $n_2$, $n_3$, and $n_4$ terms are the number of items in the observed pattern inconsistent with the true pattern for each of the four ability levels.
Table 1
A Summary of the Coefficient Procedures for the Validation of Learning Hierarchies

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Proponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of Positive Transfer</td>
<td>Gagne and Paradise (1961)</td>
</tr>
<tr>
<td>PPT Based Statistics</td>
<td></td>
</tr>
<tr>
<td>Consistency Ratio</td>
<td>Walbesser (1968)</td>
</tr>
<tr>
<td>Adequacy Ratio</td>
<td></td>
</tr>
<tr>
<td>Completeness Ratio</td>
<td></td>
</tr>
<tr>
<td>Inverse Consistency Ratio</td>
<td>Walbesser and Eisenberg (1971)</td>
</tr>
<tr>
<td>Inverse Adequacy Ratio</td>
<td></td>
</tr>
<tr>
<td>Non-PPT Based Statistics</td>
<td></td>
</tr>
<tr>
<td>Phi Coefficient</td>
<td>Capie and Jones (1971)</td>
</tr>
<tr>
<td>Coefficient of Reproducibility</td>
<td>Hofman (1977), Resnick and Wang (1964)</td>
</tr>
<tr>
<td>Proportion of Disconfirmatory Response Patterns</td>
<td>Airasian and Bart (1975), Bart and Krus (1973)</td>
</tr>
<tr>
<td>Conditional Item Difficulty</td>
<td>Airasian (1971)</td>
</tr>
</tbody>
</table>

To apply this model, it is assumed that the frequencies of the observed response patterns are distributed multinomially with probabilities given by the model. A test of the fit of the multinomial model (with the probabilities from the model as parameters) to the observed frequencies can be performed.

White and Clark Model. A somewhat different model was introduced by White and Clark (1973). This model tests the hypothesis that all subjects who possess a certain skill form a subset of the group of subjects who possess a second skill. This model is called the C statistic model.

In using this model, a matrix of response frequencies is developed using the format shown in Table 2. The population subgroups are defined as follows:

- $P_0$ = proportion of the population having neither skill,
- $P_B$ = proportion of the population having both skills,
- $P_I$ = proportion of the population having only Skill I, and
- $P_{II}$ = proportion of the population having only Skill II.
Table 2
Response Frequency Matrix for White and Clark Model

<table>
<thead>
<tr>
<th>Number of Lower Skill Questions Correctly Answered</th>
<th>Number of Higher Skill Questions Correctly Answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>P_{20} P_{21} P_{22} n_{12}</td>
</tr>
<tr>
<td>1</td>
<td>P_{10} P_{11} P_{12} n_{11}</td>
</tr>
<tr>
<td>0</td>
<td>P_{00} P_{01} P_{02} n_{10}</td>
</tr>
<tr>
<td>TOTAL</td>
<td>n_{20} n_{21} n_{22} N</td>
</tr>
</tbody>
</table>

Note: N is the total number of subjects in the sample, while $n_{ij}$ is the number of subjects correctly answering $j$ items for skill $i$.

The conditional probabilities of group members having answered the appropriate number of items correctly are:

- $\theta_a = \text{probability of examinee with Skill I answering correctly any item for Skill I}$
- $\theta_b = \text{probability of examinee without Skill I answering correctly any item for Skill I}$
- $\theta_c = \text{probability of examinee with Skill II answering correctly any item for Skill II}$
- $\theta_d = \text{probability of examinee without Skill II answering correctly any item for Skill II}$

The total probability of a cell is thus defined as the product of the probability of the examinee being in a group and the conditional probability of members of that group answering the appropriate number of items correctly. If the various $P$'s and $\theta$'s are known, a probability estimate for each cell in Table 2 can be calculated. Estimates of the $P$'s and $\theta$'s can be calculated using the cell frequencies or the marginal totals using a maximum likelihood procedure (see White and Clark, 1973). In a two-item case, $P_{ij}$ can be derived by substituting the estimates of $P$ and $\theta$ in the equation:

$$P_{ij} = P_i (1 - \theta_i)^2 \theta_j^2 + P_i (1 - \theta_i)^2 \theta_j^2 + P_i (1 - \theta_i)^2 \theta_j^2 + P_i (1 - \theta_i)^2 \theta_j^2 \theta_i$$

Once the probabilities of all the cells have been computed, the probability of the observed distribution under $H_0$ can be calculated using the observed cell frequencies and estimated probabilities as parameters of the multinomial distribution. A test of significance can be performed by summing the
probabilities of all possible distributions which show a deviation from the hypothesis as great or greater than that of the observed distribution (White and Clark, 1973). In order to reduce the amount of computation required, White and Clark (1973) suggest forming the test using only the (0,2) cell. The observed frequency and estimated probability for this cell could be used as parameters for the binomial distribution, and H₀ would be rejected whenever the observed frequency exceeded a critical value of C. The value of C, of course, depends on the desired error rate, the sample size, and the magnitude of the probability estimated for the (0,2) cell.

Dayton and Macready Model A third mathematical model for the validation of learning hierarchies was proposed by Dayton and Macready (1976). This model is essentially a generalization of the Proctor model, and also subsumes the White and Clark model (Dayton and Macready, 1976).

For any K dichotomously scored tasks, the scores on those tasks may be summarized by a column vector, U, composed of 0's and 1's. A score of 0 may arise from an incorrect response or from an omission. The product \( S = U^T U \) is the number of items or tasks successfully completed by a respondent. Assume there exists an a priori hierarchy, and there exists a set of q distinct pattern vectors, \( V_j \), comprised of 0's and 2's, which for the hypothesized hierarchy defines acceptable response patterns. (The values 0 and 2 are used instead of 0 and 1 so that the vector difference \( V - U \) will provide information necessary to compute the values of the exponents in the model.) For example, when \( K = 4 \), a linear hierarchy would be represented by \( q = 5 \) pattern vectors:

\[
V_1 = (0000), \\
V_2 = (000), \\
V_3 = (2000), \\
V_4 = (220), \\
V_5 = (2222).
\]

In the most general form the probabilistic model may be written as:

\[
P(U) = \sum_{j=1}^{q} \left( \frac{K}{\prod_{i=1}^{K} a_{ij}(1 - a_i) b_{ij} c_{ij}(1 - b_i) d_{ij}} \theta_j \right),
\]

(3)

where \( a_{ij}, b_{ij}, c_{ij} \), and \( d_{ij} \) are defined as follows. Let \( g_{ij} \) be the \( i \)th element in \( V_j - U \). Then,
The parameter $\theta_j$ represents the probability that the jth pattern vector occurs. It is the hypothetical population proportion of respondents that achieves Level $j$ of the hierarchy. The parameter $\alpha_i$ represents the probability that a respondent produces a correct response to a task, which relative to a specific pattern vector should not have been correctly completed. The $\alpha_i$ parameter is referred to as the guessing parameter. The parameter $\beta_i$ is the probability that a respondent produces a response that is incorrect which should have been completed correctly relative to a specific pattern vector. The parameter $\beta_i$ is the forgetting parameter. For this model, the occurrence of guessing and forgetting are assumed to be independent across items (Dayton and Macready, 1976).

A restriction on $\theta_j$ is

$$\sum_{j=1}^{q} \theta_j = 1. \tag{4}$$

Also, the parameters $(\theta_j, \alpha_i, \beta_i)$ are meaningful only over the interval 0 to 1 (Dayton and Macready, 1976).

For any given task, three of the four variables $a_{ij}, b_{ij}, c_{ij}$, and $d_{ij}$ will be equal to 0, and one of the four will take on the value of 1. This model allows for $2^n - 1$ independent parameters, where $n$ is the number of items. Once the parameters have been estimated, the Pearson chi-square or likelihood ratio chi-square test may be used to provide a goodness-of-fit test (Dayton and Macready, 1976).
Loglinear Models  Another class of models have been described by Goodman (1975) and Davison (1981). These models are based on an application of loglinear analysis to the problem of discovering or confirming learning hierarchies. Loglinear analysis utilizes data in the form of a contingency table, but is not limited to dichotomous variables, nor to pairwise comparisons (Davison, 1981). For a discussion of loglinear models, see Bishop, Feinberg and Holland (1975).

For simplicity, the presentation of the model described by Goodman (1975) will be limited to the case of three variables, each having three response categories. The variables used could be items (for which the response categories would be item responses), or tests (for which the response categories would be test score categories). The three variables and their response categories form a contingency table, in this case, a three-way contingency table. Cell entries represent frequencies of subjects exhibiting the response vector represented by each cell. That is, the entry in cell \((a, b, c)\) would be the number of subjects responding \(a\) to the first variable, \(b\) to the second, and \(c\) to the third.

Given an a priori or hypothesized sequence among the three variables, the contingency table can be broken down into two classes of cells: those cells representing response vectors compatible with the hypothesized sequence, and those cells representing the vectors inconsistent with the hypothesized sequence. Subjects are divided into \(k + 1\) classes, where \(k\) is the number of cells representing response vectors compatible with the hypothesized sequence. The \((k + 1)\)th class is the group of subjects exhibiting incompatible response vectors. The proportion of subjects in each class is represented by \(P_i\) \((i = 0, 1, \ldots, k)\), where \(P_0\) is the proportion in the class having inadmissible responses.

Within the class of inadmissible cells, which is called the unscaleable class, the variables are assumed to be independent, so that the probability of observing any given response vector is simply the product of the probabilities of each of the responses to the three variables. That is, \(P(a, b, c) = P(a)P(b)P(c)\) for any unscaleable subject. Therefore, the joint probability of a randomly selected subject being in the unscaleable class and having response vector \((a, b, c)\) is \(P_0 P(a)P(b)P(c)\), regardless of whether cell \((a, b, c)\) is admissible or inadmissible. The probability of a subject from the scaleable class being in an inadmissible cell is zero. If \((a, b, c)\) is an admissible response vector, the probability of a scaleable subject having that response vector is \(P_k\), if \((a, b, c)\) is the \(k\)th admissible cell. Of course, the probability of a subject from the \(k\)th class having a response vector represented by the \(j\)th cell, where \(j \neq k\), is zero. This leads to the following equation, which is fundamental to the model:

\[
P(a, b, c) = \begin{cases} 
0 & \text{if } (a, b, c) \text{ is inadmissible} \\
P_0 P(a)P(b)P(c) & \text{if } (a, b, c) \text{ is admissible}
\end{cases}
\]

This model is fitted to the data by one of several proposed algorithms (Goodman, 1975; Davison, 1980; Bishop, Feinberg and Holland, 1975; Fienberg 1977), and Pearson chi-square and likelihood ratio chi-square fit statistics can be obtained.
Summary of Model-Based Procedures  A number of mathematical models for use in validating learning hierarchies have been presented. These procedures are summarized in Table 3. Although there are four procedures listed in Table 3, those four procedures actually represent only two distinct types. One type is represented by the Dayton and Macready procedure, which subsumes the Proctor model and the White and Clark model. The other type of procedure is represented by the Goodman and Davison model.

Table 3
A Summary of the Procedures for the Validation of Learning Hierarchies Based on Mathematical Models

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Proponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guttman Model</td>
<td>Proctor (1970)</td>
</tr>
<tr>
<td>C Statistic Model</td>
<td>White and Clark (1971)</td>
</tr>
<tr>
<td>General Probabilistic Model</td>
<td>Dayton and Macready (1976)</td>
</tr>
<tr>
<td>Loglinear Model</td>
<td>Goodman (1975), Davison (1981)</td>
</tr>
</tbody>
</table>

Evaluation of Procedures for Validating Hierarchies

The process of establishing a valid hierarchy involves much more than the application of models and statistics to the response patterns of individuals or the mere application of Gagné's task analysis procedure. Before evaluating the procedures that have been proposed for the validation of learning hierarchies, a program in which the procedures should be applied will be discussed. Afterward, the criteria to be used to evaluate the procedures within this context will be presented, followed by an evaluation of the procedures.

A Program for Constructing and Validating Hierarchies

White (1974a) has proposed a nine stage program for the construction and validation of learning hierarchies that is comprehensive and directed at solving some of the problems encountered in the construction of learning hierarchies. The nine stages in the program are:

(a) Define in behavioral terms the element that is to be the highest stage of the hierarchy;
(b) Derive the hierarchy by applying Gagné's task analysis procedure;
(c) Check the reasonableness of the postulated hierarchy with experienced teachers and subject matter experts;
(d) Invent possible divisions of the elements of the hierarchy, so that very precise definitions are obtained;
(e) Carry out an investigation of whether the invented divisions do in fact represent different skills;
Write a learning program for each element, and embed in it mastery
tests for each element;

Have at least 150 suitably chosen subjects work through the
program, taking the tests as they come to them;

Analyze the results to see whether any of the postulated connections
between elements should be rejected; and

Remove all rejected connections from the hierarchy.

Cotton, Gallagher, and Marshall (1977) have proposed a tenth stage
which is to be used only if 15% or more of the connections were rejected.
This stage would involve repeating stages (f) through (i), using the revised
hierarchy with a different group of subjects.

The procedures that have been proposed for validating hierarchies are
not employed until stage (h) of the White model. At this stage of the
process a hierarchy has been postulated, and its reasonableness checked by
experts and experienced teachers. What is needed at stage (h) is a procedure
for validating the postulated hierarchical relationships between the elements
of the hierarchy. The criteria for selecting such a procedure will now be
discussed.

Selection Criteria. The first criterion for selecting a procedure for
validating the hierarchical relationships between elements in a postulated
hierarchy is that the procedure must provide information as to the direction
of the relationship. Simply providing an indication of the strength of an
association between elements is not sufficient. There could be a strong
association between a subordinate and a superordinate element in a proposed
hierarchy simply because the two elements measured the same skill. The
strong association does not imply that the superordinate element could not
have been attained without prior success on the subordinate element. It
must be demonstrated that success on the first element is not only sufficient,
but a necessary condition for success on the second element.

Another criterion is whether the procedure provides information useful
for correcting the structure of the hierarchy if any connections are rejected
as invalid. This includes information that would indicate whether the
elements are out of order, whether any unnecessary elements are included,
and whether any necessary elements were omitted from the hierarchy.

A third criterion for selecting a procedure for validating hierarchies
is whether the procedure indicates how successful a subject must be on an
element before success is likely on the subsequent element. When a single
dichotomously scored item is used for each element, this is not an issue.
But when each element involves a multi-item test, it is possible to reject
a hierarchical relationship simply because success was poorly or incorrectly
defined for that element.

There are undoubtedly other criteria that could be included in this
list. However, these should be adequate for evaluating the overall approaches
to validating learning hierarchies. A detailed comparison of individual
models with an eye toward selecting one or the other might require a more
complete list of selection criteria. That is not the purpose here. The
purpose here is to evaluate the approaches to validating hierarchies that
have been followed by researchers in order to make recommendations on the
direction future research should follow.
Evaluation of the Coefficient Approach

On the basis of the criteria for evaluation set out above, it must be concluded that the coefficient approach to the validation of learning hierarchies is inadequate. There are numerous problems with each of the coefficients, but a more serious problem is the basic inadequacy of the approach itself. The first criterion for evaluation was that a validation procedure must indicate direction as well as strength of a relationship. Coefficients such as those that have been proposed for validating learning hierarchies do not indicate direction. For instance, consider the case where \( N_{00} = 20, N_{01} = 0, N_{10} = 20, \) and \( N_{11} = 60 \). For this case, the PPT statistic has a value of .80, CR = .75, \( r_{OR} = 1.0, \) COR = .75, and phi = .61. If the two tasks are reversed, \( N_{01} = 20, N_{10} = 0, N_{00} = 20, \) and \( N_{11} = 60 \). These values yield a PPT statistic equal to 1.0, CR = 1.0, \( r_{OR} = .75, \) COR = .75, and phi = .61. Reversing the two tables did not change the values of phi and COR, and though it did change the values of the PPT and CR statistics, those statistics had high values regardless of which task was labeled Task 1. Clearly these statistics are not indicants of the direction of the relationship between the two tasks.

Coefficients such as these also do not reliably indicate when unnecessary elements are included in the hierarchy or necessary elements are omitted. If an unnecessary element is redundant, it will probably be correlated with another element or perhaps several elements in the hierarchy. Such a correlation might result in high coefficient values and acceptance of the unnecessary element. When a connection is rejected, there is no indication as to whether the connection was invalid or whether a necessary element was omitted. Task 1 might be necessary but not sufficient for attainment on Task 3. Without Task 2, even subjects successful on Task 1 might be unable to attain success on Task 3. In this case the hierarchy might be rejected without there being any indication that the hierarchy might have been acceptable had Task 2 been included. The same situation might arise if Task 2 were included, but Tasks 2 and 3 were reversed in order.

White (1974b) has listed some other inadequacies of these coefficients. These inadequacies include the following: because the sampling distributions for many of the coefficients are unknown there is no way to determine the error in estimation; and, these coefficients do not provide information useful for deciding success on multi-item tests. Based on White's criticisms of the coefficient approach and those listed above, it must be concluded that the use of coefficients for validating learning hierarchies does not appear to be a productive direction for future research.

Evaluation of the Model Approach

Applying the criteria for evaluation to the model-based procedures for validating hierarchies that have been proposed leads one to conclude that this is a very promising approach, but that the procedures so far developed fall short of fully exploiting their potential. The two types of models that have been proposed are a considerable improvement over the coefficient approach. When the fit of a particular model to the observed data is satisfactory, then a reasonable amount of confidence can be placed in the validity of the hierarchy. Because the mathematical model specified in the procedure describes the way in which performance on one element in a hierarchy is related to performance on another element, the direction of the relationship is indicated. Moreover, because the mathematical model is generally in the form of a probability statement conditioned on a latent variable, the results are generalizable beyond the sample used.
However, there are several problems with the models that have been proposed. One problem is that very little information is obtained for use in defining the criterion for success on multi-item tests. Another serious problem is encountered when the fit of the model to the observed data is inadequate. When the model is rejected, it is not easy to determine why it was rejected. A model could be rejected because the hierarchy is incorrectly specified, because the mathematical model is of an inappropriate form, or both. No information is provided as to whether unnecessary elements were included or necessary elements were omitted. Because of these inadequacies, the model-based procedures that have been proposed need improvement. More research on this approach is needed.

Summary and Recommendations

Summary The area of individualized instruction was identified as one of the fastest growing areas in the field of education. One of the most crucial components of individualized instruction was found to be the sequencing of instructional units. A review of the literature was undertaken to identify the major alternatives available for sequencing units of instruction in such a way as to facilitate education.

Procedures for validating sequences of instructional units, or learning hierarchies, were found to fall into two general categories. One category included procedures based on coefficients of dependence, while the other category contained procedures based on more complete descriptions of the relationships between units of instruction, usually a mathematical model. An evaluation of these two approaches to validating learning hierarchies was undertaken in order to facilitate the formulation of recommendations for the direction of future research in this area.

Procedures based on coefficients were found to provide insufficient information for the validation of hierarchies or for correcting deficiencies in the structure of a proposed hierarchy. Procedures based on mathematical models were also found to be inadequate, but the model-based approach was found to have potential for dealing with the problem of learning hierarchy validation. Recommendations for future research follow.

Recommendations In future research on the validation of learning hierarchies, emphasis should be placed on improving the currently available model-based procedures in several ways. A validation procedure should provide information useful for defining success on the elements of a learning hierarchy, as well as information as to whether important elements have been omitted, inappropriate elements have been included, or some elements have just been placed in the wrong order. Developing a procedure that will provide all this information will make it possible not just to accept or reject a hierarchy, but to correct deficiencies in the structure of rejected hierarchies. It appears that the development of such a procedure would greatly facilitate future development of the field of individualized instruction.
REFERENCES


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

Mr. James Baker
Systems Manning Technical Area
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Beatrice J. Farr
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Myron Fischl
U. S. Army Research Institute for the Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Michael Kaplan
U. S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

Dr. Milton S. Katz
Training Technical Area
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. James L. Raney
U. S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

Mr. Robert Ross
U. S. Army Research Institute for the Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Army

Dr. Robert Sasmor
U. S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Commandant
US Army Institute of Administration
Attn: Dr. Sherrill
FT Benjamin Harrison, TN 46256

Dr. Joseph Ward
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

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

U. S. Air Force Office of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

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

Klievin Hadad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

David R. Hunter
AFHRL/IOAM
Brooks AFB, TX 78235
Civil Govt

1 Dr. Paul G. Chapin
Linguistic's Program
National Science Foundation
Washington, DC 20550

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. John Mays
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Arthur Melmed
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

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

1 Dr. Frank Withrow
U. S. Office of Education
400 Maryland Ave. SW
Washington, DC 20202

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

Non Govt

1 Dr. James Algina
University of Florida
Gainesville, FL 32611

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

1 Anderson, Thomas H., Ph.D.
Center for the Study of Reading
174 Children's Research Center
51 Gerty Drive
Champaign, IL 61820

1 Dr. John Annett
Department of Psychology
University of Warwick
Coventry CV4 7AL
ENGLAND

1 Psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

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

1 Ms. Carole A. Bagley
Minnesota Educational Computing
Consortium
2354 Hidden Valley Lane
Stillwater, MN 55082

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Capt. J. Jean Belanger
Training Development Division
Canadian Forces Training System
CFTSHQ, CFB Trenton
Astra, Ontario K0K 180
Non Govt

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

2. Dr. Werner Pirke
DezWPs im Streitkraefteamt
Postfach 20 50 03
D-5300 Bonn 2
WEST GERMANY

3. Dr. R. Darrel Boeck
Department of Education
University of Chicago
Chicago, IL 60637

1. Liaison Scientists
Office of Naval Research,
Branch Office, London
Box 39 FPO New York 09510

1. Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

1. DR. JOHN F. BROCK
Honeywell Systems & Research Center
(MN 17-2318)
2600 Ridgeway Parkway
Minneapolis, MN 55413

1. DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 S0. STATE ST.
OREM, UT 84057

1. Dr. John B. Carroll
Psychometric Lab
Univ. of N.C. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1. Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90027

1. Dr. William E. Coffman
Director, Iowa Testing Programs
334 Lindquist Center
University of Iowa
Iowa City, IA 52242

1. Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, D.C. 20026

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

1. Dr. Fritz Drasgow
Department of Psychology
University of Illinois
603 East Daniel
Champaign, IL 61820

1. Mike Durmeyer
Instructional Program Development
Building 90
NET-PDCD
Great Lakes NTC, IL 60088

1. ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1. Dr. Benjamin A. Fairbank, Jr.
McFann-Gray & Associates, Inc.
5925 Callaghan
Suite 225
San Antonio, Texas 78228
1 Dr. Leonard Feldt
Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

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

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

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

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Daniel Gopher
Industrial & Management Engineering
Technion-Israel Institute of Technology
Haifa
ISRAEL

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

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

1 Dr. Delwyn Harnischf
University of Illinois
242b Education
Urbana, IL 61801

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Dustin H. Heuston
Wicat, Inc.
Box 986
Orem, UT 84057

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Prof. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Jack Hunter
2122 Coolidge St.
Lansing, MI 48906

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

1 Professor John A. Keats
University of Newcastle
AUSTRALIA 2308

1 Mr. Jeff Kelety
Department of Instructional Technology
University of Southern California
Los Angeles, CA 92007
Dr. Stephen Kosslyn  
Harvard University  
Department of Psychology  
33 Kirkland Street  
Cambridge, MA 02138

Dr. Marcy Lansman  
Department of Psychology, NI 25  
University of Washington  
Seattle, WA 98195

Dr. Alan Lesgold  
Learning R&D Center  
University of Pittsburgh  
Pittsburgh, PA 15260

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

Dr. Charles Lewis  
Faculteit Sociale Wetenschappen  
Rijksuniversiteit Groningen  
Oude Boteringestraat 23  
9712GC Groningen  
Netherlands

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

Dr. Frederick M. Lord  
Educational Testing Service  
Princeton, NJ 08540

Dr. James Lumsden  
Department of Psychology  
University of Western Australia  
Nedlands W.A. 6009  
AUSTRALIA

Dr. Gary Marco  
Educational Testing Service  
Princeton, NJ 08440

Dr. Scott Maxwell  
Department of Psychology  
University of Houston  
Houston, TX 77204

Dr. Samuel T. Mayo  
Loyola University of Chicago  
820 North Michigan Avenue  
Chicago, IL 60611

Professor Jason Millman  
Department of Education  
Stone Hall  
Cornell University  
Ithaca, NY 14853

Dr. Allen Munro  
Behavioral Technology Laboratories  
1845 Elena Ave., Fourth Floor  
Redondo Beach, CA 90277

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

Dr. Jesse Orlansky  
Institute for Defense Analyses  
400 Army Navy Drive  
Arlington, VA 22202

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

Dr. James A. Paulson  
Portland State University  
P.O. Box 751  
Portland, OR 97207

MR. LUIGI PETRULLO  
2431 N. EDGEWOOD STREET  
ARLINGTON, VA 22207
Dr. Brad Symson  
Psychometric Research Group  
Educational Testing Service  
Princeton, NJ 08541

Dr. Kikumi Tatsuoka  
Computer Based Education Research Laboratory  
252 Engineering Research Laboratory  
University of Illinois  
Urbana, IL 61801

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

Dr. Douglas Towne  
Univ. of So. California  
Behavioral Technology Labs  
1845 S. Elena Ave.  
Redondo Beach, CA 90277

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

Dr. David Vale  
Assessment Systems Corporation  
2395 University Avenue  
Suite 306  
St. Paul, MN 55114

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

Dr. Susan E. Whitely  
PSYCHOLOGY DEPARTMENT  
UNIVERSITY OF KANSAS  
LAWRENCE, KANSAS 66044

Wolfgang Wildgrube  
Streitkraefteamt  
Box 20 50 03  
D-5300 Bonn 2