DOCUMENT RESUME

ED 111 845 TM 004 812

AUTHOR	Besel, Ronald
TITLE	Mastery-Learning Decision Variables.
INSTITUTION	Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
REPORT NO	SWRL-TM-5-71-08
PUB DATE	18 Aug 71
NOTE	24p.; For related documents, see TM 004 811 and 814
EDRS PRICE	MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS	*Criterion Referenced Tests; Decision Making; Groups; Individual Tests; Instructional Design; *Mathematical Models; *Matrices; Probability; Psychometrics; Statistical Analysis; Test Interpretation; Test Reliability; *Test Results
IDENTIFIERS	Learning Hierarchies; *Mastery Learning Test Model (Emrick and Adams); Mastery Tests

ABSTRACT

The Mastery-Learning test model is extended. Methods for estimating prior probabilities are described. The use of an adjustment matrix to transform a probability of mastery measure and empirical methods for estimating adjustment matrix parameters are derived. Adjustment matrices are interpreted as indicators of instructional effectiveness and as evidence of the existence of learning hierarchies. Two decision variables are considered: probability of mastery for an individual and proportion in mastery for an instructional group. Discussion of the reliability, complexity, and interpretability of these decision variables and comparison with decision variables for other test models is also included. (Author/DEP)

SOUTHWEST REGIONAL LABORATORY
TECHNICAL MEMORANDUM

DATE: August 18, 1971

NO: TM 5-71-08

TITLE: MASTERY-LEARNING DECISION VARIABLES

AUTHOR: Ronald Besel

ABSTRACT

The Mastery-Learning test model described in TM 5-71-04 is extended. Methods for estimating prior probabilities are described. The use of an adjustment matrix to transform a probability of mastery measure and empirical methods for estimating adjustment matrix parameters are derived. Adjustment matrices are interpreted as indicators of instructional effectiveness and as evidence of the existence of learning hierarchies. Two decision variables are considered: probability of mastery for an individual and proportion in mastery for an instructional group. Discussion of the reliability, complexity, and interpretability of these decision variables and comparison with decision variables for other test models is also included.

MASTERY-LEARNING DECISION VARIABLES

Introduction

This paper focuses on the analysis of test data by a mastery-learning test model. The inputs for the test model are the responses of individuals to test items; these responses are classified as either correct or incorrect. The outputs of the test model are called decision variables. The test model described here is algorithmic, i.e., a mathematical model is used to compute numerical values for the decision variables. It is an extension of the Mastery Learning Model (Emrick and Adams, 1970) as described in TM 5-71-04 (Besel, 1971). Two decision variables are considered: probability of mastery for an individual and proportion in mastery for an instructional group.

Symbols and Notation

Variables will be represented by capital letters. If a variable represents a vector, the elements of the vector will be represented by the equivalent lower case symbol.

Lower case letters and numerals will be used as subscripts. The subscript (i) refers to a test item, (j) to an individual, (k) to a performance measure for an objective. If a relationship among variables does not refer to a particular test item, individual or objective, the corresponding subscripts will be deleted from the symbolic representation of the relationship. The subscript (j) will be used to refer to a second test item in an expression relating two composite tests.

Capital letter P will refer to a probability. Functional relationships will be denoted by enclosing the independent variable in parentheses (e.g., P(X/M) will represent the conditional probability of a response X given that the individual is in the mastery (M) state). The set theoretic notation $\bar{M}$ will be used to denote not being in the mastery (M) state.

The Mastery-Learning Model

The mastery-learning model (TM 5-71-04) assumes that a test measures proficiency with respect to a single skill and that there are only two states of proficiency for that skill. Each individual tested is in either the mastery (M) or non-mastery ($\bar{M}$) state at the time of testing. The only true scores are assumed to be 0 and K (for a K-item test); all intermediate scores are due to measurement error.

There are two classes of measurement errors: wrong responses by individuals in the mastery state (β errors) and correct responses by individuals in the non-mastery state (α errors).

α_i = the probability that an individual in the $\bar{M}$ state will give a correct response to the i^{th} item.

β_i = the probability that an individual in the M state will give an incorrect response to the i^{th} item.

The α_i and β_i parameters are assumed to have true values which are characteristic of the test. Emrick and Adams' model has been modified to permit item parameters rather than single α and β test parameters. The assumption is still made that α_i and β_i have the same value for every individual belonging to a common instructional group.

Let,

x_{ij} represent the response of individual j to item i,

$$x_{ij} = \begin{cases} 1 \text{ if a correct response is given} \\ 0 \text{ if an incorrect response is given} \end{cases} \quad (1)$$

$$X_j = \left[x_{1j}, x_{2j}, x_{3j}, \ldots, x_{Kj} \right] \quad (2)$$

represent the response vector for individual j

$$S_j = x_{1j} + x_{2j} + x_{3j} + \ldots + x_{Kj} \quad (3)$$

represent the test score for individual j

Sequential analysis of the item responses will be assumed in the derivation of the mastery-learning model. The (j) subscript will be deleted to simplify the notation.

For any individual tested the following conditional probabilities are associated with his response to the first item:

$$P(x_1 = 1 / M) = 1-\beta_1 \quad (4)$$

$$P(x_1 = 0 / M) = \beta_1 \quad (5)$$

$$P(x_1 = 1 / \bar{M}) = \alpha_1 \quad (6)$$

$$P(x_1 = 0 / \bar{M}) = 1-\alpha_1 \quad (7)$$

The probability that a response indicates a particular state can be computed using Bayes formula:

$$P(M / x_1) = \frac{PRM \cdot P(x_1 / M)}{P(x_1)} \quad (8)$$

where,

PRM represents the prior probability of the mastery state,

$P(x_1)$ represents the prior or expected distribution of x_1 given by,

$$P(x_1) = PRM \cdot P(x_1 / M) + [1-PRM] \cdot P(x_1 / \bar{M}) \quad (9)$$

For a correct response,

$$P(M / x_1 = 1) = \frac{PRM\ (1-\beta_1)}{PRM \cdot (1-\beta_1) + [1-PRM] \cdot \alpha_1} \quad (10)$$

For an incorrect response,

$$P(M / x_1 = 0) = \frac{PRM \cdot (\beta_1)}{PRM \cdot \beta_1 + [1-PRM] \cdot [1-\alpha_1]} \quad (11)$$

Methods for estimating prior probabilities will be discussed later in this paper.

The conditional probability of mastery based on the first item response (equation 8) is used as the prior probability for the second item response.

$$P(M / x_1, x_2) = \frac{P(M/x_1) \cdot P(x_2/M)}{P(M/x_1) \cdot P(x_2/M) + [1-P(M/x_1)] \cdot P(x_2/\bar{M})} \quad (12)$$

Substituting equations (8) and (9) for $P(M/x_1)$ yields;

$$P(M/x_1, x_2) = \frac{PRM\ [P(x_1/M) \cdot P(x_2/M)]}{PRM \cdot [P(x_1/M) \cdot P(x_2/M)] + [1-PRM]\ [P(x_1/\bar{M}) \cdot P(x_2/\bar{M})]} \quad (13)$$

This procedure can be repeated, sequentially computing the conditional probability of mastery given the i^{th} item response, with the prior probability based on the previous responses. For a K-item test:

$$P(M/X) = \frac{PRM \prod_{i=1}^{K} (P(x_i/M)}{PRM \prod_{i=1}^{K} P(x_i/M) + [1-PRM] \prod_{i=1}^{K} P(x_i/\bar{M})} \quad (14)$$

For any length test,

$$P(\bar{M}/X) = 1-P(M/X) \quad (15)$$

If average values of α_i and β_i are estimated rather than item parameters, equation (14) becomes:

$$P(M/S) = \frac{PRM \cdot (1-\beta)^S (\beta)^{K-S}}{PRM \cdot (1-\beta)^S (\beta)^{K-S} + [1-PRM] \cdot (\alpha)^S (1-\alpha)^{K-S}} \quad (16)$$

Estimating the Proportion of Students in the Mastery State

The proportion-in-mastery for a group of students can be estimated from the observed mean score for the group. The following derivation of the relationship between mean score and proportion-in-mastery makes two independence assumptions:

(1) The responses of a student to each test item are independent of the responses of all other students in the group.

(2) An individual's responses to the separate items on a test can be treated as a sequence of independent trials.

Let, $E(S_j)$ represent the expected score for the j^{th} individual and,

$$U = \frac{1}{N}\left[\sum_{j=1}^{N} S_j\right] \qquad (17)$$

represent the observed sample mean for a group of N students. From the first assumption, the expected value of the observed sample mean is:

$$E(\bar{U}) = \frac{1}{N} \cdot \left[\sum_{j=1}^{N} E(S_j)\right] \qquad (18)$$

For an item with parameters (α_1, β_1),

$$E(x_i)\Big|_{M} = 1-\beta_1 \quad \text{for the } N_m \text{ individuals in the mastery state.} \qquad (19)$$

$$E(x_i)\Big|_{\bar{M}} = \alpha_1 \quad \text{for the } (N-N_m) \text{ individuals in the non-mastery state.} \qquad (20)$$

From the second assumption the expected scores for individuals in the M and $\bar{M}$ states are:

$$E(S)\Big|_{M} = \sum_{i=1}^{K} (1-\beta_1) \quad \text{for the } N_m \text{ individuals in M state.} \qquad (21)$$

$$E(S)\Big|_{\bar{M}} = \sum_{i=1}^{K} \alpha_1 \quad \text{for the } N-N_m \text{ individuals in } \bar{M} \text{ state.} \qquad (22)$$

The expected value of the observed sample mean is then,

$$E(U) = \frac{1}{N}\left[N_m \cdot E(S)\Big|_{M} + (N-N_m) \cdot E(S)\Big|_{\bar{M}}\right] \qquad (23)$$

$$= \frac{1}{N}\left[N_m \cdot \sum_{i=1}^{K} (1-\beta_i) + (N-N_m) \cdot \sum_{i=1}^{K} \alpha_i\right] \qquad (24)$$

Define the proportion-in-mastery to be,

$$MP = \frac{N_m}{N} \tag{25}$$

and the estimated value of proportion in mastery for a particular group to be GMP.

Then,

$$E(U) = MP \cdot \sum_{i=1}^{K} (1-\beta_i-\alpha_i) + \sum_{i=1}^{K} \alpha_i \tag{26}$$

Using the observed sample mean, U, as an estimate of E(U) and solving for GMP yields:

$$GMP = \frac{U - \sum_{i=1}^{K} \alpha_i}{\sum_{i=1}^{K} (1-\beta_i-\alpha_i)} \tag{27}$$

If average values of α and β are employed rather than item values, this relationship simplifies to,

$$GMP \simeq \frac{U/K - \alpha}{1-\beta-\alpha} \tag{28}$$

An estimate of the proportion of students in the mastery state can be used for two distinct purposes: (1) it can be used directly as a decision variable by the evaluator to judge the effectiveness of an instructional unit or by the teacher in selecting an appropriate strategy for remedial instruction. If only a small percentage of an instructional group has achieved mastery, review or second instruction for the entire group may be warranted while tutorial assistance for non-masters may be preferred if a large percentage has achieved mastery; (2) it can also be

used to estimate prior probabilities needed in the computation of probability of mastery.

Prior Probability Estimates

Two general classes of prior probability estimates can be used in the computation of probability of mastery. The first includes all methods which assign the same prior probability to each individual in a group. The second class includes all methods which use other test data obtained from an individual to estimate a "personalized" prior probability.

The proportion of students in mastery estimate can be used directly as a Class 1 prior probability. Each student is assigned the same prior probability of mastery, GMR; as the group mean score increases, the prior probability estimate increases for each member of the group.

Emrick and Adams (1970) suggested a different type of Class 1 estimate. The anticipated instructional effects, i.e., the anticipated proportion of students in mastery after instruction, can be estimated from relevant past experiences. The performance of a similar group of students during the previous year may be relevent if the instructional activities are comparable. This approach may be best suited for "small" instructional groups and replicated instructional treatments while the GMP estimate can be used for "first time" instructional treatments applied to sufficiently large instructional groups.

The only type of Class 2 estimate which will be considered here is based on test data which have been analyzed in terms of probability of mastery of some performance objective. The only requirement is that

the responses of an individual to the test items currently being analyzed be independent from the responses to test items used for the prior probability estimate. Thus, the test used may be a parallel test for the same performance objective, it may be a concurrently administered test for a different performance objective or a previous measure of a performance objective.

Adjustment Matrices

It will generally be desirable to transform a probability of mastery measure in order to use it as a personalized prior probability estimate. A linear transformation of a prior performance measure to estimate a personalized prior probability will be represented by an adjustment matrix.

$$A_{c(k)} = \begin{bmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{bmatrix} \qquad (29)$$

$A_{c(k)}$ is the adjustment matrix used to estimate a personalized prior probability for a current performance measure (c) from performance measure (k). The personalized prior probability for measure (c) is then,

$$\begin{bmatrix} 1\text{-}PRM_{j,c(k)} \\ PRM_{j,c(k)} \end{bmatrix} = \begin{bmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{bmatrix} \begin{bmatrix} 1\text{-}PRM_{j,k} \\ PM_{j,k} \end{bmatrix} \qquad (30)$$

where $PM_{j,k}$ is the probability of mastery for the j^{th} individual for performance measure (k).

The $a_{c,k}$ parameters in an adjustment matrix are conditional probabilities. For example,

$a_{1,0}$ is the expected probability that the individual will be in the mastery state for the current objective on the condition that he was in the non-mastery state for the prior objective.

Since the $a_{c,k}$ parameters are probabilities,

$$a_{0,0} + a_{1,0} = 1 \tag{31}$$

and

$$a_{0,1} + a_{1,1} = 1 \tag{32}$$

Adjustment matrices have educational interpretations for particular choices of prior performance measures. If the prior performance measure is a pretest in the form of a parallel measure of the current performance measure, the adjustment matrix can be interpreted as a representation of instructional effectiveness. The $a_{1,1}$ value for an instructional effectiveness matrix should be very close to 1.0 since it can usually be assumed that instruction will not lead to a transition from mastery to a non-mastery state. The $a_{1,0}$ value is, for the case of $a_{1,1}$ equal to 1.0, a useful index of instructional effectiveness.

If a hierarchial relationship between two performance measures exists, a second type of adjustment matrix with a meaningful interpretation can be constructed. Suppose objective (k) must be mastered before the current objective, (c), can be mastered. The adjustment matrix, $A_{c(k)}$, should have an $a_{1,0}$ value equal to zero.

Unbiased Prior Probabilities

A prior probability estimate for current performance measure (c) based on performance measure (k) is defined to be unbiased if:

$$GMP_c = a_{1,0} + [a_{1,1} - a_{1,0}] GMP_k \quad (33)$$

where GMP_c is the proportion-in-mastery for the current performance measure.

A graphical representation of an adjustment matrix facilitates the explanation of this definition (see Figure 1).

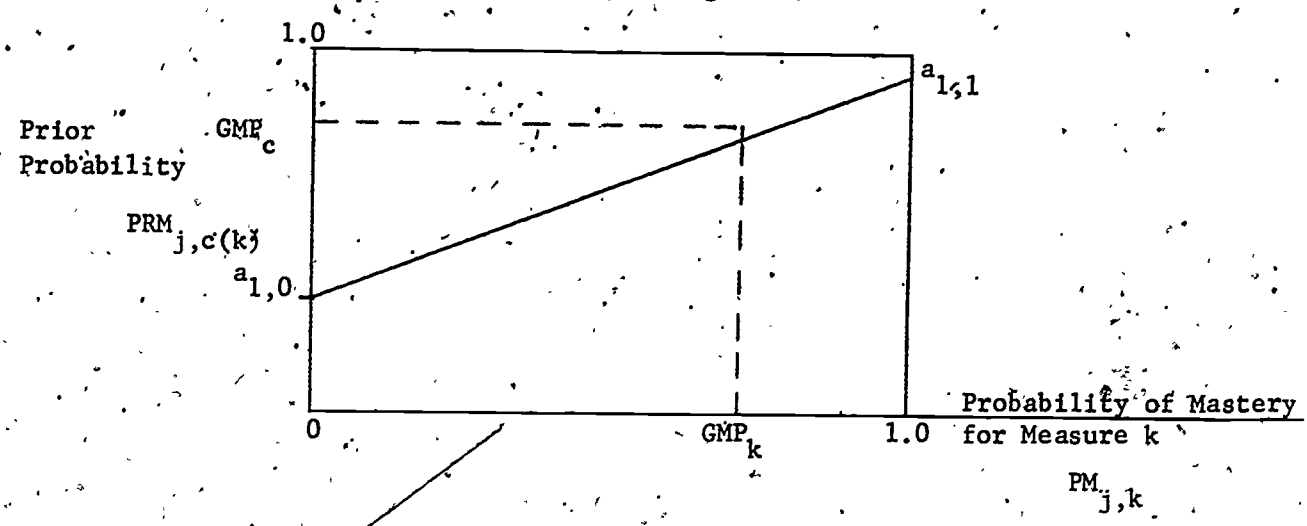

Figure 1. Adjustment matrix for estimating a prior probability for current performance measure (c) from a probability of mastery value computed for measure (k).

An adjustment matrix is represented graphically by a straight line which has an intercept equal to $a_{1,0}$ and a slope equal to $(a_{1,1} - a_{1,0})$. Equation (33) results from substituting GMP_k for $PM_{j,k}$ and GMP_c for $PRM_{j,c(k)}$.

A Correlational Method for Estimating Adjustment Matrix Parameters

The correlation between two dichotomous measures can be computed using the formula for a Phi coefficient:

$$\text{Phi} = \frac{P_{1,1} P_1 - P_1 P_2}{\sqrt{P_1 (1-P_1)} \sqrt{P_2 (1-P_2)}} \qquad (34)$$

where P_1 and P_2 represent the marginal probabilities of being in the (1) state for the first and second measure respectively and $P_{1,1}$ represents the conditional probability $P(x_2 = 1/x_1 = 1)$.

A contingency table can be used to display the correlation between two dichotomous measures.

Measure 2 \ Measure 1	$1-P_1$	P_1
$1-P_2$	$P_{0,0}$	$P_{0,1}$
P_2	$P_{1,0}$	$P_{1,1}$

Figure 2. Contingency table relating Variable 1 to Variable 2.

The following relations among conditional and marginal probabilities will be used in later derivations:

$$P_{0,0} + P_{1,0} = 1 \qquad (35)$$

$$P_{0,1} + P_{1,1} = 1 \qquad (36)$$

$$P_2 = P_1 \cdot P_{1,1} + [1-P_1] \cdot P_{1,0} \qquad (37)$$

$$P_1 = P_2 \cdot P_{1,1} + [1-P_2] \cdot P_{0,1} \qquad (38)$$

Two contingency tables are of particular interest: the contingency tables for an adjustment matrix and for relating observed response to true state.

Current Measure \ Prior Measure	$1-MP_k$	MP_k
$1-MP_c$	$a_{0,0}$	$a_{0,1}$
MP_c	$a_{1,0}$	$a_{1,1}$

Figure 3. Prior Measure - Current Measure Contingency Table

Observed Response \ True State	$\bar{M}$	M
	$1-MP$	MP
Incorrect (0) $1-PC_i$	$1-\alpha_i$	β_i
Correct (1) PC_i	α_i	$1-\beta_i$

Figure 4. True State - Observed Response Contingency Table for a single test item.

The symbol PC_i will be used to represent the probability of a correct response to item (i) if a respondent is randomly selected from the population.

Observed score will be defined to be related to true score for measure (k) by the following equation:

$$x_{i,k} = t_k + e_{i,k} \qquad (39)$$

where t_k is the true score;

$$t_k = \begin{cases} 0 & \text{for an individual in the non-mastery state,} \\ 1 & \text{for an individual in the mastery state,} \end{cases}$$

and $e_{i,k}$ is the measurement error;

$$e_{i,k} = \begin{cases} 1 & \text{if a non-master gives a correct response,} \\ -1 & \text{if a master gives an incorrect response,} \\ 0 & \text{otherwise.} \end{cases}$$

Since $x_{i,k}$ and t_k are dichotomous variables assuming only the values zero and one,

$$\sigma^2(x_{i,k}) = PC_{i,k}\,(1-PC_{i,k}) \tag{40}$$

and,

$$\sigma^2(t_k) = MP_k\,(1-MP_k) \tag{41}$$

The correlation between true score and observed score is computed using equation (34).

$$\rho(x_{i,k}\,,\,t_k) = \frac{MP_k\,(1-\beta_i) - MP_k \cdot PC_{i,k}}{\sqrt{MP_k\,(1-MP_k)}\;\sqrt{PC_{i,k}\,(1-PC_{i,k})}} \tag{42}$$

Using relation (37) to express $PC_{i,k}$ in terms of MP_k,

$$\rho(x_{i,k}\,,\,t_k) = \frac{MP_k\,(1-MP_k)\,(1-\beta_i-\alpha_i)}{\sqrt{MP_k\,(1-MP_k)}\;\sqrt{PC_{i,k}\,(1-PC_{i,k})}} \tag{43}$$

which simplifies to,

$$\rho(x_{i,k}\,,\,t_k) = \frac{\sigma^2(t_k)\,(1-\beta_i-\alpha_i)}{\sigma(t_k)\;\sigma(x_{i,k})} \tag{44}$$

from the general relationship for a correlation,

$$\rho(x,t) = \frac{\sigma(x,t)}{\sigma(x)\cdot\sigma(t)} \tag{45}$$

it can readily be seen that the numerator of (44) is the covariance of $x_{i,k}$ and t_k;

$$\sigma(x_{i,k}, t_k) = \sigma^2(t_k)\cdot\gamma_{i,k} \tag{46}$$

where $\gamma_{i,k}$ is defined to be,

$$\gamma_{i,k} = 1-\alpha_i-\beta_i \tag{47}$$

it can be shown that the covariance of any two dichotomous items is,

$$\sigma(x_i, x_j) = \sigma(t_1, t_2)\,\gamma_{i,1}\cdot\gamma_{j,2}$$

where t_1 is the true score for measure 1, t_2 is the true score for measure 2 and x_i, x_j are observed responses to items selected from measures 1 and 2 respectively.

Relations for the covariance and correlation of composite variables S_K (see equation 3) are derived directly from item relations.

Define γ_k to be the average value of $\gamma_{i,k}$

$$\gamma_k = \frac{1}{K}\sum_{i=1}^{K}\gamma_{i,k}$$

$$\sigma(S_1, S_2) = \sum_{i=1}^{K1}\sum_{j=1}^{K2}\sigma(x_{i,1}, x_{j,2})$$

$$= \sum_{i=1}^{K1}\sum_{j=1}^{K2}\sigma(t_1, t_2)\,\gamma_{i,1}\cdot\gamma_{j,2}$$

$$= \sigma(t_1 , t_2) \left[\sum_{i=1}^{K1} \gamma_{i,1} \right] \left[\sum_{j=1}^{K2} \gamma_{j,2} \right]$$

$$= \sigma(t_1 , t_2) \, (K1 \cdot \gamma_1) \, (K2 \cdot \gamma_2) \qquad (48)$$

$$\sigma(S_1 , t_1) = \sum_{i=1}^{K1} \sigma(x_{i,1} , t_1)$$

$$= \sigma^2(t_1) \cdot \sum_{i=1}^{K1} \gamma_{i,1}$$

$$= \sigma^2(t_1) \, (K1 \cdot \gamma_1) \qquad (49)$$

For Classical Test Theory, the following correction for attenuation formula is used to estimate the correlation between true scores from the correlation between observed scores (Lord and Novick, 1968).

$$\rho(t_1 , t_2) = \frac{\rho(S_1 , S_2)}{\rho(S_1 , t_1) \, \rho(S_2 , t_2)} \qquad (50)$$

It will be shown that this relation holds for the mastery learning model as well.

Substituting (45) into (50) yields

$$\rho(t_1 , t_2) = \frac{\sigma(S_1 , S_2) \, \sigma(t_1) \, \sigma(t_2)}{\sigma(S_1 , t_1) \, \sigma(S_2 , t_2)} \qquad (51)$$

then using relations (48) and (49)

$$\rho(t_1 , t_2) = \frac{\sigma(t_1 , t_2) \, (K1 \cdot \gamma_1) \, (K2 \cdot \gamma_2) \, \sigma(t_1) \, \sigma(t_2)}{\sigma^2(t_1) \, (K1 \cdot \gamma_1) \, \sigma^2(t_2) \, (K2 \cdot \gamma_2)}$$

$$= \frac{\sigma(t_1 , t_2)}{\sigma(t_1)\ \sigma(t_2)} \qquad (52)$$

which is an identity.

For an adjustment matrix,

$$\rho(t_c , t_k) = \frac{a_{1,1} \cdot MP_k - MP_k \cdot MP_c}{\sigma(t_c) \cdot \sigma(t_k)} \qquad (53)$$

Solving for $a_{1,1}$ yields

$$a_{1,1} = \frac{\sigma(t_c , t_k)}{MP_k} - MP_c \qquad (54)$$

Using GMP_k to estimate MP_k, GMP_c to estimate MP_c and the relation between $\sigma(t_c , t_k)$ and $\sigma(S_c , S_k)$ leads to,

$$a_{1,1} = \frac{\sigma(S_c , S_k)}{GMP_k\ (K1 \cdot \gamma_1)\ (K2 \cdot \gamma_2)} + GMP_c \qquad (55)$$

If the adjustment matrix is restricted to be unbiased, equation (33) can be used to solve for the $a_{1,0}$ parameter.

$$a_{1,0} = \frac{GMP_c - a_{1,1} \cdot GMP_k}{1 - GMP_k}$$

$$= GMP_c - \frac{\sigma(S_c , S_k)}{(1 - GMP_k) \cdot (K1 \cdot \gamma_1)\ (K2 \cdot \gamma_2)} \qquad (56)$$

Formulas (55) and (56) can be used to obtain adjustment matrix parameters empirically. Methods for estimating the α_i, β_i parameters will be discussed in a subsequent paper.

The Adjustment Matrix for Parallel Tests

Two mastery tests are parallel if the correlation between true scores for the tests is equal to 1. The correlation between two dichotomous measures can be unity only if:

$$P_1 = P_2 \tag{57}$$

which means that the proportion-in-mastery values are equal, and,

$$a_{1,1} = 1.0 \tag{58}$$

From equation (37), it can be seen that perfect correlation also implies that:

$$a_{1,0} = 0 \tag{59}$$

In the derivation of the Mastery Learning Model, sequential analysis of item responses was assumed to be proper. This is equivalent to the assumption that one (or several) items can be used as a prior measure--with an adjustment matrix satisfying (58) and (59)--in computing probability of mastery given one additional item response. The assumption reduces to the requirement that the test items are parallel measures.

It should be noted that the α_i, β_i parameters and the item difficulties (PC_i) need not be equal for items to be parallel. The assumption of parallel items is thus less restrictive than is the case for classical test theory (Lord and Novick, 1968).

Reliability, Complexity, and Interpretability

The following properties influence the utility of decisions variables as inputs to a decision process:

(a) Reliability: the measurement property of being repeatable. A second testing results in the same rank-ordering of individuals (or groups) or the same classification into categories.

(b) Complexity: to make effective decisions requires the derivation and application of complex decision rules. A decision procedure which is difficult to explain to individuals responsible for the decisions made will be labeled "complex" as well as procedures requiring lengthy computation.

(c) Interpretability: the decision variable has an intuitive or easily understood meaning for the decision maker.

It is much less complicated to derive an objective basis for comparing the reliability of decision variables than it is for the complexity and interpretability properties. This alone is not a valid reason for selecting decision variables solely on the basis of reliability. The relative importance of each of these properties for effective decision making must be considered; this will depend both upon the type of decision being made and the characteristics of the decision maker.

Comparison with two distinctive tests models will be used in discussing reliability, complexity, and interpretability.

Test Model 1: Standardized-Normal

(a) Individual Student Decision Variable--standard (z) score

(b) Group Decision Variable--group mean score

Test Model 2: Criterion-Referenced

(a) Individual Student Decision Variable--raw or percentage correct score

(b) Group Decision Variable--percentage exceeding criterion score

Reliability indices for the Standardized-Normal Test Model attempt to assess whether the rank-ordering of individuals is preserved with repeated testing. Preservation of rank-order is important for applications requiring correlational analysis and some selection problems. A Criterion-Referenced Test Model is more frequently used for different types of applications (performance assessment and placement). A reliability index which assesses the repeatability of classification into several (usually two) categories may be most appropriate for these applications.

The appropriate reliability index is thus dependent upon the particular application employing the decision variable. Reliability assessment for the Mastery-Learning Model is complicated by the possibility of having a number of measures available which are potential prior measures. A different reliability coefficient can be computed for each choice of a prior performance measure. For those who accept the notion of multiple indices of test validity, this complication should impose computational rather than conceptual difficulties. It should be possible to select prior measures which improve both the reliability and validity of mastery learning decision variables for those applications for which Criterion-Referenced analysis is currently used.

As defined in this paper, assessing the complexity of a decision variable requires the prior existance of decision rules which have been evaluated in terms of effectiveness. A simple decision rule which is adequate in one situation may be ineffective in another. Theoretically derived decision rules using standardized scores tend to be complex

(Cronbach and Gleser, 1965); they would be quite difficult to explain to the non-mathematicians and may require considerable computational power. Decision procedures for the Criterion-Referenced Model are degraded by the absence of an assumed distribution of true scores. This tends to both complicate the decision procedures and reduce its effectiveness (Besel, 1971).

It may be possible to use approximate decision rules which satisfy the complexity criterion and achieve acceptable results but the interpretability of the results may suffer. Arbitrary choice of criterion-scores for Criterion-Referenced Models and cutting scores for Standardized-Normal Model may simplify decision procedures at the expense of understanding the effects of altering the choice.

Theoretically based decision rules can be derived for the Mastery-Learning Model which are relatively simple to explain and calculate. For interpretability, they depend only on the concepts of subjective probability and expected loss. Emperical verification of these claims remains to be attempted.

REFERENCES

Besel, R. R. 1971, "A Comparison of Emrick and Adam's Mastery-Learning Test Model with Kriewall's Criterion-Referenced Test Model" (TM 5-71-04), Southwest Regional Laboratory.

Cronbach, L. J., and Gleser, G. C. Psychological Tests and Personnel Decisions, University of Illinois Press.

Emrick, J. A., and Adams, E. N. "An Evaluation Model for Individualized Instruction," paper presented at the Annual Meeting of the American Educational Research Association, Minneapolis, Minnesota, 1970.

Lord, F. M., and Novick, M. R. Statistical Theories of Mental Test Scores, Addison-Wesley, 1968.