DOCUMENT RESUME

ED 128 422 TM 005 607

AUTHOR Thomas, David B.
TITLE Discriminant Analysis Applied to Sequential Testing.
SPONS AGENCY Iowa Univ., Iowa City.
PUB DATE [Sep 75]
NOTE 21p.; Paper presented at the Annual Conference of the Military Testing Association (17th, Fort Benjamin Harrison, Indiana, September 15-19, 1975); Also included in TM 005 585

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Computer Assisted Instruction; *Computer Oriented Programs; *Discriminant Analysis; Educational Technology; *Individualized Instruction; Predictive Validity; Probability; *Sequential Approach; Sequential Programs; *Student Placement; *Testing; Test Reliability; Tests

ABSTRACT

Individualized instruction programs have imposed an increased reliance on tests as a means of selecting and routing students through sometimes complex programs. Testing which occurs within the training sequence is particularly vulnerable to inefficient use of both trainee and instructor time, but computer-based instruction system can provide a means for monitoring and scheduling the routings. Both branched tests and short linear tests have been used, but both have disadvantages not present with the use of the sequential testing model. The sequential test may be described as one in which an examinee is given a test item, the response to the current item as well as previously administered items is evaluated in a certain way, and the examinee is either assigned to one of a number of classes and testing stopped, or judgment is suspended and testing continues. If single items from a test are employed for the derivation of discriminant functions, then functions could be computed to provide classifications after a small number of test items had been administered. This technique results in probabilities of group membership every time a linear combination is computed, and may be adapted to sequential testing by the addition of a termination rule. (BW)

Discriminant Analysis Applied to Sequential Testing

by

David B. Thomas, Ph.D.
University of Iowa

Support for this research was provided by the Graduate College and the University Computer Center at the University of Iowa. The author would like to express his appreciation to the Iowa Testing Programs for their Cooperation.

Discriminant Analysis Applied to Sequential Testing

David B. Thomas, Ph. D.

Computer Assisted Instruction Laboratory
University of Iowa

Individually prescribed instruction programs have imposed an increased reliance on tests as a means of selecting students, and routing them through the sometimes complex programs. Instructional prescription models are being developed which depend upon the assessment of various learner characteristics including achievement, personality factors, aptitudes, attitude, and personal interests. Assessment tests are scheduled prior to instruction and may be rescheduled a number of times during the instructional process. Test batteries are also administered to select students for a particular instructional program - mechanics or electronics for example.

The potentially large number of tests that could be administered as part of a training program conducted within an individualized framework could easily impose a burden on resources. Costs of administering, scoring, and interpreting tests and of supporting trainees could reach a substantial amount. Reducing the time devoted to testing may therefore be considered a high priority activity. It is important, however, that the amount of information obtained through a testing program not be reduced significantly.

Testing which occurs within the training sequence is particularly vulnerable to inefficient use of both trainee and instructor time. Students complete units and could be tested at different times, perhaps over different content skills. Scheduling and monitoring of imbedded tests could present a major administrative problem.

Computer-based instruction systems provide one means for monitoring and scheduling the complex routings which may take place within an individualized training system. Test administration may also take place under computer control. The turnaround time required for scoring each test could be eliminated by programming the computer to score the test as it is administered.

A computer-administered test is not necessarily a shorter test, and although a computer-based system may reduce administrative costs somewhat, a means to reduce the testing time from the student perspective is desirable. There have been numerous attempts to shorten testing time: short linear tests composed of the most highly discriminating items from a longer test; branched or tailored tests in which items are presented with a difficulty level selected so as to approximate a dynamically estimated examinee ability level; and sequential tests in which items are administered until a criterion is satisfied. Branched tests are described by Wood (1973) and Weiss (1974). Cleary, Linn and Rock (1968) describe short linear tests and a sequential test based on a procedure developed by Wald (1947). Branched tests require that a large item pool be developed for a test of moderate length. Short linear tests reduce testing time, but at the expense of non-adaptability to student item responses. The sequential test is adaptive to each examinee's response protocol and requires a small item pool. This paper presents the rationale for a specific sequential testing model, describes a validation of the technique, and discusses implications of the procedure for military training programs.

The Sequential Testing Model

The sequential test may be described as one in which an examinee is given a test item, the response to the current item as well as previously administered items is evaluated in a certain way, and the examinee is either assigned to one of a number of classes and testing stopped, or judgment is suspended and testing continues.

One method which may be employed to classify examinees requires the use of a discriminant function (Tatsuoka, 1971). Discriminant analysis is a procedure which determines the coefficients of a linear combination of variables (the discriminant function) which best discriminates between groups of individuals. Typically, discriminant functions are derived employing test scores from a battery of tests. If single items from a test are employed rather than single total test scores from a battery, then functions could be computed to provide classifications after a small number of test items had been administered. If coefficients were computed for linear combinations of all items administered at any given time, such as a linear combination of the first six items, the first seven items, and so on for as many items as are in the test, then each individual could be assigned to a classification using the discriminant functions sequentially.

As the number of variables in a discriminant function must equal or exceed the number of assignment classes, n-g sets of functions would be required for a test composed of n items and used to classify individuals into one of g groups. If these sets of functions are numbered g, g+1, g+2, ...,n, then an individual could be administered g test items, the items could be scored and the scores converted to linear combinations using discriminant function set g. If the assignment were to be into one of four classes, for example, then four values, posterior probabilities of group membership, would emerge from the discriminant functions, each probability corresponding to one of the four classes. The individual could then be assigned to the class indicated by the highest probability of group membership.

The approach described provides a potential assignment following each test item administered, for the technique results in probabilities of group membership every time a linear combination is computed. Thus, one could administer four test items, compute probabilities of group membership, assign the individual to the appropriate class, and stop testing. As an examination of Table 1 indicates, this action would not be desirable. Consider the two individuals whose scores are indicated in Table 1. Person 1 would be assigned to Group 3, as the probability of membership in Group 3 is .90. If the discriminant function had been shown to be accurate in the past, one would probably be willing to accept that classification and terminate testing. Person 2, on the other hand, would also be assigned to Group 3 as the maximum probability of group membership, .55, corresponds to Group 3. Clearly, additional testing would be desired for this person as the probability of group membership in Group 3 is low. Thus, it is desirable to adapt the discriminant function procedure to a sequential testing procedure using the probability of group membership as a guide.

The discriminant function procedure may be adapted to sequential testing by the addition of a termination rule. The varieties of termination rules are endless, but a natural class of termination rules would be related to the probability of group membership. For example, testing could continue until the maximum probability of group membership reached a certain level, say .85. In the example given in Table 1, this termination rule would permit Person 1 to be

Table 1

Classification of Two Individuals by Discriminant Analysis

	Class			
Person	1	2	3	4
1	.01	.04	.90	.05
2	.05	.10	.55	.30

Entries are posterior probabilities of group membership.

assigned to Group 3 after four items had been administered and testing would terminate for that person. Person 2 would be administered four items, but judgment would be suspended as the maximum probability of group membership, .55, did not meet or exceed the minimum required for termination of testing. Person 2 would be administered another item, and additional items, until the maximum probability exceeded .85. To assure that an assignment was made and item economy maintained, an additional rule could be incorporated which would terminate testing following some maximum number of items, say 30, if termination had not already occured. Other termination rules could be suggested, but the rule mentioned is easy to use and interpret and can serve as a point of departure for research in this area.

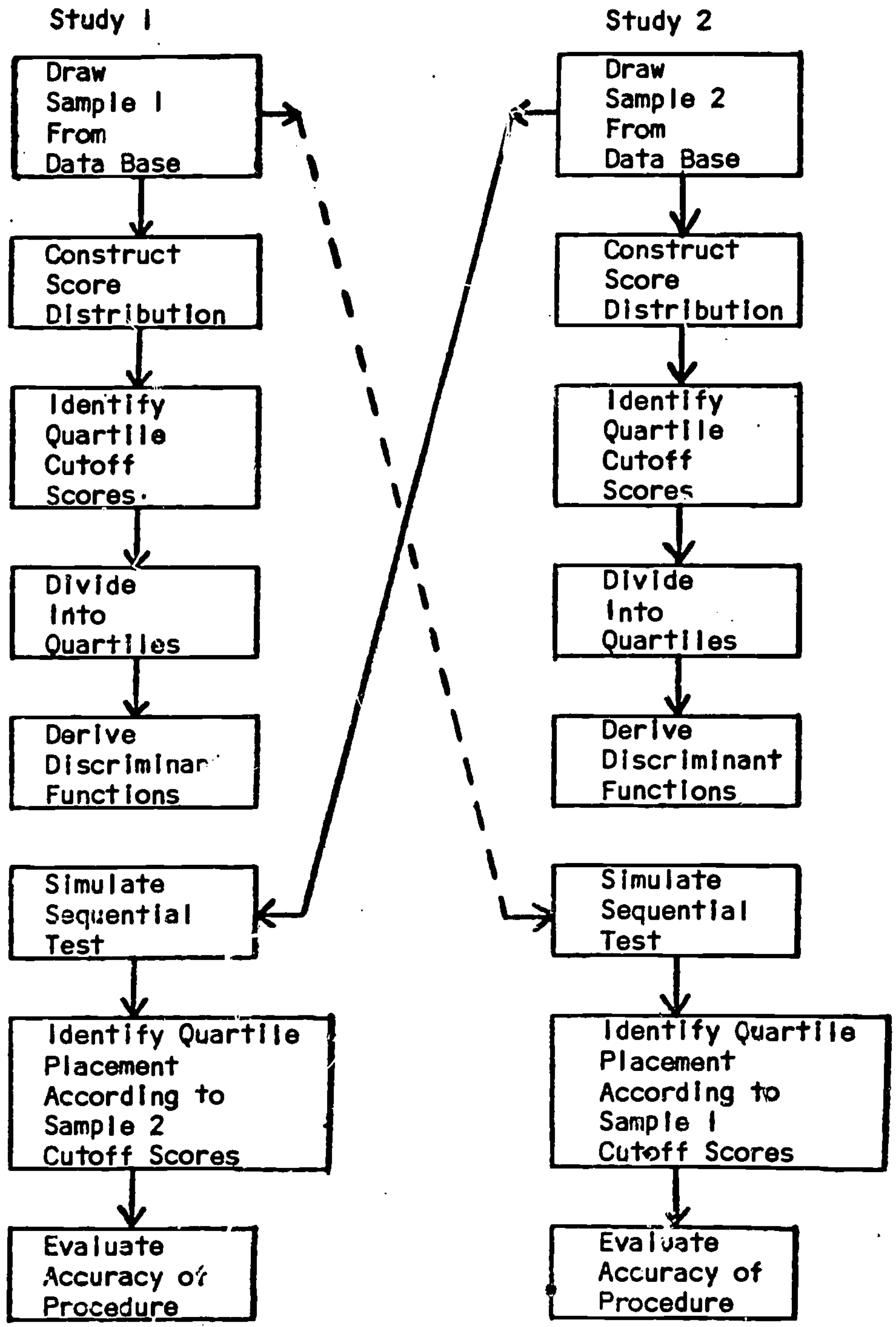
Method

Item response data from the 1971-1972 administration of the Iowa Tests of Basic Skills (Hieronymus and Lindquist, 1974) were employed to simulate the sequential testing procedure. Two samples representing 256 examinees each were drawn from a data base consisting of the response protocols of over 24,000 examinees who had completed Form 6, Level 14 (Grade 8) of the test battery. The vocabulary section consisting of 48 items was used in this investigation.

Two identical studies were conducted, with the second study constructed as a mirror-image of the first. The total number of correct responses was identified for each examinee. Score distribution tables were derived in order to identify

FIGURE I

Sequential Testing Procedure



cutoff scores for placing each examinee into one of four groups representing a quartile. In each study one sample was identified as the Developmental Sample and one sample was identified as the Validation Sample. Discriminant functions were derived employing the 256 Developmental Sample examinees in each case. The sequential testing procedure was simulated by employing the 256 Validation Sample examinees. Sample 1 was designated the Developmental Sample and Sample 2 was designated the Validation Sample for Study 1. In the second study, Samples 1 and 2 were designated the Validation and Developmental Samples respectively. Thus, each sample was alternately employed as either a validation or a developmental sample. Figure 1 depicts this procedure graphically. Table 2 presents the sample sizes and inclusive scores for each sample in each of the two studies.

The discriminant functions were derived to classify the 256 Developmental Sample examinees into four groups representing quartile placement based on the full 48 items of the test. Forty-four sets of discriminant functions were computed. Functions were derived for four variables (scores for Items 1, 2, 3, 4, 5), and in like manner, for all items through Item 48. At each stage, four sets of coefficients were computed representing the four groups to which an individual could be assigned. The coefficients generated were retained for later use by the sequential testing simulation employing the validation subjects.

Successive simulations were performed using the 45 sets of discriminant functions derived from the Developmental group's responses. The first four items were scored and converted to linear combinations using the set of coefficients for the first four items in the test. The posterior probability of membership in each group was computed and retained for later analysis. The first five test item scores were then converted to linear combinations and evaluated in a similar fashion. The process was repeated until all 48 items were included. At each step in the procedure, the probabilities of group membership were computed and retained. As each examinee's total score on the test was known, the "true group" membership was available for assessment of the accuracy of each classification. Each of the simulations concluded with a compilation of each examinee's classification history, a summary of the frequency of correct and incorrect classifications, and a summary of the average

Table 2

Sample Sizes of Developmental and Validation Groups for Studies 1 and 2

Group	Inclusive Scores	Sample Size	
		Sample 1	Sample 2
		STUDY 1	
		Developmental	Validation
1	1-21	56	71
2	22-29	68	85
3	30-35	63	42
4	36-48	69	58
Total		256	256
		STUDY 2	
		Validation	Developmental
1	1-20	49	60
2	21-26	52	67
3	27-33	67	58
4	34-48	88	71
Total		256	256

Table 3

Percent Correct Classifications and Mean Number of Items Required for Six Termination Levels in Studies 1 and 2

Term. Level	STUDY 1 Number Classified	STUDY 1 Percentage Correct	STUDY 1 Mean Items Required	STUDY 2 Number Classified	STUDY 2 Percentage Correct	STUDY 2 Mean Items Required
OVERALL						
.70	256	62	12.00	256	66	13.45
.75	255	64	13.80	253	70	15.52
.80	255	67	16.64	247	74	17.90
.85	249	71	18.34	237	81	20.75
.90	240	74	21.69	225	85	23.77
.95	209	83	25.60	193	90	27.31
GROUP 1						
.70	71	73	7.82	49	86	7.29
.75	71	75	8.44	49	90	8.74
.80	71	79	9.28	49	90	9.81
.85	70	80	10.01	48	94	11.27
.90	70	77	12.57	46	91	11.93
.95	63	86	14.52	43	95	14.51
GROUP 2						
.70	85	54	14.06	52	48	14.98
.75	85	55	16.00	51	53	17.06
.80	85	59	19.02	47	64	19.89
.85	83	65	20.82	42	74	23.47
.90	78	72	24.82	40	80	28.05
.95	67	82	29.04	29	79	33.28

Table 3 (cont)

Percent Correct Classifications and Mean Number of Items Required for Six Termination Levels in Studies 1 and 2

	STUDY 1			STUDY 2		
Term. Level	Number Classified	Percentage Correct	Mean Items Required	Number Classified	Percentage Correct	Mean Items Required
GROUP 3						
.70	42	40	16.86	67	33	17.69
.75	41	44	19.88	65	40	20.00
.80	41	49	22.56	63	44	23.98
.85	39	56	26.00	60	57	28.96
.90	35	63	28.17	54	69	33.56
.95	29	76	33.69	40	77	38.08
GROUP 4						
.70	58	76	10.59	88	91	12.75
.75	58	79	12.86	88	91	15.10
.80	58	79	17.97	88	93	16.99
.85	57	81	19.70	87	94	18.99
.90	57	79	24.63	85	95	21.93
.95	50	84	30.26	81	98	26.65

number of items required to classify the examinees under each of six termination levels. These summaries provided the primary basis for evaluation of the discriminant function approach to sequential testing.

Results

The sequential testing procedure was evaluated by inspection of the extent to which examinees were accurately assigned to the quartile in which they would be assigned on the basis of their total score. Item economy was also an evaluation factor. Overall and quartile-group tabulations of the simulation results from both studies are presented in Table 3. The table shows that for Study 1 the number classified was less than the total sample for termination levels exceeding .70 overall, and exceeding .80, .80, .70, and .80 for groups 1, 2, 3, and 4 respectively. For Study 2 the corresponding termination levels were .70, .80, .70, .70, and .80 for the entire sample and the four respective groups. The difference between the total number classified and the total sample size (Validation Sample sizes indicated in Table 2) represents those examinees for whom all 48 items were "administered" in the simulation. Data for those unclassified examinees are not represented in the computation of the results presented.

Examination of Table 3 reveals that 62% to 83% of those examinees classified were classified correctly in Study 1. Sixty-six to 90 percent were correctly classified in Study 2. A somewhat higher accuracy rate was found in Study 2 at all termination levels. Differential accuracy levels were found for the four quartile-groups in both Study 1 and Study 2. In both studies, higher accuracy was shown in groups 1 and 4 than in groups 2 and 3. In the majority of instances, the percentage correctly classified increased as the termination level increased. The number of examinees classifed overall and in each group decreased with the increase in termination level. Eighteen percent of the examinees in Study 1 and 25 percent of those in Study 2 failed to be classified by the model at the .95 termination level. At the more moderate .80 termination level, virtually all (99% and 96% for Studies 1 and 2 respectively) examinees were classified, approximately 70% correctly. Kolmogorov-Smirnov one sample tests were performed to test whether the obtained accuracy levels were those obtainable from a population with equal accuracy across groups. The tests were repeated for each termination level in the two studies. The tests showed no evidence of departure from equal accuracy levels ($p > .10$) at any termination level in Study 1 and for termination levels of .80 and above in Study 2. Significant differences from distribution congruence were indicated for termination levels of .70 ($p < .01$) and .75 ($p < .05$).

The procedure required an average of 12 to 26 items from the 48 item test in Study 1 for the 6 termination levels, and 14 to 28 items in Study 2. Differential item ecomony was shown for the four ability levels in both studies. For nearly all termination levels the ranked mean number of items required was Group 3-Group 2-Group 4-Group 1, with Group 3 requiring the highest mean number of items for classification, more than double the mean number required for Group 1.

The high error rate suggested the analysis depicted by Table 4. The table indicates that the majority of errors resulted from assignment of an examinee to an ability level one level removed from the true level (degree of error:1). Thus, at a termination level of .75, for example, all errors for Group 3 of Study 1 (56.1% of the examinees) were for classifications to either Group 4 or Group 2; no errors occurred by assigning an examinee to Group 1. Table 5 further explicates the nature of the classification errors. A termination level of .80 was selected for further analysis of the classification errors. For each error of classification, the difference between the examinees' total test score and the minimum or maximum score required for the assigned group was computed as a measure of the extent of the error. This error score is referred to as the Boundary Score Deviation in Table 5. The Boundary Score Deviation may be described as follows: If an examinee in Group 3 (Study 1) has a total test score of 30 and was classified in Group 2, the boundary score deviation is 30-29=1, as a total score of 29 would have placed the examinee in Group 2. An assignment to Group 4 would yeild a boundary score deviation of 36-30=6 as the minimum score for Group 4 was 36.

Table 5 shows, as would be expected, that the majority of classifcation errors occurred for examinees obtaining total scores close to the boundaries of the groups in which

Table 4

Frequency and Percentage of Examinees
Misclassified by 1, 2, and 3 Categories

GROUP	DEGREE OF ERROR	STUDY 1 N	STUDY 1 FREQ	STUDY 1 %	STUDY 2 N	STUDY 2 FREQ	STUDY 2 %
	TERMINATION LEVEL .70						
1	1	71	13	18.3	49	6	12.0
	2	71	1	1.4	49	0	-
	3	71	5	7.0	49	1	2.0
2	1	85	32	37.6	52	25	48.1
	2	85	7	8.2	52	2	3.8
3	1	42	24	57.1	67	40	59.7
	2	42	1	2.4	67	5	7.5
4	1	58	10	17.2	88	8	9.1
	2	58	3	5.2	88	0	-
	3	58	1	1.7	88	0	-
	TERMINATION LEVEL .75						
1	1	71	13	18.3	49	5	10.2
	2	71	1	1.4	49	0	-
	3	71	4	5.6	49	0	-
2	1	85	32	37.6	51	24	47.1
	2	85	6	7.1	51	0	-
3	1	41	23	56.1	65	35	53.8
	2	41	0	-	65	4	6.2
4	1	58	10	17.2	88	8	9.1
	2	58	1	1.7	88	0	-
	3	58	1	1.7	88	0	-

Table 4 (cont)

Frequency and Percentage of Examinees Misclassified by 1, 2, and 3 Categories

GROUP	DEGREE OF ERROR	STUDY 1 N	STUDY 1 FREQ	STUDY 1 %	STUDY 2 N	STUDY 2 FREQ	STUDY 2 %
	TERMINATION LEVEL .80						
1	1	71	12	16.9	49	5	10.2
	2	71	0	-	49	0	-
	3	71	3	4.2	49	0	-
2	1	85	32	37.6	47	17	36.2
	2	85	3	3.5	47	0	-
3	1	41	21	51.2	63	31	49.2
	2	41	0	-	63	4	6.3
4	1	58	11	19.0	88	6	6.8
	2	58	0	-	88	0	-
	3	58	1	1.7	88	0	-
	TERMINATION LEVEL .85						
1	1	70	12	17.1	48	3	6.3
	2	70	0	-	48	0	-
	3	70	2	2.9	48	0	-
2	1	83	26	31.3	42	11	26.2
	2	83	3	3.6	42	0	-
3	1	39	17	43.6	60	24	40.0
	2	39	0	-	60	2	3.3
4	1	57	10	17.5	87	5	5.7
	2	57	0	-	87	0	-
	3	57	1	1.8	87	0	-

Table 4 (cont)

Frequency and Percentage of Examinees Misclassified by 1, 2, and 3 Categories

GROUP	DEGREE OF ERROR	STUDY 1 N	STUDY 1 FREQ	STUDY 1 %	STUDY 2 N	STUDY 2 FREQ	STUDY 2 %
	TERMINATION LEVEL .90						
1	1	70	15	21.4	46	4	8.7
	2	70	0	-	46	0	-
	3	70	1	1.4	46	0	-
2	1	78	19	24.4	40	8	20.0
	2	78	3	3.8	40	0	-
3	1	35	13	37.1	54	16	29.6
	2	35	0	-	54	1	1.9
4	1	57	12	21.1	85	4	4.7
	2	57	0	-	85	0	-
	3	57	0	-	85	0	-
	TERMINATION LEVEL .95						
1	1	63	9	14.3	43	2	4.7
	2	63	0	-	43	0	-
	3	63	0	-	43	0	-
2	1	67	11	16.4	29	6	20.7
	2	67	1	1.5	29	0	-
3	1	29	7	24.1	40	8	20.0
	2	29	0	-	40	1	2.5
4	1	50	8	16.0	81	2	2.5
	2	50	0	-	81	0	-
	3	50	0	-	81	0	-

Table 5
Frequency and Percentage of Examinees Misclassified in Each Group by Boundary Score Deviation for Termination Level = .80

Study	Boundary Score Deviation	Group 1 F	Group 1 %	Group 2 F	Group 2 %	Group 3 F	Group 3 %	Group 4 F	Group 4 %
		ONE DEGREE ERROR							
1	1	8	11	6	7	5	12	5	9
	2	3	4	10	12	6	15	4	7
	3	0	0	5	6	1	2	1	2
	4	1	1	5	6	1	2	1	2
	5	0	0	4	5	4	10	0	0
	6	0	0	1	1	4	10	0	0
	7	0	0	1	1	0	0	0	0
	Total	12	17	32	38	21	51	11	19
2	1	4	8	1	2	13	21	3	3
	2	0	0	8	17	5	8	2	2
	3	0	0	2	4	6	10	0	0
	4	1	2	2	4	2	3	0	0
	5	0	0	2	4	2	3	0	0
	6	0	0	2	4	0	0	0	0
	7	0	0	0	0	1	2	0	0
	Total	5	10	17	36	31	49	6	7
		TWO DEGREE ERROR							
1	7	0	0	2	2	0	0	0	0
	8	0	0	0	0	0	0	0	0
	9	0	0	0	0	0	0	0	0
	10	0	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0
	12	0	0	1	1	0	0	0	0
	Total	0	0	1	1	0	0	0	0
2	7	0	0	0	0	0	0	0	0
	8	0	0	0	0	1	2	0	0
	9	0	0	0	0	1	2	0	0
	10	0	0	0	0	1	2	0	0
	11	0	0	0	0	1	2	0	0
	12	0	0	0	0	0	0	0	0
	Total	0	0	0	0	4	6	0	0

they were classified by the procedure. Boundary score deviations of 3 points or less account for over 90% of the errors in Groups 1 and 4, and over 50% of the errors in Groups 2 and 3. The errors are again shown to be of greater magnitude for examinees in Groups 2 and 3 in both studies.

Discussion

Evaluation of the sequential testing procedure was performed using two criteria: classification accuracy and item economy. Classification accuracy is indicated by the percentages given in Table 3 and 4. The most striking result of the analysis is the apparent differential classification accuracy from group to group. This finding is not at all clear in view of the results of the Kolmogorov-Smirnov tests which in general did not reveal significant variance from equal group to group accuracy. The nature of the K-S tests leads this researcher to suspend judgment on the differential accuracy question pending review of additional data. It may be reported, however, that the findings of two other studies by the author, lacking K-S tests, have indicated differential accuracy of the same direction and magnitude of that indicated in the present studies.

As Table 4 indicates, major classification errors occur for about 5% of the examinees, primarily those in Groups 2 and 3. The review of individual examinee classifications reported in Table 5 indicates that the incorrect classifications of a single category occur for those examinees whose total scores lie near the quartile boundaries (e.g., an examinee with a total score of 21 may be classified incorrectly in Group 2, the lowest score for which was 22 in Study 1). Thus, the classification errors are generally not major ones.

An important question is, "How accurate does the test have to be to adequately meet an instructional need?" The answer may lie in an analysis of the variety of instructional alternatives at any decision point. If three alternatives are available, then a testing procedure which differentiates between those students for whom alternative A is appropriate and those for whom alternatives B or C are more appropriate should be sufficient. It appears that the classificaton errors are not of sufficient magnitude as to preclude use of the procedure for the purpose described.

It may be concluded that the accuracy of the sequential testing procedure is probably adequate for many instructional classification decisions. Those examinees who were incorrectly classified by one quartile group, were not badly served by the procedure, as their total scores very nearly would have placed them into the assigned group. Were instruction to follow at a higher or lower level (pace, content, difficulty) any significant imprecision would doubtless be corrected through subsequent monitoring and analysis of the student's progress.

A review of Table 3 indicates that a substantial reduction in the number of items required for classification is possible in comparison to the number of items in the full test. For examinees in Groups 1 and 4, classification may be made with an average of 1/3 the items in the full test. The middle groups require approximately 1/2 the items. Thus, is is concluded that item economy is significant. The time saving could amount to 10-12 minutes per administration. Stated differently, twice the number of examinees could be administered a long series of tests in a given time period.

Certain examinee's responses could not result in classification for termination levels higher than .80. A procedure is required to provide a classification for these otherwise non-classified examinees. One procedure is to administer all items to those few non-classified examinees, if necessary, and to transform the usual total score into the quartile assignment. Another procedure would be to retain a number correct score and continue testing until sequential classification had succeeded or until the total number correct at Item N plus the sum of all future items (assumed correct) fell within the same range of the classification current for the total number correct at Item N (e.g., if 44 items had been administered from a 48 item test, and 31 were correct, testing could terminate if sequential classificaton had not occurred, as 31 and [48-44]+31 are both within the range of Group 3 [see Table 2, Study 1]). The effect of such procedures has not been evaluated. Such modifications to the method presented here are certainly possible and would increase accuracy somewhat for the few examinees affected.

Cost and time have long been important factors in the design and administration of military training. The cost of computer technology has been dropping rapidly and is currently available to many sectors of the educational enterprize, but the relative cost and effectiveness of computer-based instruction and testing is still of concern. The developmental cost of computer-based tests and the time (both human and computer) to administer them at a terminal is in many cases prohibitive. The use of adaptive tests, whether branching or sequential, will permit a more comprehensive array of tests to be administered by computer within a short time period. When existing tests are prefereable, time and cost a consideration, and computer administration desirable, the sequential procedure would be preferred over the branching procedure. This research is not conclusive as research which includes actual computer terminal administration and scoring has yet to be performed using the technique, but the results are sufficiently encouraging that research should continue on the technique, especially for developing a means for improving classification accuracy. Item economy and improved accuracy would permit one the time to measure many of the learner characteristics necessary for a prescriptive learning model.

References

Cleary, T. A., Linn, R. L., & Rock, D. A. Reproduction of total test score through the use of sequential programmed tests. Journal of Educational Measurement, 1968, 5, 183-187.

Hieronymus, A. N., & Lindquist, E. F. Iowa tests of basic skills, Forms 5 and 6: Manual for administrators, supervisors and counselors. Boston: Houghton Mifflin Company, 1974.

Tatsuoka, M. M. Multivariate analysis: Techniques for educational and psychological research. New York: John Wiley and Sons, Inc., 1971.

Wald, A. Sequential analysis. New York: John Wiley and Sons, Inc., 1947.

Weiss, D. J. Strategies of adaptive ability measurement (Research Report 74-5) Minneapolis, MN: Department of Psychology, Psychometric Methods Program, December, 1974.

Wood, R. Response contingent testing. Review of Educational Research, 1973, 43(4), 529-544