DOCUMENT RESUME

ED 187 760 TM 800 267

AUTHOR	Weiss, David J.
TITLE	Computerized Adaptive Performance Evaluation: Final Report, February 1976 through January 1980.
INSTITUTION	Minnesota Univ., Minneapolis. Dept. of Psychology.
SPONS AGENCY	Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE	Feb 80
CONTRACT	N00014-76-C-0627
NOTE	29p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Achievement Gains; *Achievement Tests; Aptitude Tests; *Computer Assisted Testing; Criterion Referenced Tests; Feedback; Intellectual Disciplines; *Latent Trait Theory; Mastery Tests; Postsecondary Education; Pretests Posttests; Scoring; Testing Problems; *Test Theory; Time Factors (Learning)
IDENTIFIERS	*Adaptive Testing; Test Length

ABSTRACT

During a three-year project (1977-1980) on computerized adaptive achievement testing, item characteristic curve theory (ICC) and adaptive testing strategies designed almost exclusively for ability testing were applied to achievement testing. Adaptive techniques substantially reduced test length without reducing quality, when applied to three problem areas: the use of unidimensional item pools, the measurement of multiple content areas, and mastery testing. The use of unidimensional ICC in achievement testing was further supported by the lack of effect on dimensionality of immediate knowledge of results. Because some results were inconclusive, more research is needed on inter-subtest branching, dimensionality of achievement over time in instruction, and adaptive mastery testing. Finally achievement measurement lacks a coherent framework. Norm referenced and criterion referenced testing approaches have little in common; and they do not relate to the crucial problem of measuring individual achievement gains during or after instruction. To incorporate inter-subtest branching, inter-time branching, and mastery testing, an adaptive self-referenced testing model is proposed. (CP)

Final Report: COMPUTERIZED ADAPTIVE PERFORMANCE EVALUATION

David J. Weiss

February 1980

COMPUTERIZED ADAPTIVE TESTING LABORATORY
PSYCHOMETRIC METHODS PROGRAM
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF MINNESOTA
MINNEAPOLIS MN 55455

TM 800267

Final Report of Project NR150-389, N00014-76-C-0627
Supported by the
Defense Advanced Research Projects Agency, Navy Personnel
and Research Development Center, Army Research
Institute, Air Force Human Resources Laboratory,
and Office of Naval Research, and monitored by
the Office of Naval Research
David J. Weiss, Principal Investigator

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) Final Report: Computerized Adaptive Performance Evaluation	5. TYPE OF REPORT & PERIOD COVERED Final Report February 1976 - January 1980
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David J. Weiss	8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0627
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology University of Minnesota Minneapolis, MN 55455	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P.E.: 61153N PROJ.:RR042-04 T.A.: RR042-04-01 W.U.: NR150-389
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research Arlington, VA 22217	12. REPORT DATE February 1980
	13. NUMBER OF PAGES 19
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This research was supported by funds from the Defense Advanced Research Projects Agency, Navy Personnel and Research Development Center, Army Research Institute, Air Force Human Resources Laboratory, and Office of Naval Research, and monitored by the Office of Naval Research.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Testing
Achievement Testing
Computerized Testing
Adaptive Testing
Sequential Testing
Branched Testing
Individualized Testing
Tailored Testing
Programmed Testing
Response-Contingent Testing
Automated Testing
Item Characteristic Curve Theory
Latent-Trait Test Theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Three years of research on computerized adaptive achievement testing are summarized. The goals of the research program are described, and the research approach is summarized and related to the nine technical reports completed under this contract. Major research findings are presented. The implications of the research methods and results for future research in computerized adaptive achievement testing are described. Also included are abstracts of the nine technical reports and a list of other project reports and papers.

Contents

FINAL REPORT:
COMPUTERIZED ADAPTIVE PERFORMANCE EVALUATION

Objectives

The original objectives of this research program were concerned with (1) the development of a psychometric basis for the construction, development, and evaluation of criterion-referenced performance tests for use in the measurement of achievement and (2) the development of psychometric methodology for computerized adaptive performance simulation tests. A performance simulation test was defined as an interactive problem-solving test in a particular area of achievement.

Research in pursuance of these objectives began in February 1976 and continued through January 1979. Technical reports were completed during the period February 1979 through January 1980.

Approach

Literature Review

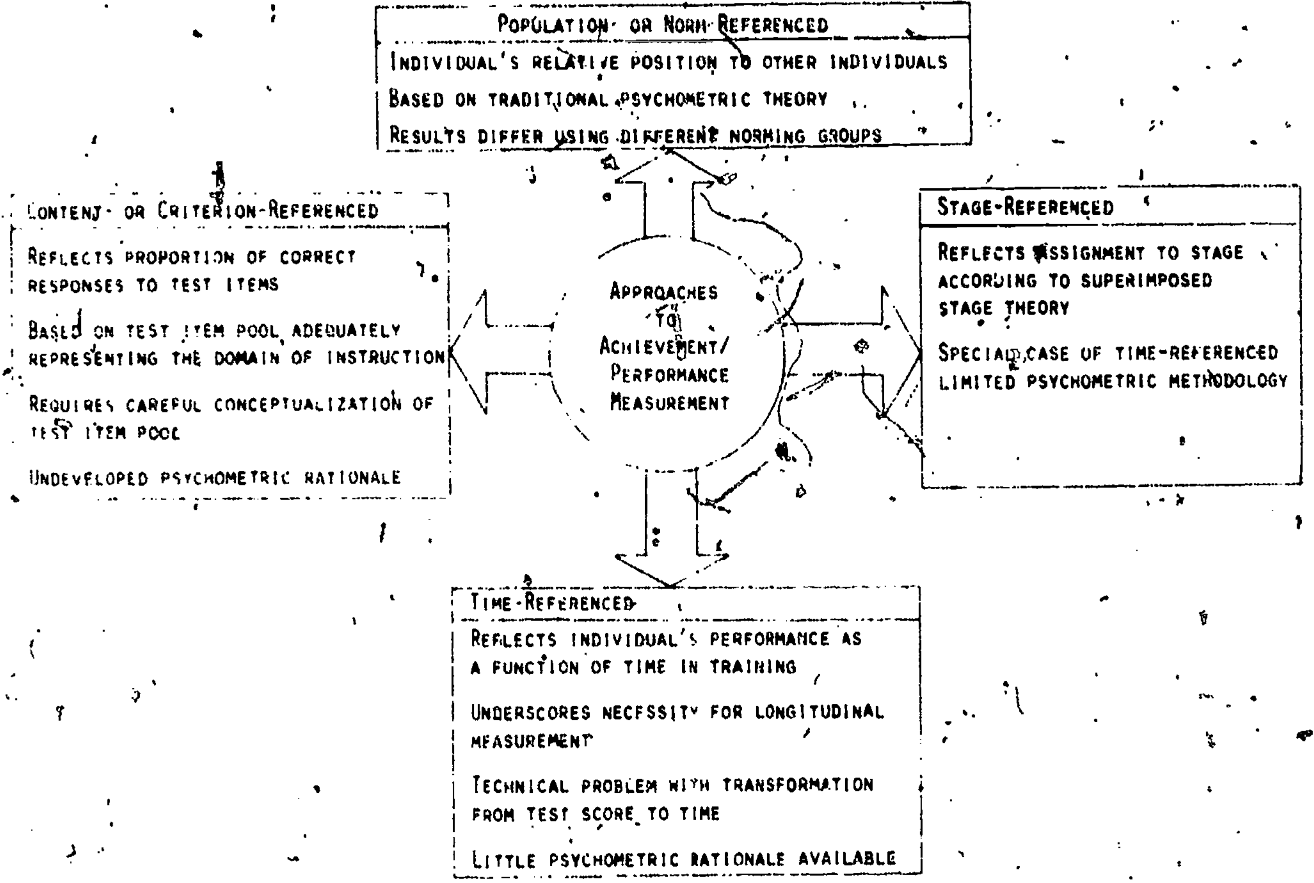
Research began with a review of the literature on the problem of the measurement of performance and achievement. Analysis of the literature concerned with the measurement of achievement led to a restructuring of project objectives.

Figure 1 summarizes the several approaches to the measurement of achievement or performance that were identified in the review of the literature. As Figure 1 shows, the measurement of achievement was determined to be considerably more complex than the related problem of ability measurement. The most prominent trend in the achievement measurement literature is the use of population- or norm-referenced techniques borrowed from the field of ability measurement. In general, these techniques have been based on classical psychological test theory, with the result that the obtained measurements and statements of achievement or performance have differed for a given individual based on the particular norming group to which the individual has been compared. In addition, the use of classical test theory for achievement measurement makes it difficult to apply adaptive testing techniques, because of the relatively large numbers of items required for adaptive testing methods based on classical test theory (Weiss, 1974).

The second major trend identified in the achievement measurement literature was that of content- or criterion-referenced measurement. The problem of criterion-referenced testing (also known as mastery testing) is quite different from that of ability testing. As a result, a serious limitation of the area of criterion-referenced measurement is that the psychometric rationale for it was relatively undeveloped. In addition, virtually no methodologies had been developed for the application of adaptive testing techniques to the problem of criterion-referenced measurement. Thus, an important objective of the project was to devise adaptive testing methodologies uniquely applicable to the problem of criterion- (or content-) referenced measurement.

FIGURE 1

APPROACHES TO ACHIEVEMENT/PERFORMANCE EVALUATION



The literature review identified a second important problem uniquely characteristic of achievement measurement, which was not characteristic of ability measurement. This problem was the fact that the measurement of achievement frequently occurs as a result of an individual's exposure to a restricted environment, such as a class or a training course. Typical of these environments is a relatively short time-frame in which the change in an individual's observed achievement level is to occur. Thus, an important problem in the area of achievement measurement is measuring an individual's achievement level over relatively short periods of time, including changes in that achievement level as a function of time.

Such an approach to measurement can be called "time-referenced" measurement, which evidences several important problems. Among these are the problem of measuring change in an individual's achievement level from one point in time to another relatively close point in time. Similar to the area of criterion-referenced measurement, there was very little psychometric rationale available in the literature for the measurement of individual gain as required by a time-referenced measurement perspective.

A special case of time-referenced measurement is that of "stage-referenced" measurement. In stage-referenced measurement, a particular theoretical structure describing stages of achievement is superimposed on the measurement problem. Thus, the achievement measurement problem becomes that of determining whether an individual is progressing in achievement levels according to the particular stage theory describing levels of achievement in the specified achievement domain. Similar to the problems of time-referenced and criterion-referenced measurement of achievement, there was very little psychometric rationale available in the literature for the stage-referenced measurement of achievement.

The review of the literature also identified several other problems that are characteristic of the measurement of achievement, as compared to the measurement of ability. One of these is that the goals of achievement measurement are frequently embodied in the specification of particular achievement domains. Frequently, these achievement domains are relatively specific; and in the process of constructing achievement tests to measure these domains, only a limited number of test items can be generated due to the specificity of the domains. Thus, the measurement of achievement frequently requires a multidimensional approach measuring specific content domains using relatively small numbers of test items in comparison to those used for the measurement of ability. As a result, traditional adaptive testing models developed in the ability testing area may not be directly applicable to the measurement of achievement. The literature thus suggested that it might be necessary to develop adaptive testing strategies for the measurement of achievement that were specifically designed to operate efficiently with a large number of small content domains.

Finally, the review of the literature and some subsequent analysis of instructional environments indicated that the measurement of performance by computerized adaptive simulation techniques was considerably more complex than had originally been anticipated. Additionally, the review indicated that there was virtually no psychometric rationale available in the literature for the measurement of performance by simulation. Although there were some applications of performance simulation to the measurement of achievement, analysis of the methodologies and attempts to apply those methodologies in relevant instructional environments indicated that the measurement of achievement by performance simulation was seriously situation-bound. That is, it was extremely unlikely that any generalizable methodologies could be developed that would be transferable across instructional situations of different types. Consequently, after some preliminary trial work with performance simulations, the objective of developing a psychometric rationale for the measurement of achievement by performance simulation was abandoned until more generalizable methodologies could be identified.

Revised Objectives

The review of the literature thus led to a redefinition of project goals. The revised project objectives were oriented around the development of adaptive testing strategies designed to address the unique problems of the measurement of achievement. The approach used was to first examine the applicability of adaptive testing strategies developed in the ability testing domain to relevant problems in the achievement testing domain. Then, further efforts were oriented toward the development of adaptive testing techniques specifically designed for the unique demands of achievement testing, and an investigation of some of the unique problems of achievement testing and analysis of some of the psychological aspects of the achievement testing environment.

Results

Applications of Item Characteristic Curve Models and Adaptive Testing Strategies

ICC models. The first technical report from the project (Research Report 77-5) investigated the question of whether item characteristic curve (ICC) theory methods utilized in ability testing were applicable to data derived from the measurement of achievement. This report described the ICC calibration of an achievement testing item pool. Data used were derived from a general biology course at the University of Minnesota. The item pool was a multiple-choice set of items written by course instructors. In addition to analyzing the applicability of ICC item calibration techniques to this item pool, the dimensionality of the pool was examined in order to determine whether unidimensional ICC theory was applicable to the measurement of this domain of achievement. Results showed that the pool was generally unidimensional and that it was possible to derive appropriate ICC parameters from this pool.

Adaptive testing strategies. Using this item pool, the next question investigated was whether adaptive testing techniques developed for the ability testing domain were applicable to the measurement of biology achievement (Research Report 77-7). A stratified-adaptive (stradaptive) test was administered to a group of students and compared with a conventional classroom test derived from the same item pool, as well as with an improved conventional test developed from the pool. Tests were compared in terms of information (precision of measurement). Results showed that, as expected, the adaptive test provided measurement of greater precision than did the conventional tests. The results also indicated that the adaptive test provided measurement of equal precision with considerably fewer numbers of items than did the conventional tests. When the average number of items administered in the adaptive test was equal to that of the conventional tests, adaptive test scores were more precise than either the classroom conventional test or the improved conventional test.

Although the demonstration of improved precision of measurement from adaptive testing in comparison to conventional testing is supportive of the general value of adaptive testing for measuring achievement, the question of the relative validity of the two techniques was also important. In Research Report 78-4 the comparative validity of adaptive and conventional achievement tests was studied. Since it is very difficult in the achievement domain to obtain a criterion against which the relative validity of two testing techniques can be evaluated, the problem was approached by comparing the respective construct validity of the two testing techniques. The results of this study showed that the construct validity of the adaptive tests was effectively higher than that of the conventional tests, since equal validities were achieved for the two testing strategies, but the adaptive tests required 25% to 35% fewer items than did the conventional tests.

Thus, these studies demonstrated the applicability of ICC techniques previously applied almost exclusively in the area of ability testing, as well as adaptive testing strategies developed for ability testing, to the problem of achievement testing. Results indicated both higher precision of measurement and higher effective levels of validity for the adaptive test.

ICC scoring methods. The process of examining the problem of the applicability of ICC theory and adaptive testing techniques to the measurement of achievement led to the development of a set of computer programs for scoring achievement test data with ICC models. Since these programs were written as general purpose programs, they were made available in Research Report 79-1 for other researchers who desired to use ICC methodologies in scoring achievement or ability tests.

In the process of implementing the reliability and validity studies comparing adaptive and conventional testing strategies, decisions had to be made about the appropriate ways of scoring the achievement test data using ICC models. These decisions were necessary for both the conventional tests and the adaptive tests. Thus, a relevant question concerned the relationships among achievement level estimates using the one-, two-, and three-parameter ICC models, as well as the maximum likelihood normal, maximum likelihood logistic, and Bayesian methods for scoring ability test data with ICC models.

To compare these scoring methods and models with each other, live data from an achievement test were scored by all combinations of ICC models and methods. The results (Research Report 79-3) indicated that highly similar achievement level estimates were derived from the one- and two-parameter data but that when the third (guessing) parameter was added to the scoring procedures, the similarities among achievement level estimates decreased. The data also indicated that the three-parameter model resulted in less similar achievement level estimates for adaptive test data than for conventional test data. However, at the same time, there were fewer convergence failures for maximum likelihood scoring in adaptive test data than there were in conventional test data.

Unique Problems of Achievement Testing

In addition to studying the applicability of ICC models and adaptive test procedures derived from ability testing to the problems of achievement testing, the project was concerned with the development of solutions to some of the unique problems raised in achievement testing, as well as the analysis of the implications of some other unique characteristics of achievement testing.

Multiple content areas. As indicated previously, one problem characteristic of achievement testing, in contrast to ability testing, is the necessity to measure an individual's achievement levels in a number of content areas at the same time. In addition, in many cases the number of items available in a content area is very restricted, resulting in relatively short tests that would not permit the application of many standard adaptive testing strategies.

Consequently, an adaptive testing strategy designed specifically for achievement test batteries was developed (Research Report 77-6). This strategy is one that is applicable to achievement tests composed of any number of short subtests. The strategy is designed to utilize both intra-subtest adaptive item selection as well as inter-subtest adaptive branching in order to reduce test battery length to a minimum for each individual. The testing strategy utilizes a maximum information ICC-based item selection technique combined with Bayesian scoring to adaptively select items within a subtest until there are no items left that provide more than trivial amounts of information about an individual's achievement level. Having obtained an achievement level estimate from one subtest, that estimate is then used in a bivariate regression equation to obtain a prior achievement level estimate in the next subtest in the test battery. The adaptive testing strategy then adaptively selects items in the next subtest, using the prior ability estimate, until no further items are available for administration in that subtest. At the end of the second subtest, multiple regression is used to obtain a prior achievement level estimate to begin testing in the third subtest, and the process is repeated until all subtests have been administered.

Results of applying this adaptive testing strategy to an achievement test battery in a military testing environment, using real-data simulation techniques, indicated an average 50% reduction in test length for the individuals tested, with no loss in the quality of the obtained measurements. Test length reductions varied from 18% to 80% across individuals. Thus, considerable reductions in number of test items administered were achieved while maintaining the quality of the measurements obtained from the conventional test.

Mastery testing. Since the methodologies required to adaptively measure mastery within a criterion-referenced framework are not the same as those available for the measurement of ability levels, an adaptive testing strategy for making mastery decisions was developed (Research Report 79-5). This testing strategy utilized ICC theory and methodologies in conjunction with a maximum information adaptive testing technique and Bayesian scoring. The testing strategy was designed to use a prespecified and flexible mastery level for comparison with each individual's performance.

The adaptive mastery testing strategy was compared with a conventional mastery test in a military training environment, using real-data simulation. When the results for the two testing strategies were compared, the adaptive mastery testing strategy reduced the average test length from 30% to 81% over all mastery decisions examined, with modal test length reductions up to 92%, yet it reached the same decision as the conventional test for 96% of trainees. Thus, again, considerable savings in the number of test items administered were observed for the adaptive test, while it made decisions which were highly similar to those made by the conventional test.

Dimensionality of achievement over time. As indicated above, a unique problem in the area of the measurement of achievement is that of measuring a person's change in achievement level over a relatively short period of time. If ICC theory is to be used in the measurement of achievement, it will gain its highest degree of potential usefulness if it can be used to measure the growth in one individual's achievement level from the beginning of instruction to later points in instruction. However, the implementation of this paradigm for the measurement of individual growth requires the demonstration that an achievement test given at two or more points in time measures the same achievement dimension and that the dimension measured is a undimensional variable. Research Report 79-4 reported results addressed to this question.

Dimensionality was investigated within the pretest-test paradigm for measuring change in achievement levels and within the test-posttest paradigm for measuring retention. Data indicated that there were some questions about the utility of the pretest-test paradigm, since a comparison of the ICC parameter estimates obtained from achievement test items at two points in time 4 weeks apart suggested a change in the dimensionality of achievement over that period of instruction. These results were also supported by the results of factor analyses. The data did, however, support the test-posttest paradigm to measure retention, since a regression comparison of students' achievement level estimates did not indicate any differences in the achievement metric up to 1 month after the completion of instruction. However, additional research is necessary in order to further verify and examine these conclusions.

Effects of knowledge of results. The advent of computerized adaptive testing also brings with it the potential of administering to students during the process of testing immediate feedback as to the correctness or incorrectness of their test responses. Previous research in the ability testing domain (Betz & Weiss, 1976a, 1976b; Prestwood & Weiss, 1978) suggests that the administration of immediate knowledge of results for each test item during the process of testing reduces the effects of extraneous variables on ability test scores. However, if immediate feedback is to be administered to students in an achievement testing environment, it is possible that the information gained from feedback on prior items may affect a student's performance on subsequent items in the test. A basic assumption of ICC theory is that of local independence, that is, that the response of a student to a given test item is the result only of the underlying achievement variable, and not of other variables. If knowledge of results from prior items in an achievement test affected a student's performance on subsequent items, the assumption of local independence would be violated.

Research Report 80-1 was concerned with this issue. In two studies, data derived from two groups of students (one of which received immediate knowledge of results while the other received no knowledge of results) on computer-administered tests were compared with each other. The results indicated essentially no systematic differences in achievement level estimates or in the dimensionality of the students' responses as a result of the administration of immediate knowledge of results. Thus, the data indicated that this added benefit of computerized administration of achievement tests did not affect the assumptions under which ICC theory could be applied in the achievement testing environment.

Major Findings

Summarized below are the major findings from this research program, with references to the research reports in which these findings are reported. In addition to these major findings, the original research reports should be consulted for additional important results and conclusions.

1. The successful application of ICC theory to achievement testing requires that the item pool be reasonably unidimensional. Analyses of a large item pool, constructed by the instructional staff of a university level course, indicated that the pool was essentially unidimensional (Research Reports 77-5 and 80-1).

2. When ICC item parameters were estimated from this item pool, the majority of the items resulted in parameter estimates that were suitable for operational testing purposes (Research Report 77-5).

3. The ICC parameter estimates obtained from this item pool reflected sufficiently high levels of discrimination and a sufficient range of difficulty to be useful in adaptive testing (Research Reports 77-5, 77-7, and 78-4).

4. Using operational achievement tests from military instructional environments, it was possible to obtain usable ICC item parameter estimates even in narrowly defined content domains (Research Reports 77-6 and 79-5).

5. The item parameter data indicate that some caution might be necessary, however, when estimating ICC item parameters in achievement test data. Relatively high discrimination parameter estimates in conjunction with high guessing parameter estimates (Research Reports 77-5, 77-6, and 79-5) may reflect a restriction in range on the achievement variable. If the effect of instruction is to eliminate individual differences in measured achievement, ICC parameter estimates of discrimination and guessing obtained on groups at their peak of instruction will be artificially inflated. Additional research on this problem is necessary.

6. ICC theory and methods, combined with specially designed adaptive testing strategies, can be useful in substantially reducing the number of items administered to trainees in an achievement test battery composed of a number of specific content domains (Research Report 77-6).

7. Both adaptive testing techniques and ICC theory and methods are useful in reducing test lengths for tests used to make mastery decisions (Research Report 79-5).

8. In a variety of applications to the problem of achievement testing--including measuring achievement with a large unidimensional item pool, measuring achievement levels in a number of specific content domains, and measuring achievement against a defined mastery criterion--adaptive testing techniques using ICC theory can substantially reduce the numbers of items required in an achievement test without reducing the quality of the measurements (Research Reports 77-6 and 79-5).

9. Adaptive testing can improve the quality of achievement measurements in terms of both precision and validity while reducing the numbers of items required (Research Reports 77-7 and 78-4).

10. ICC test scoring methods (Research Report 79-1) can be fruitfully applied to achievement testing data (Research Report 79-3). However, maximum likelihood ICC scoring is less useful in conventional tests because of its nonconvergence problem when the test is too easy or too difficult for a testee. Although non-convergences occur much less frequently in adaptive test data, use of the three-parameter ICC model with different scoring methods tends to result in somewhat different achievement level estimates. More research on this problem is indicated.

11. Because of its ability to equate testings and link item pools onto a common metric, ICC theory has the potential of offering solutions to the problem of measuring gains in achievement levels during the process of instruction. However, examination of the dimensionality of an achievement test item pool from pre-instruction to the peak of

instruction shows changes in the dimensionality of achievement during instruction (Research Report 79-4). These results, if verified with other data, suggest potential problems in the applicability of unidimensional ICC theory to the measurement of individual growth in achievement levels due to instruction.

12. The use of ICC methods to measure retention following instruction was supported by the data (Research Report 79-4). These results show that the same achievement variable was measured up to 1 month after instruction as was measured at the peak of instruction.

13. The post-instruction data (Research Report 79-4) also support the use of computerized adaptive testing in operational instructional environments. Since these data indicate that the same achievement variable is measurable up to a month after the end of instruction, instructional environments with a limited number of testing terminals can obtain similar measurements from trainees when tests are administered on different days.

14. The use of unidimensional ICC theory in achievement testing is further supported by the lack of effect on dimensionality of the administration of immediate knowledge of results during the process of achievement testing (Research Report 80-1).

Implications for Further Research

The findings and experience of this 3-year research program strongly support the use of ICC theory and methods and computerized adaptive testing for the measurement of achievement. However, many new questions were raised by the research (some of which were described above) and some of the original questions addressed are still in need of further research. Portions of the research described below are being pursued under a contract entitled "Computerized Adaptive Achievement Testing," NR150-433, with the Personnel and Training Research Programs of the Office of Naval Research, with funds from the Defense Advanced Research Projects Agency, Army Research Institute, Air Force Office of Scientific Research, and the Office of Naval Research.

Inter-Subtest Branching

Although Research Report 77-6 demonstrated that an adaptive testing strategy using intra-subtest adaptive item selection in conjunction with inter-subtest adaptive branching could substantially reduce test battery length in one achievement test battery, the generality of this finding needs to be examined. In addition, the relative efficiency of alternative approaches to inter-subtest branching needs to be studied.

The scoring strategy used in Research Report 77-6 was based on the maximum information item selection strategy using Bayesian scoring. However, the use of Bayesian scoring, which has a tendency to regress achievement estimates toward the mean, may result in the premature termination of the intra-subtest item selection, particularly when used in conjunction with the minimum information termination criterion. Thus, a relevant area of research is that of the evaluation of intra-subtest item selection strategies that may eliminate this problem and identification of situations under which use of Bayesian scoring in conjunction with maximum information item selection is less than optimal.

A second problem in intra-subtest adaptive item selection for inter-subtest branching strategies is that of the termination criterion. Research to date has utilized a termination criterion based on minimum information at the current estimated level of achievement. However, if Bayesian scoring is to be used, it is possible to terminate on the basis of a minimum posterior Bayesian variance of the achievement level estimate. The relative performance of these two termination criteria as well as their interactions with the intra-subtest item selection strategy, needs to be investigated.

With regard to branching between content areas, previous research has identified one means of ordering subtests for inter-subtest branching and has relied exclusively on linear multiple regression as the inter-subtest achievement level estimation technique. Other prediction strategies are available for making predictions between content areas and there are other ways of ordering subtests to be used in inter-subtest predictions. In addition, the use of linear multiple regression equations brings up the question of shrinkage with regard to the application of regression equations based on one sample of individuals when utilized on another sample from the same population. The effect of overestimation and shrinkage needs to be investigated within this inter-subtest branching strategy.

Finally, previous research has indicated that there is wide variability in the range of reduction in number of items administered across subtests. Thus, a relevant question is the nature of the subtests resulting in larger or smaller reductions due to the use of the inter-subtest branching strategy. This latter question is most efficiently investigated by monte carlo simulation studies in which characteristics of the subtests are systematically varied.

Dimensionality of Achievement Over Time

As indicated above, ICC theory has the potential of permitting the measurement of individual growth in achievement over time in instruction. But the initial results in Research Report 79-4 suggest that the achievement dimension changes from pretest to end-of-course-unit testing. Thus, further examination of this problem is indicated.

The investigation of the dimensionality of achievement over time is being studied in a number of achievement domains, including domains that are primarily cognitive as well as those that are primarily conceptual. Obtained data on achievement measured at various points in time will be factor analyzed. In each case, items will be parameterized by ICC models and the change of these parameters over time will be studied. In addition, achievement level estimates based on factors identified at relevant points in time will be obtained and the relationship among these achievement level estimates over time will be studied. The relative saliency of factors identified at different points in time will also be analyzed to determine whether the same factors are evident at different points in time but at different levels of saliency. If the latter hypothesis is supported by the data, it may then be possible to investigate inter-time branching, taking into account the relevant saliency of those dimensions at different points in time.

Depending on the results of the analyses of achievement level data at different points in time over a number of instructional contexts, adaptive testing strategies for inter-time branching will be developed and evaluated. If the same dimension is found to exist with different saliencies at different points in time, the utility of the information provided at the prior point in time with respect to adaptive testing at later points in time will be studied by live testing and by real-data simulation. One obvious approach would be to simply use the correlation of achievement level estimates on a normative group from earlier points in time with later points in time as entry points into later time achievement level estimation. When data are available at more than one prior point in time, the use of multivariate prediction strategies becomes relevant, and the relative advantages of different strategies will need to be investigated.

Adaptive Mastery Testing

An adaptive testing strategy for making mastery decisions was developed in Research Report 79-5. Although the data in that report indicate some promise for this ICC-based mastery testing approach, considerable additional study of its potential as a solution to the mastery testing problem is appropriate.

First, the adaptive mastery testing (AMT) strategy needs to be studied in additional mastery tests. In addition, its operating characteristics need to be examined in comparison with competitive strategies for mastery testing, including strategies based on Waldian decision theory.

The strategy also needs to be examined in a wide variety of classification situations. In one application of the AMT strategy, error may be associated primarily with the criterion, as would be the case where the items in a mastery test are all of similar difficulty and discrimination; hence, the maximum information in the item pool is concentrated around the criterion cutoff value. In a more realistic situation, errors are associated with both the criterion and the individual being measured. These different approaches to adaptive mastery testing should be compared in both real-data simulation studies and monte carlo simulation studies. The real-data simulation studies will use existing data administered in a conventional test format, from mastery tests utilized in military and educational environments, to determine the operating characteristics of these two major approaches to AMT as well as to evaluate the outcomes when both the criterion and the individual are measured with error. If differential results are obtained using these strategies in real-data simulation, it will then be appropriate to design monte carlo simulation studies to model the relevant parameters of the situation (e.g., levels of item difficulty, discrimination, and numbers of items, as well as various degrees of error on the criterion) and to compare these results with results obtained by competing strategies.

A final area of research with regard to AMT is the generalization of the methodologies to the multi-subtest mastery testing problem. Similar to the multi-subtest achievement testing problem, decisions made with regard to one subtest may be related to decisions made with regard to another subtest. Thus, research is indicated for determining how much information derived from one subtest in a multi-subtest mastery test can be used in adaptive testing in other subtests. For example, achievement level estimates generated from subtests in one content area can be used to begin adaptive testing in another content area. In some cases, only very few items will be necessary to make the mastery decisions in later subtests because of the intercorrelations among the mastery decisions and/or achievement level estimates derived from other content areas.

Adaptive Self-Referenced Testing

The review of the achievement measurement literature indicated the lack of a coherent framework for the measurement of achievement. Approaches to the measurement of achievement such as norm- or population-referenced testing and criterion-referenced (mastery) testing appeared to have nothing in common with each other and little or no implication for what appears to be the important problem in the measurement of achievement--that of measuring individual improvement in achievement levels during the process of, or as a result of, instruction. The activities of the present research program have led to the notion of Adaptive Self-Referenced Testing (ASRT), which appears to represent a coherent framework for the measurement of achievement. ASRT can incorporate into a single framework the notions of inter-subtest branching, inter-time branching, and mastery testing.

ASRT is only possible by combining computerized adaptive testing and ICC theory. It involves the measurement of growth on an individual basis, incorporating knowledge of the student's level of performance at an earlier point in time, which is used as a starting point for measurement at a later point in time. ASRT is designed to track an individual's growth in one area of achievement as a function of time. It thus can be used to identify the degree and extent of learning as it occurs and the point at which learning occurs or fails to occur during the process of instruction. The generalization of unidimensional self-referenced testing to the multidimensional case (i.e., where more than one content area is being measured) incorporates the inter-test branching problem. The objective is to utilize, on an individual basis, information gained both on other tests and at prior time periods for the measurement of growth in learning (achievement).

ASRT is unique in that the sequence of measurements taken to measure each individual's learning history is based only on that individual's prior performance at earlier points in time in the same content domain. It is also designed to operate uniquely within both computer-assisted and computer-managed instruction. If properly implemented, it should be an extremely powerful approach for measuring achievement in these contexts, permitting a continuous evaluation of student progress and a non-normative definition of "when learning has occurred and how much has been learned," while reducing testing time to a minimum for each student.

References

Betz, N. E., & Weiss, D. J. Effects of immediate knowledge of results and adaptive testing on ability test performance (Research Report 76-3). Minneapolis: Department of Psychology, Psychometric Methods Program, June 1976. (NTIS No. AD A027147)

Betz, N. E., & Weiss, D. J. Psychological effects of immediate knowledge of results and adaptive testing on ability test performance (Research Report 76-4). Minneapolis: Department of Psychology, Psychometric Methods Program, June 1976. (NTIS No. AD A027170)

Prestwood, J. S., & Weiss, D. J. The effects of knowledge of results and test difficulty on ability test performance and psychological reactions to testing (Research Report 78-2). Minneapolis: Department of Psychology, Psychometric Methods Program, September 1978.

Weiss, D. J. Strategies of adaptive ability measurement (Research Report 74-5). Minneapolis: Department of Psychology, Psychometric Methods Program, December 1974. (NTIS No. AD A004270)

Abstracts of Research Reports

Research Report 77-5
Calibration of an Item Pool for the Adaptive Measurement of Achievement
Isaac I. Bejar, David J. Weiss, and G. Gage Kingsbury
September 1977

The applicability of item characteristic curve (ICC) theory to a multiple-choice test item pool used to measure achievement is described. The rationale for attempting to use ICC theory in an achievement framework is summarized, and the adequacy for adaptive testing of a conventional classroom achievement test item pool in a college biology class is studied. Using criteria usually applied to ability measurement item pools, the item difficulties and discriminations in this achievement test pool were found to be similar to those used in adaptive testing pools for ability testing. Studies of the dimensionality of the pool indicate that it is primarily unidimensional. Analysis of the item parameters of items administered to two different samples reveals the possibility of a deviation from invariance in the discrimination parameter, but a high degree of invariance for the difficulty parameter. The pool as a whole, as well as two subpools, is judged to be adequate for use in adaptive testing. It is also concluded that the ICC model is not inappropriate for application to typical college classroom achievement tests similar to the one studied.

Research Report 77-6
An Adaptive Testing Strategy for Achievement Test Batteries
Joel M. Brown and David J. Weiss
October 1977

An adaptive testing strategy is described for use with achievement tests that cover multiple content areas. The testing strategy combines adaptive item selection both within and between the subtests in the multiple-subtest battery. A real-data simulation was conducted in order to compare the results from computerized adaptive testing with those from conventional paper-and-pencil testing, in terms of test information and test length. Data for the simulation consisted of test results for 365 fire-control technicians on a paper-and-pencil administration of a 232-item achievement test, which was divided into 12 subtests, each covering a different content area. Correlations between subtest scores from adaptive and conventional testing were .90 or higher for 11 of the 12 content areas. An information analysis showed that for all 12 subtests, the subtest information curves from adaptive testing were essentially identical to the corresponding subtest information curves from conventional testing. On the average, the number of items administered with adaptive testing was half as many as was required with conventional testing; the shortest adaptive test battery used 18% of the total number of items in the conventional test, while the longest used 18%. The adaptive testing strategy, therefore, provided a considerable reduction in test length and virtually no loss in precision of measurement when compared with the conventional administration of the achievement test battery.

Research Report 77-7

An Information Comparison of Conventional and Adaptive Tests in the Measurement of Classroom Achievement

Isaac I. Bejar, David J. Weiss, and Kathleen A. Gialluca

October 1977

The information provided by typical and improved conventional classroom paper-and-pencil achievement tests is compared with the information provided by an adaptive test covering the same subject matter. Both tests were administered to over 700 students in a general biology course. Using the same scoring method, adaptive testing was found to yield substantially more precise estimates of achievement level than the conventional test throughout the entire range of achievement, while at the same time reducing the length of the test. The comparison of the improved conventional test with the stradaptive test also indicated that the scores derived from the adaptive test were more precise, even in the range of achievement where the improved test was designed to be optimal. An analysis of the effects of expanding an adaptive test item pool indicates that even when slightly more discriminating items are added to the pool, improved precision of measurement can result. A comparison of response pattern information values (observed information) with test information values (theoretical information) shows that the observed information consistently underestimates theoretical information, although the pattern of results from the two procedures is quite similar. It is concluded that the adaptive measurement of classroom achievement results in scores that are less likely to be confounded by errors of measurement and, therefore, are more likely to reflect a testee's true level of achievement. In addition, the reduction in number of test items administered by the adaptive measurement of achievement can result in additional time spent in instruction.

Research Report 78-4

Construct Validation of Adaptive Achievement Testing

Isaac I. Bejar and David J. Weiss

November 1978

The construct validities of conventional classroom paper-and-pencil and adaptive achievement tests were compared using data from two independent groups of 269 and 230 college students. Two adaptive achievement tests were computer administered to each group using the stradaptive testing strategy; each group also completed two conventional classroom paper-and-pencil achievement tests. All achievement tests were drawn from the same pool of achievement test items on which item characteristic curve (ICC) parameters had been determined. Students were also administered two stradaptive vocabulary tests. All tests were scored by maximum likelihood estimation using the three-parameter logistic model. A nomological net was specified, describing the relationships of the achievement tests to the achievement constructs and their relationships with the vocabulary construct and the vocabulary tests. The parameters of the net were estimated by fitting the observed intercorrelations among the test scores to the nomological net, using the methodology of linear structural equations. Maximum likelihood estimates of the parameters of the nomological net indicated essentially equal validities for the conventional and adaptive tests in four comparisons. However, the validity of the adaptive tests was effectively higher than that of the conventional tests, since equal validities were achieved with from 25% to 31% fewer items. The data also permitted an analysis of the effects of verbal ability on achievement test performance, separately for the conventional and adaptive tests. The results from a confirmatory maximum likelihood factor analysis showed a larger influence of verbal ability on achievement test performance at the first administration of the adaptive test. This result was attributed to a necessity to learn how to use the computer equipment with verbal instructions, which may have further reduced the validity of the adaptive tests. Combined with the facts that the adaptive tests were obtained under volunteer conditions while the conventional tests were obtained under "motivated" grading conditions, the results of this study indicate that computer-administered adaptive tests can provide more valid measurement of achievement than conventional classroom paper-and-pencil tests.

Research Report 79-1

Computer Programs for Scoring Test Data with Item Characteristic Curve Models

Isaac I. Bejar and David J. Weiss

February 1979

Three computer programs are described for scoring test response data using item characteristic curve (ICC), or latent trait, models. The rationale and mathematical basis of both maximum likelihood and Bayesian ICC scoring methods are presented, as well as some data comparing the two methods of scoring. The three computer programs are designed for scoring conventional (linear) test data (LINDSCO) in dichotomous response format, adaptive test dichotomous data (ADADSCO), and conventional (linear) test data scored by polychotomous ICC models (LINPSCO). Options available in these three general purpose programs are described, and examples of the input and output are given for each program. Complete FORTRAN listings of the three programs are included.

Research Report 79-3

Relationships Among Achievement Level Estimates from Three Item Characteristic Curve Scoring Methods

G. Gage Kingsbury and David J. Weiss

April 1979

This study compared achievement level estimates from three item characteristic curve (ICC) scoring methods using the one-, two-, and three-parameter ICC models. The three scoring methods were maximum likelihood normal, maximum likelihood logistic, and Owen's (1975) Bayesian scoring method. Data included all possible response patterns from a hypothetical five-item test, as well as response patterns from live administration of a conventional classroom and a computerized adaptive achievement test. For the conventional and adaptive test data, correlations among achievement level estimates were examined as a function of test length. Results for all data sets showed a high degree of similarity among θ estimates for the one- and two-parameter data, with slight decreases in correlations as information on the discrimination parameter was used in scoring. When the third ("guessing") parameter was used in scoring the item response data, correlations among θ estimates were reduced, particularly for the adaptive test data. The data also showed an increasing tendency for the maximum likelihood methods to result in convergence failures as the third parameter of the ICC was used in scoring. In general, however, the adaptive test data were less likely to result in convergence failures than were the conventional test data. The data also illustrated how each of the three scoring methods tend to utilize ICC parameter information in arriving at θ estimates and the relationships of these estimates to a number correct scoring philosophy. Advantages and disadvantages of each of the scoring methods are discussed. It is suggested that future research examine the relative validities of scoring methods and model combinations.

Research Report 79-4

Effect of Point-in-Time in Instruction on the Measurement of Achievement

G. Gage Kingsbury and David J. Weiss

August 1979

Item characteristic curve (ICC) theory has potential for solving some of the problems inherent in the pretest-test and test-posttest paradigms for measuring change in achievement levels. However, if achievement tests given at different points in the course of instruction tap different achievement dimensions, the use of ICC approaches and/or change scores from these tests is not desirable. This problem is investigated in two studies designed to determine whether or not achievement tests administered at different times during a sequence of instruction actually measure the same achievement dimensions. To investigate possible changes in dimensionality between different points in instruction, aspects of the dimensionality of achievement test data were examined prior to instruction, at the peak of instruction, and up to a month following the peak of instruction. Data used were conventional and adaptive achievement test data administered to students in a general biology course at the University of Minnesota. Results raised questions about the utility of the pretest-test paradigm for measuring change in achievement levels, since a comparison of ICC parameter estimates indicated that a change in the dimensionality of achievement had occurred within the short (4-week) period of instruction. This change was also observed using a factor analytic comparison. Use of the test-posttest paradigm to measure retention was supported, since a regression comparison of students' achievement level estimates did not indicate any significant change in the achievement metric up to 1 month after the peak of instruction. The significance of this result for the use of adaptive testing technology in measuring achievement is described. Implications of these studies and the use of ICC theory in the measurement of achievement, as well as some potential limitations in terms of generalizability of these results, are discussed.

Research Report 79-5

An Adaptive Testing Strategy for Mastery Decisions

G. Gage Kingsbury and David J. Weiss

September 1979

In an attempt to increase the efficiency of mastery testing while maintaining a high level of confidence for each mastery decision, the theory and technology of item characteristic curve (ICC) response theory (Lord & Novick, 1968) and adaptive testing were applied to the problem of judging individuals' competencies against a prespecified mastery level to determine whether each individual is a "master" or a "nonmaster" of a specified content domain. Items from two conventionally administered classroom mastery tests administered in a military training environment were calibrated using the unidimensional three-parameter logistic ICC model. Then, using response data originally obtained from the conventional administration of the tests, a computerized adaptive mastery testing (AMT) strategy was applied in a real-data simulation. The AMT procedure used ICC theory to transform the arbitrary "proportion correct" mastery level used in traditional mastery testing to the ICC achievement metric in order to allow the adaptation of the test to each trainee's achievement level estimate, which was calculated after each item response. Adaptive testing continued until the 95% Bayesian confidence interval around the trainee's achievement level estimate failed to contain the prespecified mastery level. At that point testing was terminated, and a mastery decision was made for the trainee. Results obtained from the AMT procedure were compared to results obtained from the traditional mastery testing paradigm in terms of the reduction in mean test length, information characteristics, and the correspondence between decisions made by the two procedures for three different mastery levels and for each of the two tests. The AMT procedure reduced the average test length 30% to 81% over all circumstances examined (with modal test length reductions of up to 92%) while reaching the same decision as the conventional procedure for 96% of the trainees. Additional advantages and possible applications of AMT procedures in certain classroom situations are noted and discussed, and further research questions are suggested.

Research Report 80-1

[illegible] of Immediate Knowledge of Results on Achievement Test Performance and Test Dimensionality

Kathleen A. Gialluca and David J. Weiss

January 1980

These two studies investigated the effects of administering immediate knowledge of results (KR) concerning the correctness or incorrectness of each item response on a computerized adaptive test of Biology achievement. In the case of incorrect responses, the correct answers were provided to the student. The results of these studies indicate that the provision of informative KR did not systematically increase total test scores, as would be expected if students were using information from previously administered items to help them answer subsequent items. Furthermore, provision of informative KR did not alter the dimensionality of the achievement tests administered, indicating that the latent trait model assumption of local independence among the items was not affected to any significant degree.

Other Papers and Publications

Bejar, I. I. A comparison of conventional and computer-based adaptive achievement testing. In D. J. Weiss (Ed.), Proceedings of the 1977 computerized adaptive testing conference. Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, 1978.

Bejar, I. I. Application of adaptive testing in measuring achievement and performance. In D. J. Weiss (Ed.), Applications of computerized adaptive testing (Research Report 77-1). Minneapolis: University of Minnesota Department of Psychology, Psychometric Methods Program, March 1977. (Paper presented at a symposium at the 18th annual convention of the Military Testing Association, October 1976.)

Kingsbury, G. G., & Weiss, D. J. Differences among three latent-trait-based methods for estimating achievement levels. Paper presented at the 1978 Convention of the American Psychological Association, Toronto, Ontario, Canada, August 1978.

Weiss, D. J. Computerized adaptive testing for the measurement of achievement and performance. Paper presented at the conference on Innovations in Instructional Systems Development, Wayzata, MN, August 1977.

Weiss, D. J. Computerized adaptive achievement testing. In Harold F. O'Neil (Ed.), Procedures for instructional systems development. New York: Academic Press, 1979.

Weiss, D. J., & Brown, J. M. Multi-content adaptive measurement of achievement. In D. J. Weiss (Ed.), Proceedings of the 1977 computerized adaptive testing conference. Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, 1978.

DISTRIBUTION LIST

Navy

1 Dr. Jack R. Borsting
Provost & Academic Dean
U. S. Naval Postgraduate School
Monterey, CA 93940

1 [illegible] (Pers-11B)
Dept. of Navy
Washington, DC 20370

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

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

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Patrick R. Harrison
Psychology Course Director
LEADERSHIP & LAW DEPT. (7b)
DIV. OF PROFESSIONAL DEVELOPMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. George Moeller
Head, Human Factors Dept.
Naval Submarine Medical Research Lab
Groton, CN 06340

1 Dr William Montague
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. William Moonan
Code 303
Navy Personnel R&D Center
San Diego, CA 92152

1 LCDR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, CA 93940

1 Commanding Officer
U.S. Naval Amphibious School
Coronado, CA 92155

1 Library
Naval Health Research Center
P. O. Box 85122
San Diego, CA 92138

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

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

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

1 Director, Navy Personnel R&D Center
Washington Liason Office
Building 200, 2N
Washington Navy Yard, DC 20374

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Scientific Director
Office of Naval Research
Scientific Liaison Group/Tokyo
American Embassy
APO San Francisco, CA 96503

1 Special Asst. for Education and
Training (OP-01E)
Rm. 2705 Arlington Annex
Washington, DC 20370

1 Office of the Chief of Naval Operations
Research, Development, and Studies Branc
(OP-102)
Washington, DC 20350

1 Long-Range Manpower, Personnel, and
Training Planning Branch (OP-110)
Room G828
Arlington Annex
Washington, DC 20350

1 Head, Manpower Training and Reserves
Section (Op-964D)
Room 4A478, The Pentagon
Washington, DC 20350

1 Captain Donald F. Parker, USN
Commanding Officer
Navy Personnel R&D Center
San Diego, CA 92152

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 DR. RICHARD A. POLLAK
ACADEMIC COMPUTING CENTER
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Dr. Gary Poock
Operations Research Department
Code 55PK
Naval Postgraduate School
Monterey, CA 93940

1 The Principal Deputy Assistant
Secretary of the Navy (MRA&L)
4E780, The Pentagon
Washington, DC 20350

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Ronald Weitzman,
Code 54 WZ
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 DR. H.M. WEST III (OP-01)
Head, Program Development Branch(OP-120)
ARLINGTON ANNEX
WASHINGTON, DC 20350

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Director, Research & Analysis Division
Plans and Policy Department
Navy Recruiting Command
4015 Wilson Boulevard
Arlington, VA 22203

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA [illegible]

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL [illegible]

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC [illegible]

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL [illegible]

Army

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

1 Mr. J. Barber
HQS, Department of the Army
DAPE-ZBR
Washington, DC [illegible]

1 Col Gary W. Bloedorn
US Army TRADOC Systems Analysis Activity
Attn: ATAA-TH
WSMR, NM [illegible]

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA [illegible]

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

1 Dr. Ed Johnson
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA [illegible]

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

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

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

1 LTC Michael Plummer
Chief, Leadership & Organizational
Effectiveness Division
Office of the Deputy Chief of Staff
for Personnel
Dept. of the Army
Pentagon, Washington DC 20301

1 DR. JAMES L. RANEY
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

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

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

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

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

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

Air Force

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

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

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

1 Dr. Philip De Leo
AFHRL/TT
Lowry AFB, CO 80230

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Research and Measurment Division
Research Branch, AFMPC/MPCYPR
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Dr. Marty Rockway (AFHRL/TT)
Lowry AFB
Colorado 80230

1 Dr. Frank Schufletowski
U.S. Air Force
ATC/XPTD
Randolph AFB, TX 78148

1 Jack A. Thorpe, Maj., USAF
Naval War College
Providence, RI 02846

1 Brian K. Waters, Lt.Col, USAF
Air War College (EDV)
Maxwell AFB, AL 36112

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

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

1 Major Michael L. Patrow, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

1 DR. A. L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

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

[illegible]

12 Defense [illegible] Center
Cameron Station, Bldg.
Alexandria, VA [illegible]
Attn: [illegible]

1 Dr. [illegible] Fletcher
ADVANCED RESEARCH PROJECT AGENCY
1400 WILSON BLVD.
ARLINGTON, VA [illegible]

1 Dr. William Gr[illegible]
Testing Directorate
[illegible]
Ft. Sheridan, [illegible]

1 Director, Research and Data
OASD(MRA&L)
3B919, The Pentagon
Washington, DC [illegible]

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

Civil Govt

1 Dr. Lorraine D. Eyde
Personnel R&D Center
Office of Personnel Management of USA
1900 EStreet NW
Washington, D.C. 20415

1 Jerry Lehnus
REGIONAL PSYCHOLOGIST
U.S. Office of Personnel Management
230 S. DEARBORN STREET
CHICAGO, IL 60604

1 William J. McLaurin
Rm. 301, Internal Revenue Service
2221 Jefferson Davis Highway
Arlington, VA 22202

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

1 Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

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

Non Govt

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

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

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

1 Dr. Jackson Beatty
Department of Psychology
University of California
Los Angeles, CA 90024

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

1 DezWPs im Streitkraefteamt
Postfach 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Lyle Bourne
Department of Psychology
University of Colorado
Boulder, CO 80309

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

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

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

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

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, DC 20036

1 Dr. Kenneth B. Cross
Anacapa Sciences, Inc.
P.O. Drawer Q
Santa Barbara, CA 93102

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

1 Dr. Leonard Feldt
Lindquist Center for Measurment
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
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

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

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
Suite 900
4330 East West Highway
Washington, DC 20014

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

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

1 Dr. Ross Greene
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

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

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

1 Dr. Frederick Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

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

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

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

1 Dr. Douglas H. Jones
Rm T-255
Educational Testing Service
Princeton, NJ 08450

1 Journal Supplement Abstract Service
American Psychological Association
1200 17th Street N.W.
Washington, DC 20036

1 Dr. Arnold F. Kanarick
Honeywell, Inc.
Honeywell Plaza MN12-3166
Minneapolis, MN 55408

1 Professor John A. Keats
University of Newcastle
AUSTRALIA 2308

1 Dr. Mazie Knerr
Litton-Mellonics
Box 1286
Springfield, VA 22151

1 Dr. Stephen Kosslyn
Harvard University
Department of Psychology
33 Kirkland Street
Cambridge, MA 02138

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

1 Dr. Michael Levine
210 Education Building
University of Illinois
Champaign, IL 61820

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

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

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

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

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77004

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

1 Dr. Mark Miller
Computer Science Laboratory
Texas Instruments, Inc.
Mail Station 371, P.O. Box 225936
Dallas, TX 75265

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

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

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

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

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

1 MINRAT M. L. RAUCH
P.II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
4 Hill Hall
Columbia, MO 65211

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

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

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

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

1 Dr. Kathryn T. Spoehr
Department of Psychology
Brown University
Providence, RI 02912

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 Dr. David Stone
ED 236
SUNY, Albany
Albany, NY 12222

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

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

1 Dr. Brad Sympson
Psychometric Research Group
Educational Testing Service
Princeton, NJ 08541

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

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

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

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 Dr. John Wannous
Department of Management
Michigan University
East Lansing, MI 48824

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

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

PREVIOUS PUBLICATIONS

Proceedings of the 1977 Computerized Adaptive Testing Conference. July 1978.

Research Reports

80-1. Effects of Immediate Knowledge of Results on Achievement Test Performance and Test Dimensionality. January 1980.
79-7. The Person Response Curve: Fit of Individuals to Item Characteristic Curve Models. December 1979.
79-6. Efficiency of an Adaptive Inter-Subtest Branching Strategy in the Measurement of Classroom Achievement. November 1979.
79-5. An Adaptive Testing Strategy for Mastery Decisions. September 1979.
79-4. Effect of Point-in-Time in Instruction on the Measurement of Achievement. August 1979.
79-3. Relationships among Achievement Level Estimates from Three Item Characteristic Curve Scoring Methods. April 1979.
Final Report: Bias-Free Computerized Testing. March 1979.
79-2. Effects of Computerized Adaptive Testing on Black and White Students. March 1979.
79-1. Computer Programs for Scoring Test Data with Item Characteristic Curve Models. February 1979.
78-5. An Item Bias Investigation of a Standardized Aptitude Test. December 1978.
78-4. A Construct Validation of Adaptive Achievement Testing. November 1978.
78-3. A Comparison of Levels and Dimensions of Performance in Black and White Groups on Tests of Vocabulary, Mathematics, and Spatial Ability. October 1978.
78-2. The Effects of Knowledge of Results and Test Difficulty on Ability Test Performance and Psychological Reactions to Testing. September 1978.
78-1. A Comparison of the Fairness of Adaptive and Conventional Testing Strategies. August 1978.
77-7. An Information Comparison of Conventional and Adaptive Tests in the Measurement of Classroom Achievement. October 1977.
77-6. An Adaptive Testing Strategy for Achievement Test Batteries. October 1977.
77-5. Calibration of an Item Pool for the Adaptive Measurement of Achievement. September 1977.
77-4. A Rapid Item-Search Procedure for Bayesian Adaptive Testing. May 1977.
77-3. Accuracy of Perceived Test-Item Difficulties. May 1977.
77-2. A Comparison of Information Functions of Multiple-Choice and Free-Response Vocabulary Items. April 1977.
77-1. Applications of Computerized Adaptive Testing. March 1977.
Final Report: Computerized Ability Testing, 1972-1975. April 1976.
76-5. Effects of Item Characteristics on Test Fairness. December 1976.
76-4. Psychological Effects of Immediate Knowledge of Results and Adaptive Ability Testing. June 1976.
76-3. Effects of Immediate Knowledge of Results and Adaptive Testing on Ability Test Performance. June 1976.
76-2. Effects of Time Limits on Test-Taking Behavior. April 1976.
76-1. Some Properties of a Bayesian Adaptive Ability Testing Strategy. March 1976.
75-6. A Simulation Study of Stradaptive Ability Testing. December 1975.
75-5. Computerized Adaptive Trait Measurement: Problems and Prospects. November 1975.
75-4. A Study of Computer-Administered Stradaptive Ability Testing. October 1975.
75-3. Empirical and Simulation Studies of Flexilevel Ability Testing. July 1975.
75-2. TETREST: A FORTRAN IV Program for Calculating Tetrachoric Correlations. March 1975.
75-1. An Empirical Comparison of Two-Stage and Pyramidal Adaptive Ability Testing. February 1975.
74-5. Strategies of Adaptive Ability Measurement. December 1974.
74-4. Simulation Studies of Two-Stage Ability Testing. October 1974.
74-3. An Empirical Investigation of Computer-Administered Pyramidal Ability Testing. July 1974.
74-2. A Word Knowledge Item Pool for Adaptive Ability Measurement. June 1974.
74-1. A Computer Software System for Adaptive Ability Measurement. January 1974.
73-3. The Stratified Adaptive Computerized Ability Test. September 1973.
73-2. Comparison of Four Empirical Item Scoring Procedures. August 1973.
73-1. Ability Measurement: Conventional or Adaptive? February 1973.

Copies of these reports are available, while supplies last, from:

Computerized Adaptive Testing Laboratory
Psychometric Methods Program, Department of Psychology
N660 Elliott Hall, University of Minnesota
75 East River Road, Minneapolis, Minnesota 55455