DOCUMENT RESUME

ED 163 970 IR 006 641

AUTHOR Kaskowitz, David; Suppes, Patrick
TITLE Development and Implementation of a Computer Model for Student Management: Phases I and II. Interim Report, April 1977-December 1977.
INSTITUTION Stanford Univ., Calif.
SPONS AGENCY Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO AFHRL-TR-78-7
PUB DATE Mar 78
CONTRACT F33615-77-C-0041
NOTE 61p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Computer Managed Instruction; Continuous Progress Plan; *Curriculum Development; Educational Research; Goodness of Fit; *Mathematical Models; Military Training; Predictive Validity; Predictor Variables; Statistical Analysis; *Technical Education

ABSTRACT

Results are described for the first two phases of a study to develop and evaluate models of student progress in a technical training course being offered in a computer managed instructional environment at Lowry Air Force Base, Colorado. Several categories of models were developed in Phase I and defined by model type and by the type of variables used as predictors. In Phase II the models were evaluated using data collected on 368 students in an inventory management course. The evaluation consisted of parameter estimation and derivation of goodness of fit statistics. The results indicated that by using the performance information on the initial blocks of the course, more precise predictions of course completion time can be made than when only preassessment data are used. (Author/CMV)

AFHRL-TR-78-7



AIR FORCE HUMAN RESOURCES LABORATORY

# DEVELOPMENT AND IMPLEMENTATION OF A COMPUTER MODEL FOR STUDENT MANAGEMENT:

## PHASES I AND II

By

David Kaskowitz
Patrick Suppes
Stanford University
Stanford, California 94305

TECHNICAL TRAINING DIVISION
Lowry Air Force Base, Colorado 80230

March 1978
Interim Report for Period April 1977 — December 1977

Approved for public release; distribution unlimited.



AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This interim report was submitted by Stanford University, Stanford, California 94305, under contract F33615-77-C-0041, project 1121, with Technical Training Division, Air Force Human Resources Laboratory (AFSC), Lowry Air Force Base, Colorado 80230. Dr. Roger J. Pennell, Instructional Technology Branch, was the contract monitor.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

MARTY R. ROCKWAY, Technical Director
Technical Training Division

DAN D. FULGHAM, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER AFHRL-TR-78-7 | 2. GOVT ACCESSION NO. |
| 3. RECIPIENT'S CATALOG NUMBER | |
| 4. TITLE (and Subtitle) DEVELOPMENT AND IMPLEMENTATION OF A COMPUTER MODEL FOR STUDENT MANAGEMENT: PHASES I AND II | 5. TYPE OF REPORT & PERIOD COVERED Interim April 1977 – December 1977 |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) David Kaskowitz, Patrick Suppes | 8. CONTRACT OR GRANT NUMBER(s) F33615-77-C-0041 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Stanford University, Stanford, California 94305 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62205F 11210220 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235 | 12. REPORT DATE March 1978 |
| | 13. NUMBER OF PAGES 62 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Technical Training Division, Air Force Human Resources Laboratory, Lowry Air Force Base, Colorado 80230 | 15. SECURITY CLASS. (of this report) Unclassified |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

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

computer managed instruction
course completion times
mathematical models
student progress management

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


This report describes the results of the first two phases of a study to develop and evaluate models of student progress on a technical training course being offered in a computer managed instructional environment at Lowry Air Force Base, Colorado. Several categories of models were developed in Phase I. The categories were defined by model type and by the type of variables used as predictors.

In Phase II, the models were evaluated using data collected on 368 students in an inventory management course. The evaluation consisted of parameter estimation and derivation of goodness of fit statistics. The results indicated that by using the performance information on the initial blocks of the course, more precise predictions of course completion time can be made than when only preassessment data are used.


DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AFHRL-TR-78-7

## PREFACE

The authors wish to acknowledge the assistance received from the staff at SRI International who worked on this project under a subcontract to Stanford University. Dr. John Draper served as SRI's project supervisor. Mr. Booker Thomas served as SRI's project leader in the early portion of the project. Ms. Fran Adams made a major contribution to the literature review conducted in Phase I.

TABLE OF CONTENTS

LIST OF TABLES

# DEVELOPMENT AND IMPLEMENTATION OF A COMPUTER MODEL FOR STUDENT MANAGEMENT: Phases I and II

## I INTRODUCTION

### Background

One of the major advantages in computer-assisted instruction (CAI) and computer-managed instruction (CMI) is that they can be individualized so that students may proceed through a curriculum at their own pace. In contrast to conventional instructional organization, where students proceed in a lockstep fashion, a CAI/CMI environment enables students to proceed at a pace consistent with their own abilities and motivations. Thus, slower students can have time to master material and will not get lost or retard the progress of other students. Similarly, faster students will tend to get less bored when they can proceed at their own pace. By allowing for self-pacing, an instructional program gives students more autonomy and can be more efficient than conventional programs that may necessarily be geared to the pace of the slower students in the class.

Although self-paced instruction has many advantages, it does introduce some problems needing solution. Notably, if students are not given information regarding standards of performance, they may be at a loss as to how to pace themselves through a course. A totally self-paced system can create administrative problems as well if certain actions necessary at the completion of a course require preparation days or weeks in advance and there is uncertainty as to when a student will complete the course. Furthermore, for CMI or CAI to be run efficiently, it is important that students are progressing at a pace consistent with their ability and that students who are not progressing in such a manner are identified and appropriate remedial action is taken. Remedial action may include provision of special assistance or application of incentives tied to course progress. On the other hand, a schedule of incentives for rewarding exceptional rates of progress through a course also requires that some criterion be established regarding student progress.

A student progress management system (SPMS) can enhance the effectiveness of self-paced instruction by providing information to students, instructors, and administrators on expected and actual rates of progress through a course of instruction. The student is informed of what is expected, the instructor is provided information necessary for monitoring student progress, and the administrator is provided information necessary for planning outprocessing activities. At the heart of such a system are the procedures by which expectations of student progress are determined from baseline and initial performance information on the student. The individual identification of expected performance maintains the individualization of the instructional system while imposing enough structure for efficient operation of the system in terms of optimizing student flow.

The Technical Training Division, Air Force Human Resources Laboratory (AFHRL), Lowry Air Force Base, Colorado, funded this study by Stanford University to develop and implement a model of student progress as part of a student progress management system in the Advanced Instructional System (AIS) implemented at Lowry. This report describes the results of the design phases (Phases I and II) of the study and includes recommendations for tasks to be carried out during Phase III. Phases I and II consisted of a review of pertinent literature, formulation of alternative models of student progress, evaluation of the models using actual data collected by the AIS, and recommendations for a system to be implemented during Phase III.

Section II contains a description of the models of student progress that were examined and a description of the methods used in the evaluation of models; Section III contains a description of the results; finally, Section IV contains the conclusions of the study and recommendations for Phase III. In the remainder of this chapter is presented a brief description of the AIS implemented at Lowry AFB, including the current student progress management system. The detailed report of the literature review is contained in Appendix A.

## Overview of the AIS

The AIS at Lowry Air Force Base now includes four courses: Inventory Management, Materiel Facilities, Precision Measurement Equipment, and Weapons Management. Each course consists of a series of lessons grouped into blocks of instruction. To complete a lesson or a block of lessons, a student must satisfactorily complete a lesson or block criterion progress check consisting of a written test, a performance test, or instruction certification. For some lessons a number of alternative modules of instruction are available that differ in mode and level of presentation.

Student assignments to lessons and modules are controlled by a computerized component of the AIS called the Adaptive Model. Assignments to lessons and modules are based on student characteristics and on material and personnel utilization in order to optimize overall system efficiency by minimizing predicted student course time and maximizing the use of course materials and personnel. As a result, the sequence of lesson presentation and the type of module presented for a given lesson vary across students. The sequence with which blocks of lessons are presented, however, is usually fixed across students.

The current AIS system includes a student progress management component (Dallman, 1977) that was pilot tested early in 1977 and fully implemented beginning in July 1977. The predictive model used in the student progress management component consists of a linear regression of course completion time on baseline variables consisting of student demographic characteristics and preassessment test scores. Predicted course completion time can be converted to targeted course completion time by a policy function which can truncate extremely low or high predicted course completion times and can reduce or raise predicted course completion times by a fixed percentage. For example, at the present time predicted course completion times are decreased by a fixed percentage to take into account the two hours per day that students are to devote to study outside of the learning center. Once a targeted course completion time is calculated, targeted time on a particular block or lesson is calculated by using the proportion of mean course completion time accounted for by the mean time on the block or lesson.

At the beginning of a course, a student is instructed on how to make his own Course Completion Map, which plots his targeted completion times and allows the student to keep track of his daily progress in the course. Two checkpoints are indicated on the Course Completion Map: at the end of block 2 and 8 days before the scheduled course completion date. The student has the responsibility to see his instructor at these times to verify that he is progressing satisfactorily, and, in the case of the graduation date checkpoint, to set a graduation date on the calendar.

If a student's actual time to arrive at some point in the course is greater than his targeted time by a specified amount (2 days), the student is to initiate a Progress Counseling Session to confer with the instructor regarding problems and potential solutions. Unless a student can convince the instructor that his targeted times are incorrect, the student and instructor will enter into a "performance contract" that specifies the date on which the student will get back on the original schedule.

In the next section, we turn to a description of the specific models of student progress that we evaluated and the methods used in the evaluation.

## II DESCRIPTION OF THE MODELS AND THE METHODOLOGY

### Description of the Models of Student Progress

Very few mathematical models of student progress were found in the process of the literature review (see Appendix A). In addition, past studies (Larsen, Markosian, and Suppes, 1977; Wagner et al., 1973; Malone et al., 1977) indicated that relatively simple mathematical models tend to provide the best predictions of student performance. Therefore, the evaluation was restricted to the trajectory and milestone models described below. These models were considered to be the best candidates based on the findings of the literature review.

A trajectory model includes a hypothesized relationship between cumulative study time and cumulative achievement. Cumulative achievement is a construct representing a student's location in the curriculum. It is generally measured on some index representing average student performance. If A represents the value of the achievement index and T represents cumulative time, a trajectory model hypothesizes a functional relationship between T and A:

$$T = f(A; a_1, a_2, \ldots, a_n),$$

where the $a_i$ are parameters. By estimating the parameters, one attempts to describe or predict a student's trajectory through a course.

In a milestone model, cumulative time to any particular milestone in a course is predicted directly, without an achievement index as an intervening variable. Course milestones would consist of particular identifiable points of interest in the curriculum, such as the completion of particular blocks or lessons.

Models can be characterized by the kind of data used for predictions as well as by type. Baseline models employ baseline variables that are available before a student starts on a course. Demographic information such as age and sex, academic history such as number of years of schooling and highest degree attained, and scores on cognitive and affective tests administered before a student's entry into the course may be included as baseline variables. Performance models employ data regarding a student's initial performance on a course to predict subsequent performance. Performance variables may consist of time to particular criteria or scores on tests. Mixed models, where both baseline and performance variables are included, may also be formulated.

The specific models reported in the evaluation are described below under five headings:

| | |
|---|---|
| Baseline Milestone Models | (BMM) |
| Performance Milestone Models | (PMM) |
| Baseline Trajectory Models | (BTM) |
| Performance Trajectory Model | (PTM) |
| Mixed Milestone Model | (MMM) |

In the description of the models the following notation will be used:

$X_{kj}$ = baseline variable k for student j.

$t_{ij}$ = the cumulative time in the course for student j to complete the ith milestone.

$A_{ij}$ = the cumulative achievement index for student j after completion of the ith milestone.

$\hat{t}_{ij}$ = the predicted cumulative time in the course for student j to complete the ith milestone.

*Baseline Milestone Models*

BMM1 — Regression of Cumulative Times to Achieve Milestones on Baseline Variables. The model hypothesizes a linear relationship between cumulative time to achieve each milestone and the baseline variables. For each milestone, say i, we would have:

$$\hat{t}_{ij} = a_i + \sum_k b_{ik} X_{kj} \ .$$

BMM2 — Regression of Course Completion Time on Baseline Variables. This model is identical to BMM1 in using a linear model to predict course completion time. Time to complete a particular segment of the course, however, is predicted by using the ratio of mean time to complete the segment to mean course completion time. If there are n milestones in the course, then course completion time is expressed as

$$\hat{t}_{nj} = a_n + \sum_k b_{nk} X_{kj} \ .$$

For any particular milestone i let $p_i = \mu_i/\mu_n$, where

$\mu_i$ = the mean cumulative time to achieve milestone i

$\mu_n$ = the mean course completion time.

Then this model predicts that for milestone i

$$\hat{t}_{ij} = p_i t_{nj} \ .$$

*Performance Milestone Model*

PMM — Regression of Course Completion Time on Performance Variables. The performance milestone model uses the cumulative times on the first i milestones to predict the cumulative course time:

$$\hat{t}_{nj} = a_i + \sum_{m=1}^{i} b_{mi} t_{mj} \ .$$

*Baseline Trajectory Models*

BTM1 — Linear Trajectory Model. The linear trajectory model hypothesizes a linear relationship between cumulative time and an achievement index:

$$\hat{t}_{ij} = a_j + b_j A_{ij} \ .$$

In turn, it is assumed that $a_j$ and $b_j$ have a linear relationship to the baseline variables:

$$a_j = c + \sum_k d_k X_{kj}$$

$$b_j = c' + \sum_k d'_k X_{kj}$$

BTM2 — Nonlinear Trajectory Model. The nonlinear trajectory model hypothesizes a linear relationship between some power of cumulative time and an achievement index:

$$(\hat{t}_{ij})^{K_j} = a_j + b_j A_{ij} \ .$$

Again, the parameters $K_j$, $a_j$, and $b_j$ are assumed to have a linear relationship to the preassessment scores.

*Performance Trajectory Model*

PTM – Linear Trajectory Model. This model corresponds to BTM1 in hypothesizing a linear relationship between cumulative time and cumulative achievement. However, in this case, the parameters $a_j$ and $b_j$ are estimated directly using the first i observed cumulative times. The course completion time $\hat{t}_{nj}$ for student j is predicted to be:

$$\hat{t}_{nj} = a_{ji} + b_{ji} A_{nj} \quad ,$$

using the times on the first i milestones to estimate $a_{ji}$ and $b_{ji}$.

*Mixed Milestone Model*

The mixed milestone model (MMM) uses the baseline variables as well as the cumulative times on the first i milestones to predict course completion time:

$$\hat{t}_{nj} = a_i + \sum_{m=1}^{i} b_{mi} t_{mj} + \sum_{k} c_{ik} X_{kj}$$

*Other Models*

Several other models were considered, but complete evaluations were not conducted because intermediate results indicated that they were not viable alternatives. As is pointed out in the literature review in Appendix A, the results of Malone et al. (1977) in evaluating the predictive fit of various trajectory models indicated that the relatively simple models had the best predictive fit. Therefore, the models included in the evaluation were the more simple models that appeared to have the greatest chance of providing a good predictive fit.

**The Data Base**

Data and corresponding documentation were obtained for the Inventory Management (IM) course as it had been implemented in late 1976 and early 1977. The data consisted of 15,259 records from the Recent Data File (RDF), where each record contained information for a particular student on a particular block of instruction. It should be noted that the data used in the evaluation of alternative models were collected in the absence of a student progress management system. At the time of data collection, the SPMS was in a pilot phase, and not enough students had completed the course under the SPMS to permit use of their data in the evaluation.

The IM course is organized into six blocks of instruction. Each block consists of a series of lessons, followed by a period during which the student reviews the material on the block and takes the block test. With the exception of the first block, the lesson prior to block review consists of a "chief of supply lab." If the chief of supply labs and block reviews are counted as lessons, there are 61 lessons in the course, organized as shown in Table 1.

Each record from the RDF contained data for a particular block of instruction or contained summary data for the entire course (block zero records). After preliminary tabulations of the data, a file was created that consisted of data on 760 students who had entered the IM course during 1977 and had reliable block elapsed time data for all six blocks of instruction. Reliable block times are defined as times resulting from actual real-time system tracking. Table 2 shows the number of records retained at each step of the process by which this file was created.

Table 1

NUMBER OF LESSONS PER BLOCK OF INSTRUCTION IN THE INVENTORY MANAGEMENT COURSE

| Block | Number of Lessons |
|---|---|
| 1 | 9 |
| 2 | 10 |
| 3 | 12 |
| 4 | 9 |
| 5 | 12 |
| 6 | 9 |
| Total | 61 |

Table 2

STATISTICS ON EDIT OF THE RECENT DATA FILE

| | |
|---|---|
| Number of records on original file | 15,259 |
| Number of students represented | 2,539 |
| Number of students represented who entered the course in 1977 | 998 |
| Number of students who entered the course in 1977 who had records for all seven blocks | 766 |
| Number of above who had all preassessment data available | 760 |
| Number out of 760 students with all block elapsed times (BELTs) reliable | 550 |

As was recommended to us, only students who had entered the IM course during 1977 were included in the evaluation. Also, students had to have finished the course, as indicated by the presence of all seven blocks of information on the RDF, in order to be included. Since cumulative time to finish each block was a critical variable of interest, it was decided to include only those cases where all the block times were reliable. This resulted in deleting about 30% of the cases from the evaluation. The distribution of included students by course version is shown in Table 3.

The tabulation indicates that the great majority of students were in course version 1. This was the course version where the Adaptive Model determined lesson sequences. On particular lessons, it determined the module to be assigned from a set of alternatives. Also, the student progress management system was not implemented in course version 1 at the time the data were collected. The subsequent tabulations will include only the data for the 368 students in course version 1 with all block elapsed times reliable.

Table 3

DISTRIBUTION OF STUDENTS BY COURSE VERSION

| Course Version | Number With All BELTs Reliable | Percent of Total | Total Number |
|---|---|---|---|
| 1 | 368 | 72 | 508 |
| 2 | 10 | 59 | 17 |
| 3 | 33 | 70 | 47 |
| 4 | 43 | 77 | 56 |
| 5 | 52 | 76 | 68 |
| 6 | 44 | 69 | 64 |
| | 550 | | 760 |

The file included the variables listed in Table 4. All times on the file were expressed in minutes. In the creation of this file the following conventions were followed:

- If Measured Time to Lesson Criterion (LTMC) was greater than 600 minutes or if the flags on the RDF (LDPF and MLTR) indicated that LTMC was unreliable, then an imputed time was calculated by multiplying the mean LTMC for the lesson by the ratio of the sum of the student's available LTMCs to the sum of the corresponding mean LTMCs.

  Of the 760 students on the file, 81% had at least one LTMC that had to be imputed. The mean number of imputations was 4.6. Of all the lesson times, 8% had to be imputed. Of the 368 students in course version 1 with all block elapsed times reliable, 68% had at least one LTMC that had to be imputed. The mean number of imputations was 3.0, representing 5% of the LTMCs for this group.

- The sequence of lesson presentation was determined using the date and time of day (LCDT and LCTM) that the lesson criterion was met as indicated on the RDF. If LCDT or LCTM was missing or if the sequence of lessons for a block was inadmissible as indicated by the course hierarchy charts, the most common lesson sequence for the block was assigned to the student.

  Of the 760 students on the file, 83% had at least one lesson sequence imputed on the first four blocks. (The last two blocks had a fixed sequence of lessons.) In all, 41% of the lesson sequences on the first four blocks were imputed. Of the 368 students in course version 1 with all block elapsed times reliable, 77% had at least one lesson sequence imputed on the first four blocks and a total of 36% of the lesson sequences on the first four blocks were imputed. Although these percentages of imputed sequences are high, the effect on the analysis is probably negligible since sequencing of lessons probably has a very small effect on time to criterion.

- If a block elapsed time (BELT) was equal to zero, or if the flag on the RDF (BLTR) indicated that BELT was unreliable, then BELT was considered missing and assigned a value of –1.

Preliminary tabulations indicated that there were relatively large differences between the actual block elapsed time variables, defined as block elapsed time minus absence time, and the corresponding sums of lesson time to criterion. Table 5 summarizes the differences found.

Table 4

VARIABLES INCLUDED IN THE ANALYSIS FILE

| Variable Title | Variable Label | Description and Comments |
|---|---|---|
| Student I.D. | SI | Unique I.D. assigned to each student. |
| Course Version* | CRSVSN | Course version student enrolled in. |
| Course Entry Date* | CRSEDT | Date student entered course, (i.e., took preassessment). |
| Module Number | MNOLi | Module number for 17 lessons where alternative modules are available (i = 1, . . . , 17). |
| Sex* (b) | SEX | Code representing student's sex. |
| Highest School Year Completed* (b) | HIYEAR | Highest school year completed. |
| Student's Age at Course Entry* (b) | ENTAGE | Student's age in years, at course entry (rounded to nearest year). |
| Reading Vocabulary General Scale* (b) | RVOCGN | Student's score on the reading vocabulary, general scale (preassessment). |
| Reading Vocabulary Scientific Scale* (b) | RVOCSC | Student's score on the reading vocabulary scientific scale (preassessment). |
| Reading Vocabulary Total Scale* (b) | RVOCTL | Student's score on the reading vocabulary total scale (preassessment). |
| Pre-Course State Curiosity* (b) | STCUR | Student's score on the pre-course state curiosity scale (preassessment). |
| Pre-Course State Anxiety* (b) | STANX | Student's score on the pre-course state anxiety scale (preassessment). |
| Trait Curiosity* (b) | TRCUR | Student's score on the trait curiosity scale (preassessment). |
| Trait Anxiety* (b) | TRANX | Student's score on the trait anxiety scale (preassessment). |
| Internal-External Scale* (b) | IESCL | Student's score on the internal-external scale (preassessment). |
| Test Anxiety* (b) | TSTANX | Student's score on the test anxiety scale (preassessment). |
| Preference for Audio Mode* (b) | PREFA | Student's score on the audio preference scale of the General Media Preference Test (preassessment). |

*Documentation on this variable was taken from "DEP Variables List" provided by AFHRL.
Note: (b) indicates the variable is included in the set of baseline variables.

Table 4 (Continued)

| Variable Title | Variable Label | Description and Comments |
|---|---|---|
| Preference for Visual Mode* (b) | PREFV | Student's score on the visual preference scale of the General Media Preference Test (preassessment). |
| Preference for Printed Mode* (b) | PREFP | Student's score on the printed preference scale of the General Media Preference Test (preassessment). |
| Experience with Self Pacing* (b) | EXPSP | Student's score on the experience with self pacing scale of the General Media Preference Test (preassessment). |
| Experience with Conventional Instruction* (b) | EXPCI | Student's score on the experience with conventional Instruction scale of the General Media Preference Test (preassessment). |
| IM/MF Reading Subscale I* (b) | READS1 | Student's score on the IM/MF reading skills test, subscale 1 (preassessment). |
| IM/MF Reading Subscale 2* (b) | READS2 | Student's score on the IM/MF reading skills test, subscale 2 (preassessment). |
| IM/MF Reading Total Scale* (b) | READST | Student's score on the IM/MF reading skills test, total scale (preassessment). |
| IM/MF Logical Reasoning Scale* (b) | LOGREA | Student's score on the IM/MF logical reasoning scale (preassessment). |
| Concealed Figures Scale* (b) | CONFIG | Student's score on the concealed figures scale (preassessment). |
| Memory For Numbers Backward Scale* (b) | MEMNB | Student's score on the IM/MF memory for numbers test, backward, scale (preassessment). |
| Memory For Numbers; Total Scale* (b) | MEMNT | Student's score on the IM/MF memory for numbers test, total scale (preassessment). |
| Block Elapsed Time | BELTi | Regular elapsed classroom time while student was in block (i = 1, . . . , 6). |
| Cumulative Actual Block Elapsed Time | CABELTi | Regular elapsed classroom time up to the completion of the ith block excluding absence time (i = 1, . . . , 6). |
| Measured Time Absent | TIi | Time absent during the ith block (i = 1, . . . , 6). |
| Cumulative Block Achievement Index | CBAIi | Value of the achievement index at the end of the ith block (i = 1, . . . , 6). |
| Measured Time to Lessons Criterion | LTMCi | Measured time spent by the student on lesson i until he first passed it (i = 1, . . . , 61). |
| Cumulative Lesson Elapsed Time | CLETi | Measured time spent by the student until he first passed the ith lesson presented to him (i = 1, . . . , 61). |
| Cumulative Achievement Index | CAIi | Value of the achievement index at the end of the ith presented lesson (i = 1, . . . , 61). |

*Documentation on this variable was taken from "DEP Variables List" provided by AFHRL.
Note: (b) indicates the variable is included in the set of baseline variables.

Table 5

DIFFERENCES BETWEEN CUMULATIVE BLOCK ELAPSED TIMES AND CORRESPONDING CUMULATIVE LESSON TIMES TO CRITERION (Minutes)

| Block | Mean Difference | S.D. |
|---|---|---|
| 1 | 187 | 266 |
| 2 | 538 | 475 |
| 3 | 1185 | 839 |
| 4 | 1538 | 978 |
| 5 | 1968 | 1151 |
| 6 | 2218 | 1210 |

Discussion with AFHRL personnel indicated that differences are due primarily to:

- Inclusion of administrative lost time due to shift open and close at the block level and not at the lesson level.
- Intermittent omission of block remediation time after a failure on the initial block test.

After some further tabulations of such variables are Measured Materials Remediation Time (T3), it was decided to carry out the analysis separately on the block elapsed times and on the lesson time to criterion to see whether prediction using one of these sets was better than prediction using the other set.

Summary statistics for the variables included in the analysis are tabulated in Appendix B. These include means and standard deviations as well as selected correlations among variables.

The 24 variables indicated by a "b" in the "Variable Title" column in Table 4 comprised the set referred to as baseline (or preassessment) variables in the description of the models. A few other preassessment variables that appeared on the RDF were excluded from the evaluation because of lack of variation.

**Procedure for Creating the Achievement Index**

An index of achievement was needed for the trajectory models. After considering alternative approaches to defining the achievement index, it was decided to use mean lesson time scaled by mean course completion time. That is, each lesson was assigned an achievement value by dividing the mean measured time to lesson criterion by the sum of the mean measured times to lesson criterion. The ratio was then multiplied by 100 so that the cumulative achievement index would indicate the percentage of the course completed at any particular time. In this way, the achievement index is linearly related to the mean cumulative time to complete each block, calculated by summing the appropriate lesson times.

Table 6 shows the value of the achievement index at the end of each block of instruction and the cumulative block times normalized so that total course time equals 100. Differences between the achievement index and the normalized block times are extremely small, indicating that the values of the index at the end of blocks would have been virtually the same if the block elapsed times had been used.

 

Table 6

COMPARISON OF THE ACHIEVEMENT INDEX AND NORMALIZED BLOCK TIMES

| End of Block | Cumulative Achievement Index | Cumulative Block Times (Normalized) | Observed Percentage of Cumulative Lesson Time at the End of Each Block | |
|---|---|---|---|---|
| | | | Mean | S.D. |
| 1 | 13.9 | 13.0 | 14.4 | 3.2 |
| 2 | 32.2 | 31.0 | 33.1 | 5.4 |
| 3 | 55.4 | 55.5 | 55.8 | 6.2 |
| 4 | 70.3 | 70.2 | 70.2 | 5.8 |
| 5 | 86.9 | 87.4 | 86.8 | 3.6 |
| 6 | 100.0 | 100.0 | 100.0 | - |

The table also gives the mean and standard deviation of the percentage of cumulative lesson time at the end of each block for the 368 students in course version 1. The mean cumulative percentage of time spent to finish each block corresponds closely to the cumulative achievement index.

### Parameter Estimation

The SPSS software as operationalized on the IMSSS PDP-10 was used to generate parameter estimates. For the milestone models, the cumulative elapsed times served as dependent variables in a stepwise regression on the specified independent variables. For the baseline trajectory models, a two-stage approach was used. In the first stage, parameters were estimated for each student. The estimated parameters were then entered as dependent variables in a stepwise regression with the baseline variables as independent variables. In the performance trajectory models, the parameters were estimated for each student separately using the initial cumulative elapsed times.

The regressions that included the baseline variables were conducted in two runs. In the first run, the stepwise regression included all 24 baseline variables. The results of this run were examined and a second run was made that included only the most salient variable, that is, only those that increased the square of the multiple correlation coefficient by at least .003. This criterion was selected to reduce the number of variables in the final equation; it tended to reduce the number of baseline variables from 24 to less than 10. The criterion is rather liberal in including variables that contribute relatively little to the regression.

### Evaluation Measures

Since deviation of elapsed time from targeted time is critical to the student progress management system, the evaluation measures were selected as functions of the distribution of residuals, defined as observed elapsed time minus predicted elapsed time. The statistics generated for each model included:

- The mean residual
- The median residual

- The standard deviation of the residuals
- The mean absolute residual
- The root mean square residual.

For the baseline models, these statistics were generated for cumulative elapsed time to the end of each block. For the performance models, they were generated for cumulative elapsed time to course completion.

AFHRL-TR-78-7

# III RESULTS

The results of the parameter estimation are described below for each model. Comparisons of the goodness of fit and predictive accuracy of the models are then made.

## Results for BMM1

Table 7 contains the summary statistics for the regression of the cumulative block elapsed times on the baseline variables. The value of the multiple correlation coefficient, $R^2$, has very little change between the run with all variables entered and the run with only the most salient variables entered (see Section II for a description of the methodology). The $R^2$ values stay relatively stable across blocks, with the lowest value of .24 for the first block. The $R^2$ values indicate that the preassessment scores are accounting for between 24% and 30% of the variance of the cumulative block elapsed time. This corresponds to a multiple correlation coefficient of between .49 and .55. The standard errors of estimation also indicate, as might be expected, that the error in estimation increases with block number. We expect this effect because the magnitude of the cumulative block times will be increasing with block number.

The variables that enter the stepwise regression in the initial steps and account for most of the $R^2$ include the total score on the IM/MF reading skills test (READST), sex, and the score on the experience with conventional instruction scale of the General Media Preference Test (EXCPI).

The statistics shown in Table 8 on the regression of cumulative lesson time to the end of each block on the baseline scores are similar to those at the block level. The multiple R does tend to decrease somewhat between the first and sixth block. The three variables that entered the stepwise regression first at the block level also enter first at the lesson level. Table 9 shows selected regression statistics at the end of lessons where all students would have been presented the same set of lessons. The first few lessons in Block 1 have extremely low $R^2$ values, but by the fifth lesson the value of $R^2$ is already up to .28.

## Results for BMM2

The BMM2 model does not require additional regression runs. It merely uses the BMM1 results for course completion time.

## Results for PMM

Table 10 presents the statistics on the performance milestone model at both the block and the lesson levels. In this case, the dependent variable is the remaining time to course completion and the independent variables are initial cumulative block or lesson times. For comparison, the regression statistics for course completion time on the baseline milestone model are also included.

The column labeled "$R^2$" contains the square of the multiple correlation of the course completion time with the initial cumulative times. The column labeled "$R^2$ for Remaining Course Time" contains the square of the multiple correlation of the remaining course time with the initial cumulative times. The standard error of estimation is the same for both the case when course completion time is the dependent variable and the case when remaining course time is the dependent variable. Under the column labeled "Mean Time Remaining" are the average times remaining in the course at the end of each block. This column is included for reference to indicate the magnitude of time remaining.

At both the block and the lesson levels, there is a substantial increase in the $R^2$ values between the baseline regressions and the regression using the performance information on the first block. Of course, the

AFHRL-TR-78-7

Table 7

SUMMARY STATISTICS FOR THE REGRESSION OF THE CUMULATIVE BLOCK ELAPSED TIMES ON THE BASELINE VARIABLES

| At the end of block: | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $R^2$ with all variables entered | .24 | .31 | .30 | .30 | .29 | .28 |
| $R^2$ on the truncated run | .24 | .30 | .30 | .30 | .30 | .28 |
| Standard error of estimation | 389 | 743 | 1231 | 1466 | 1739 | 1888 |
| Variable entered ($R^2$) | | | | | | |
| Step 1 | READST (.09) | READST (.14) | READST (.14) | READST (.13) | READST (.13) | READST (.12) |
| 2 | SEX (.16) | SEX (.20) | SEX (.20) | SEX (.20) | SEX (.19) | SEX (.18) |
| 3 | MEMNB (.18) | LOGREA (.24) | EXPCI (.25) | EXPCI (.25) | EXPCI (.24) | EXPCI (.23) |
| 4 | EXPCI (.20) | EXPCI (.26) | LOGREA (.26) | LOGREA (.26) | LOGREA (.26) | LOGREA (.24) |
| 5 | LOGREA (.21) | TRCUR (.28) | TRCUR (.27) | TRCUR (.27) | TRCUR (.27) | TRCUR (.26) |
| 6 | TRCUR (.22) | HIYEAR (.28) | IESCL (.28) | IESCL (.28) | IESCL (.28) | IESCL (.26) |
| 7 | HIYEAR (.22) | STCUR (.29) | STCUR (.29) | STCUR (.29) | STCUR (.28) | HIYEAR (.27) |
| 8 | RVOCTL (.23) | TSTANX (.29) | HIYEAR (.29) | HIYEAR (.29) | HIYEAR (.29) | STCUR (.27) |
| 9 | STCUR (.23) | IESCL (.30) | MEMNT (.30) | TSTANX (.30) | TSTANX (.29) | TSTANX (.28) |
| 10 | STANX (.24) | | | | | |

AFHRL-TR-78-7

Table 8

SUMMARY STATISTICS FOR THE REGRESSION OF THE CUMULATIVE LESSON ELAPSED TIMES ON THE BASELINE VARIABLES

| AT the end of block: | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $R^2$ with all variables entered | .30 | .30 | .28 | .26 | .24 | .23 |
| $R^2$ on the truncated run | .29 | .28 | .27 | .25 | .23 | .22 |
| Standard error of estimation | 286 | 582 | 931 | 1149 | 1336 | 1468 |
| Variable entered ($R^2$) | | | | | | |
| Step 1 | READST (.12) | READST (.12) | READST (.11) | READST (.11) | READST (.10) | READST (.10) |
| 2 | SEX (.18) | EXPCI (.17) | EXPCI (.16) | EXPCI (.16) | SEX (.14) | SEX (.14) |
| 3 | EXPCI (.23) | SEX (.21) | SEX (.21) | SEX (.21) | EXPCI (.19) | EXPCI (.18) |
| 4 | RVOCGN (.25) | LOGREA (.23) | LOGREA (.23) | TRCUR (.22) | TRCUR (.20) | TRCUR (.19) |
| 5 | TRCUR (.27) | TRCUR (.25) | TRCUR (.24) | LOGREA (.23) | LOGREA (.21) | LOGREA (.20) |
| 6 | STCUR (.27) | TSTANX (.26) | TSTANX (.25) | TSTANX (.24) | CONFIG (.22) | CONFIG (.21) |
| 7 | TSTANX (.28) | STCUR (.27) | STCUR (.26) | STCUR (.24) | TSTANX (.22) | HIYEAR (.21) |
| 8 | HIYEAR (.29) | HIYEAR (.27) | CONFIG (.26) | CONFIG (.25) | TRANX (.23) | STCUR (.21) |
| 9 | | RVOCGN (.28) | PREFP (.26) | | HIYEAR (.23) | TSTANX (.22) |
| 10 | | PREFP (.28) | PREFV (.27) | | | |

Table 9

STATISTICS FROM THE REGRESSION OF CUMULATIVE LESSON TIME ON BASELINE VARIABLES FOR SELECTED LESSONS

| Block | Lesson | $R^2$ | Standard Error of Estimation | Mean Cumulative Lesson Time |
|---|---|---|---|---|
| 1 | 1 | .10 | 53 | 96 |
| | 2 | .17 | 67 | 178 |
| | 5 | .28 | 139 | 471 |
| | 13* | .28 | 286 | 1033 |
| 2 | 1 | .24 | 314 | 1133 |
| | 2 | .29 | 345 | 1263 |
| | 3 | .29 | 357 | 1304 |
| | 8 | .29 | 533 | 2149 |
| | 12* | .29 | 559 | 2316 |
| | 13* | .28 | 582 | 2378 |
| 3 | 12* | .26 | 908 | 3921 |
| | 13* | .27 | 931 | 4024 |
| 4 | 12* | .25 | 1120 | 4982 |
| | 13* | .25 | 1149 | 5063 |
| 5 | 5 | .24 | 1241 | 5571 |
| | 12* | .23 | 1312 | 6133 |
| | 13* | .23 | 1336 | 6250 |
| 6 | 5 | .22 | 1424 | 6809 |
| | 13* | .22 | 1468 | 7186 |

*A "12" represents the Chief of Supply Lab; a "13" represents the block review, consisting of the block test and remediation.

Table 10

STATISTICS FROM THE REGRESSION OF COURSE TIME ON INITIAL BLOCK OR LESSON TIMES FOR THE PERFORMANCE MILESTONE MODEL

Block Level

| | $R^2$ | Standard Error of Estimation | Mean Time Remaining | $R^2$ for Remaining Course Time |
|---|---|---|---|---|
| Baseline | .28 | 1888 | 9405 | – |
| Block: 1 | .50 | 1554 | 8185 | .34 |
| 2 | .74 | 1116 | 6490 | .46 |
| 3 | .89 | 731 | 4195 | .42 |
| 4 | .94 | 535 | 2804 | .39 |
| 5 | .98 | 294 | 1187 | .21 |

Lesson Level

| | $R^2$ | Standard Error of Estimation | Mean Time Remaining | $R^2$ for Remaining Course Time |
|---|---|---|---|---|
| Baseline | .22 | 1468 | 7186 | – |
| Block: 1 | .51 | 1152 | 6154 | .35 |
| 2 | .67 | 951 | 4808 | .34 |
| 3 | .84 | 654 | 3162 | .34 |
| 4 | .92 | 481 | 2124 | .28 |
| 5 | .98 | 258 | 936 | .27 |

gain in $R^2$ for predicting course completion time using succeeding cumulative block times is necessary since cumulative course time is the sum of cumulative block time and remaining course time. But the $R^2$ for predicting remaining course time is also much larger than the $R^2$ for the baseline regression.

In any event, the standard error of estimation decreases by about 300 minutes between the baseline regression and those using the performance data on the first block. The standard errors continue to decline at the end of each subsequent block, indicating the rate at which increasingly accurate prediction of course completion time can be made, given information on performance on succeeding blocks.

Table 11 presents the regression statistics for the performance model at the lesson level for each lesson included in the analysis. By the fifth lesson, the $R^2$ value is higher than that found in the baseline milestone models and the standard error of estimation is much less. The average cumulative time to accomplish the fifth lesson was 471 minutes, which constituted about 7% of the entire course.

## Results for BTMI

The results of the parameter estimation for the linear baseline trajectory model are described in two parts: how well the trajectory model fit the data for individuals, and how well the baseline variables could be used to predict the estimated parameters.

Table 11

STATISTICS FROM THE REGRESSION OF REMAINING COURSE TIME ON INITIAL LESSON TIMES FOR THE PERFORMANCE MILESTONE MODEL FOR SELECTED LESSONS

| Block | Lesson | $R^2$ | Standard Error of Estimation | $R^2$ for Remaining Course Time |
|---|---|---|---|---|
| 1 | 1 | .09 | 1564 | .07 |
| | 2 | .21 | 1462 | .18 |
| | 5 | .38 | 1288 | .31 |
| | 13* | .51 | 1152 | .35 |
| 2 | 1 | .52 | 1139 | .35 |
| | 2 | .56 | 1098 | .36 |
| | 3 | .56 | 1093 | .36 |
| | 8 | .65 | 965 | .36 |
| | 12* | .67 | 959 | .35 |
| | 13* | .67 | 951 | .34 |
| 3 | 12* | .83 | 686 | .32 |
| | 13* | .85 | 654 | .34 |
| 4 | 12* | .91 | 502 | .28 |
| | 13* | .92 | 481 | .28 |
| 5 | 5 | .95 | 369 | .32 |
| | 12* | .97 | 284 | .30 |
| | 13* | .98 | 258 | .27 |
| 6 | 5 | .99 | 143 | .19 |

*A "12" corresponds to the Chief of Supply Lab; a "13" corresponds to the block review.

AFHRL-TR-78-7

The first question is addressed by estimating the parameters of the model for each student individually and using the parameter estimates to generate expected cumulative times. The second question is addressed by trying to predict the estimated parameter values using the baseline variables.

Table 12 presents the goodness of fit statistics for cumulative block time and cumulative lesson time at the completion of each block, using the a and b values estimated for each student individually. The goodness of fit statistics used here are the same as those used in the comparison of fits: the mean, stanard deviation, median, mean absolute, and root mean square of the distribution of residuals, where the residual is defined as the observed time minus expected time in minutes. The results indicate that the trajectory model provides a good descriptive fit to the data at both the block and the lesson levels. The values of the mean and median residuals are relatively small across blocks, indicating negligible bias. The values of the mean absolute residual and the root mean square residual are uniformly small, indicating a good fit to the data.

The mean and standard deviation across students of the parameters at the block and lesson level are given in Table 13. The model, as we formulated it, relates cumulative elapsed time, t, to a linear function of achievement, A:

$$t = a + bA$$

In this formulation, the parameter <u>b</u> indicates the amount of time it takes to move through 1% of the course material. It would be expected that the parameter <u>a</u> would be close to zero, and this indeed is the case when it is recalled that the entire course length averages 9405 minutes using block elapsed time and 7185 minutes using lesson elapsed time. The differences in the magnitude of the estimates of <u>b</u> between the block and lesson levels of analysis may be attributed to the differences between the two levels in the estimates of course length.

The statistics on the goodness of fit of the regression of the trajectory model parameters on the baseline parameters are presented in Table 14. Judging from the extremely low values of the square of multiple correlation coefficients, .11 and .13, and the large standard errors of estimation, the fit appears rather poor for the estimate of <u>a</u>. The estimate of <u>b</u> is a bit better, with the square of the multiple correlations of .26 at the block level and .18 at the lesson level. How the adequacy of fit of <u>a</u> and <u>b</u> translates into the adequacy of fit on cumulative times will be discussed later in the comparison of models. The first three variables to enter the estimation of the <u>b</u> coefficient – the total score on the IM/MF reading skills test (READST), sex (SEX), and the score on the experience with conventional instruction scale (EXPCI) – entered first in the baseline milestone model regressions as well. The trajectory model thus appears to provide a good descriptive fit of the data and leads to prediction of progress on the basis of the same baseline variables as are used in the milestone model.

**Results for BTM2**

The BTM2 model, expressed by the equation $t_j^K = a_j + b_j A$, was evaluated only at the block level for reasons that will be discussed below. The parameters in this model were estimated by finding the <u>a</u> and <u>b</u> coefficients for selected values of K (K ranging from .3 to 1.9 in increments of .1). The value of K that minimized the sum of squares of the residuals was taken as the estimate of the K parameter and the associated <u>a</u> and <u>b</u> estimates were assigned the corresponding values.

Table 12

GOODNESS OF FIT STATISTICS FOR THE LINEAR TRAJECTORY MODEL (BTM1) AT THE BLOCK AND LESSON LEVEL

| | Block 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **Block Level** | | | | | | |
| Residual statistics | | | | | | |
| Mean | 13 | –45 | 32 | 2 | 34 | –36 |
| Standard deviation | 221 | 178 | 272 | 234 | 156 | 243 |
| Median | 34 | –42 | –9 | –1 | 24 | –21 |
| Mean absolute | 163 | 128 | 195 | 175 | 120 | 186 |
| Root mean square | 221 | 183 | 274 | 234 | 160 | 245 |
| **Lesson Level** | | | | | | |
| Residual statistics | | | | | | |
| Mean | –10 | 24 | 12 | –12 | –10 | –13 |
| Standard deviation | 134 | 176 | 217 | 174 | 148 | 307 |
| Median | –10 | 21 | 17 | 0 | –15 | –11 |
| Mean absolute | 103 | 134 | 167 | 133 | 117 | 242 |
| Root mean square | 134 | 177 | 217 | 175 | 149 | 307 |

Table 13

MEAN AND STANDARD DEVIATION OF THE ESTIMATED PARAMETERS FOR THE BASELINE TRAJECTORY MODEL 1
(n = 368)

| | Mean | S.D. |
|---|---|---|
| Block Level | | |
| $\hat{a}$ | –121 | 464 |
| $\hat{b}$ | 96 | 23 |
| Lesson Level | | |
| $\hat{a}$ | 50.7 | 297 |
| $\hat{b}$ | 72. | 17 |

AFHRL-TR-78-7

The distribution of students on the value of K that minimizes the residual sum of squares is given in Figure 1. The mode of the distribution is at K = 1.0. Of all the students, 22% had a value of K = 1.0 that minimized the residual sum of squares and 77% had values between .8 and 1.2, inclusive.

Because of the way in which the achievement index was defined, it is not surprising that most of the values of K cluster about 1. The achievement index was based on a standardization of mean lesson times so that each lesson was assigned an achievement value equal to the percentage of course time spent on the lesson. The fact that the optimum K value was approximately 1 for most students indicates that the proportion of time spent on a particular part of the course tends to be stable across students.

As was found in another examination of the trajectory model (Larsen, Markosian, and Suppes, 1977), when K is estimated individually by student, the values of K, a, and b are highly dependent. For example, Table 15 presents the mean value of the estimated b coefficient as a function of K. The range in the mean value of the estimate of the b coefficient is from .7 for K = .5 to 477766 for K = 1.9. This dependence may be explained by the way in which K enters the model as an exponent of time. For a given observed course length, the value of the a and b coefficients would be expected to increase as K increases.

Table 16 presents the summary statistics on the descriptive fit of the nonlinear trajectory model. The low values of the mean and median indicate that the bias in the model is negligible. The values of the mean absolute residuals and root mean square residuals indicate a good fit to the data. Comparison of the statistics in Table 16 with those in Table 12 show that the nonlinear model improves the descriptive fit substantially for the first and last blocks.

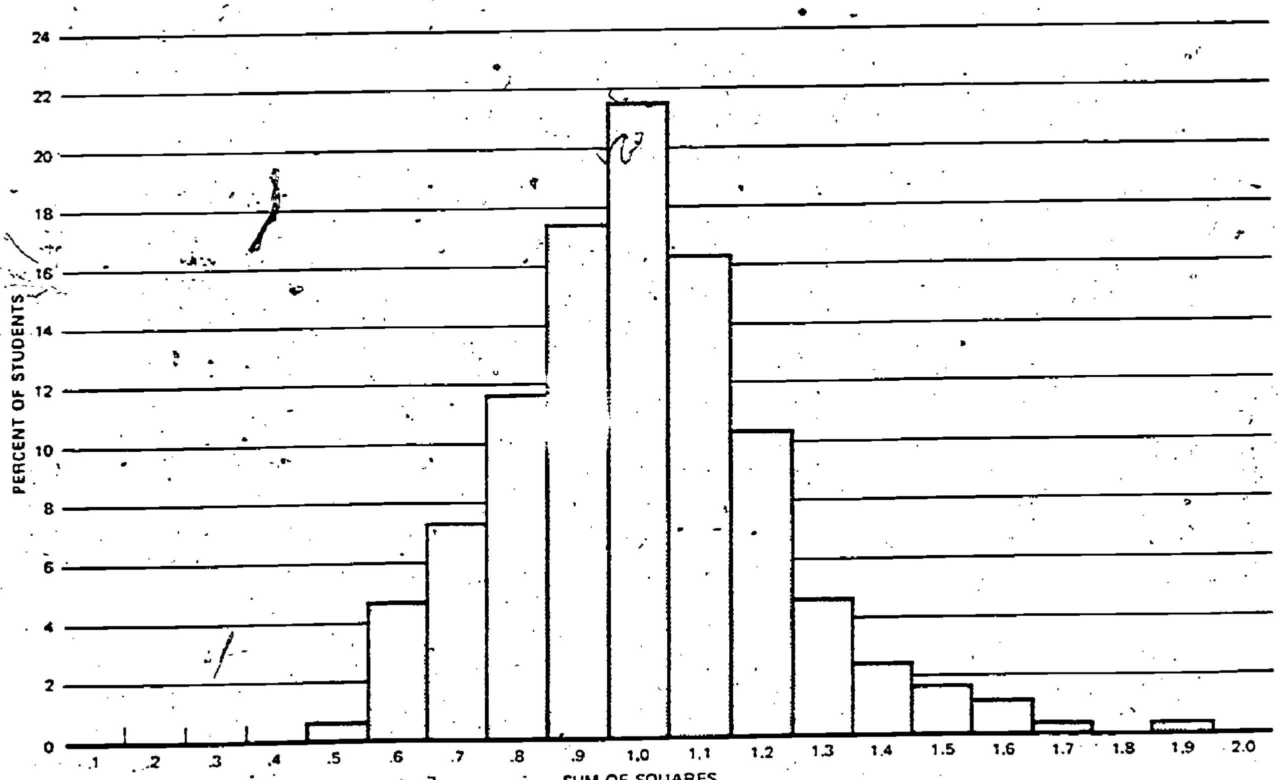


FIGURE III-1 DISTRIBUTION OF STUDENTS ON K THAT MINIMIZES THE RESIDUAL

Table 14

STATISTICS FROM THE REGRESSION OF THE LINEAR TRAJECTORY PARAMETERS ON THE BASELINE VARIABLES

| | Block Level | | | | Lesson Level | | | |
|---|---|---|---|---|---|---|---|---|
| Dependent variable | | a | | b | | a | | b |
| $R^2$ with all baseline variables | | .11 | | .26 | | .13 | | .20 |
| $R^2$ with truncated set | | .10 | | .26 | | .11 | | .18 |
| Standard error of estimation | | 445 | | 20 | | 282 | | 16 |
| Variable entered ($R^2$) | | | | | | | | |
| Step 1 | READST | (.03) | READST | (.11) | RVOCTL | (.07) | READST | (.08) |
| 2 | SEX | (.06) | EXPCI | (.16) | HIYEAR | (.08) | EXPCI | (.12) |
| 3 | MEMNB | (.08) | SEX | (.21) | STCUR | (.09) | SEX | (.16) |
| 4 | LOGREA | (.09) | IESCL | (.22) | EXPCI | (.10) | TRCUR | (.17) |
| 5 | ENTAGE | (.09) | TRCUR | (.23) | SEX | (.11) | LOGREA | (.18) |
| 6 | STCUR | (.10) | READST | (.24) | TRCUR | (.11) | CONFIG | (.18) |
| 7 | RVOCGN | (.10) | HIYEAR | (.24) | | | | |
| 8 | | | ENTAGE | (.25) | | | | |
| 9 | | | STCUR | (.25) | | | | |
| 10 | | | MEMNB | (.25) | | | | |
| 11 | | | TSTANX | (.26) | | | | |

Table 15

STATISTICS ON THE ESTIMATE OF THE COEFFICIENT AS A FUNCTION OF K

| K | Mean | S.D. | N |
|---|---|---|---|
| .5 | .7 | .06 | 2 |
| .6 | 1.9 | .35 | 17 |
| .7 | 5.0 | .89 | 27 |
| .8 | 13.5 | 2.24 | 43 |
| .9 | 35.7 | 7.09 | 64 |
| 1.0 | 94.8 | 22.45 | 79 |
| 1.1 | 257.2 | 66.93 | 60 |
| 1.2 | 721.7 | 210.07 | 38 |
| 1.3 | 1767.9 | 436.88 | 17 |
| 1.4 | 4881.0 | 1104.23 | 9 |
| 1.5 | 10723.9 | 4597.95 | 6 |
| 1.6 | 38168.0 | 17598.96 | 4 |
| 1.7 | 89632.0 | – | 1 |
| 1.8 | 477766.2 | – | 1 |

Table 16

GOODNESS OF FIT STATISTICS FOR THE NONLINEAR TRAJECTORY MODEL (BTM2) AT THE BLOCK LEVEL

| | Block | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Residual Statistics | | | | | | |
| Mean | 20 | –62 | 35 | 12 | 39 | –45 |
| Standard deviation | 74 | 195 | 146 | 152 | 154 | 116 |
| Median | 10 | –29 | 18 | 3 | 31 | –40 |
| Mean absolute | 57 | 139 | 103 | 116 | 118 | 95 |
| Root mean square | 76 | 205 | 150 | 153 | 158 | 124 |

Estimation of K, a, and b using the baseline variables has presented major problems. Because the value of K corresponds to the exponent of the dependent variable, estimation of the three parameters separately would not be fruitful.

One alternative that was explored was to attempt to predict K using the baseline variables and then to estimate a and b as a function of K. As a first step, K was entered as the dependent variable in a stepwise regression with the baseline variables as independent variables.

The results of the regression, however, indicated that the baseline variables were poor predictors of K. With all the variables entered in the regression, the square of the multiple correlation coefficient was only .13 and the standard error of estimation was .2. With such poor prediction of K and the sensitivity of the nonlinear model to the value of K, the nonlinear model does not appear to be useful for predicting student progress. Augmenting the trajectory model with a nonlinear component apparently improves the model from the point of view of describing the data base, but makes the model too sensitive for predictive purposes.

## Results for PTM

Table 17 presents the means and standard deviations of the parameter estimates for the baseline trajectory model at the end of each block. The estimates at the end of block 1 are left blank in the block level analyses because the PTM model requires at least two points of observation. The model, it may be recalled, is:

$$t = a + bA,$$

where t is cumulative time, A is cumulative achievement, and a and b are parameters to be estimated. For the block level analysis, the mean of the estimates of a are consistently negative, but with a very large standard deviation. For the lesson data, the mean of the estimates of a are substantially closer to zero, with much smaller standard deviations. The lower standard deviation of the estimates of a at the lesson level is probably due to the larger number of data points that enter into the estimation. For example, at block 2 there are 2 data points for each student at the block level and 19 data points for each student at the lesson level.

Another point that may be noteworthy about Table 17 is that the standard deviation in the estimates of b decreases substantially between the initial block and block 5. At the block level, the decrease is from 30 minutes to 24 minutes; at the lesson level the decrease is from 25 minutes to 18 minutes. The decline continues, as indicated in Table 13, when the Block 6 data are included in the estimation. This decline is probably due in part to the increase in the number of data points across blocks. It may also indicate that the learning rate is becoming more homogeneous over time. Note, however, that the parameter related to learning rate, b, still has a rather large standard deviation.

## Results for MMM

Table 18 contains the summary statistics for the regression of remaining cumulative course time on both the performance and the baseline variables. Comparison of these data with those presented in Table 10 for the performance milestone model indicates that the baseline variables add very little in predicting course completion time in addition to what is explained by the performance variables. At the block level, the standard error of estimation decreases by about 100 minutes as the result of including the baseline variables at the end of Block 1. In all other cases, the improvement in prediction is negligible.

AFHRL-TR-78-7

Table 17

MEAN AND STANDARD DEVIATION OF THE ESTIMATED PARAMETERS FOR THE PERFORMANCE TRAJECTORY MODEL

(n = 368)

| | Block | | | | Lesson | | | |
|---|---|---|---|---|---|---|---|---|
| | $\hat{a}$ | | $\hat{b}$ | | $\hat{a}$ | | $\hat{b}$ | |
| At the end of Block: | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| 1 | – | – | – | – | –4 | 69 | 75 | 25 |
| 2 | –64 | 386 | 92 | 30 | –1 | 101 | 74 | 22 |
| 3 | –140 | 413 | 96 | 28 | 17 | 159 | 73 | 20 |
| 4 | –133 | 410 | 96 | 26 | 31 | 202 | 72 | 19 |
| 5 | –140 | 428 | 96 | 24 | 45 | 257 | 72 | 18 |

Table 18

STATISTICS ON THE REGRESSION FOR THE MIXED MILESTONE MODEL

| | Truncated Run Number of Base-line Variables Entered | $R^2$ on Truncated Run | Standard Error of Estimation |
|---|---|---|---|
| **Block Level** | | | |
| Block 1 | 6 | .57 | 1449 |
| 2 | 1 | .75 | 1105 |
| 3 | 0 | – | – |
| 4 | 0 | – | – |
| 5 | 0 | – | – |
| **Lesson Level** | | | |
| Block 1 | 4 | .53 | 1137 |
| 2 | 0 | – | – |
| 3 | 0 | – | – |
| 4 | 0 | – | – |
| 5 | 0 | – | – |

Note: "–" indicates no baseline variables entered.

## Comparison of the Baseline Models

The goodness of fit statistics for the three baseline models are arrayed in Table 19 for the block level analysis and in Table 20 for the lesson level analysis. The goodness of fit statistics indicate that all three models fit the data to about the same degree. The mean residuals are uniformly small, at both the block and the lesson levels. The median residuals are consistently negative and consistently exceed 100 minutes for Blocks 3 through 6. Relative to the magnitude of the mean absolute residuals and the root mean square residuals, however, the medians appear relatively small; and the bias in the models does not appear to be large enough to be a factor in model selection.

The mean absolute deviations and the root mean square deviations increase substantially over blocks for all models. These statistics are about five times larger at the end of the sixth block than at the end of the first block. Of course, this is a direct reflection of the relative magnitude of the cumulative time to the end of the course versus the cumulative time to the end of the first block.

## Comparison of the Performance Models

The predictive accuracy statistics for the two performance models are arrayed in Table 21 for the block level and lesson level analyses. The statistics are for the distribution of observed minus predicted course time using baseline data and time-to-criterion data up to the end of each block. There is no entry for Block 1 for the trajectory model at the block level because this model requires at least two observations for purposes of estimation.

The predictive accuracy statistics indicate that the performance milestone model is better at predicting course time than the trajectory model, both at the block and lesson levels. For all blocks, the mean absolute residual and the root mean square residual are much higher for the trajectory model than for the milestone model. In fact, the size of these statistics indicates a lag of about one block between the two models. That is, it takes about one more block for the trajectory model to meet the accuracy of the milestone model.

The performance trajectory model also has a rather large bias in the first few blocks. At the block level, the large positive values of the mean and median residuals at the end of Block 2 indicate a positive bias, which means that the model is predicting too low a value for the remaining course time. At the lesson level, the large negative values for the mean and median residuals at the end of Block 1 and Block 2 indicate a negative bias, where the model is predicting too high a value for the course time. Also, it is interesting that at the lesson level the mean absolute residual and root mean square residual are larger for the performance trajectory model at the end of Block 1 than for the baseline trajectory model.

One explanation for the difference in predictive accuracy between the trajectory and milestone models is that the milestone model tends to rely heavily on the most recently observed cumulative time, whereas the trajectory model tends to rely on all the observed cumulative times. Thus, the lag in the predictive accuracy statistics for the trajectory model may be attributed to the weighting that this model gives to early cumulative times. In any event, the performance milestone model appears to be superior to the performance trajectory model, at both the block and the lesson levels.

AFHRL-TR-78-7

Table 19

GOODNESS OF FIT STATISTICS FOR THE THREE BASELINE MODELS: BLOCK LEVEL ANALYSIS

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **BMM1** | | | | | | |
| Mean | 0 | 0 | 0 | 0 | 0 | 0 |
| Standard deviation | 384 | 734 | 1216 | 1448 | 1717 | 1865 |
| Median | –39 | –80 | –28 | –135 | –95 | –117 |
| Mean absolute | 280 | 556 | 952 | 1146 | 1368 | 1480 |
| Root mean square | 384 | 733 | 1214 | 1446 | 1715 | 1862 |
| **BMM2** | | | | | | |
| Mean | – | – | – | – | – | 0 |
| Standard deviation | 397 | 746 | 1227 | 1454 | 1720 | 1865 |
| Median | –47 | –79 | –114 | –138 | –151 | –117 |
| Mean absolute | 290 | 575 | 963 | 1155 | 1372 | 1480 |
| Root mean square | 396 | 745 | 1225 | 1452 | 1718 | 1862 |
| **BTM** | | | | | | |
| Mean | 13 | –45 | 32 | 2 | 34 | –36 |
| Standard deviation | 393 | 735 | 1217 | 1446 | 1714 | 1865 |
| Median | –16 | –96 | 1 | –136 | –72 | –83 |
| Mean absolute | 283 | 561 | 949 | 1146 | 1365 | 1477 |
| Root mean square | 392 | 735 | 1215 | 1444 | 1712 | 1863 |

Note: "–" indicates between –.5 and +.5, but not identical to zero.

AFHRL-TR-78-7

Table 20

GOODNESS OF FIT STATISTICS FOR THE THREE BASELINE MODELS: LESSON LEVEL ANALYSIS; AT THE END OF EACH BLOCK

| | Block | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| **BMI** | | | | | | |
| Mean | 0 | 0 | 0 | 0 | 0 | 0 |
| Standard deviation | 283 | 574 | 918 | 1136 | 1320 | 1450 |
| Median | –32 | –75 | –108 | –166 | –121 | –131 |
| Mean absolute | 220 | 450 | 735 | 927 | 1069 | 1178 |
| Root mean square | 283 | 573 | 917 | 1135 | 1318 | 1448 |
| **BMM2** | | | | | | |
| Mean | 35 | 63 | 42 | 12 | 7 | 0 |
| Standard deviation | 295 | 589 | 932 | 1143 | 1322 | 1450 |
| Median | 4 | 6 | –83 | –201 | –142 | –131 |
| Mean absolute | 234 | 470 | 751 | 942 | 1071 | 1178 |
| Root mean square | 297 | 592 | 932 | 1142 | 1320 | 1448 |
| **BTM** | | | | | | |
| Mean | –10 | 24 | 12 | –11 | –10 | –13 |
| Standard deviation | 286 | 582 | 928 | 1142 | 1327 | 1461 |
| Median | –37 | –43 | –122 | –239 | –165 | –150 |
| Mean absolute | 222 | 458 | 747 | 937 | 1070 | 1187 |
| Root mean square | 286 | 582 | 927 | 1140 | 1325 | 1459 |

The mean absolute deviations and the root mean square deviations increase substantially over blocks for all models. These statistics are about five times larger at the end of the sixth block than at the end of the first block. Of course, this is a direct reflection of the relative magnitude of the cumulative time to the end of the course versus the cumulative time to the end of the first block.

AFHRL-TR-78-7

Table 21

GOODNESS OF FIT STATISTICS FOR THE TWO PERFORMANCE MODELS: BLOCK LEVEL AND LESSON LEVEL

| Level, Model, and Residual Statistics | Baseline | Block 1 | Block 2 | Block 3 | Block 4 | Block 5 |
|---|---|---|---|---|---|---|
| **Block** | | | | | | |
| Milestone | | | | | | |
| Mean | 0 | 0 | 0 | 0 | 0 | 0 |
| Standard deviation | 1865 | 1552 | 1113 | 728 | 532 | 292 |
| Median | –117 | –105 | –68 | –59 | –20 | –29 |
| Mean absolute | 1480 | 1219 | 875 | 563 | 415 | 230 |
| Root mean square | 1862 | 1550 | 1112 | 727 | 532 | 292 |
| Trajectory | | | | | | |
| Mean | –36 | – | 222 | –75 | –54 | –68 |
| Standard deviation | 1865 | – | 1529 | 969 | 728 | 458 |
| Median | –83 | – | 316 | 29 | 8 | –40 |
| Mean absolute | 1477 | – | 1134 | 710 | 543 | 351 |
| Root mean square | 1863 | – | 1543 | 970 | 729 | 462 |
| **Lesson** | | | | | | |
| Milestone | | | | | | |
| Mean | 0 | 0 | 0 | 0 | 0 | 0 |
| Standard deviation | 1460 | 1146 | 938 | 643 | 470 | 251 |
| Median | –131 | 2 | 16 | –71 | –37 | –21 |
| Mean absolute | 1178 | 915 | 768 | 514 | 375 | 196 |
| Root mean square | 1448 | 1144 | 937 | 642 | 470 | 251 |
| Trajectory | | | | | | |
| Mean | –13 | –280 | –232 | –124 | –66 | –26 |
| Standard deviation | 1461 | 1823 | 1297 | 907 | 691 | 477 |
| Median | –150 | –204 | –260 | –111 | –95 | –32 |
| Mean absolute | 1187 | 1411 | 1042 | 732 | 554 | 378 |
| Root mean square | 1459 | 1842 | 1316 | 914 | 693 | 477 |

Note: Each column contains the goodness of fit statistics for remaining course time using the cumulative performance data through the specified block.

# IV CONCLUSIONS AND RECOMMENDATIONS

## Conclusions Based on the Results of the Evaluation

Based on the results of the literature review summarized in Appendix A, trajectory and milestone models of student progress were formulated and evaluated. The evaluation, using block elapsed times and lesson times to criterion, was conducted separately on 368 students who had completed the Inventory Management Course early in 1977. Four models were evaluated using baseline data only; two models were evaluated using performance data only; and one model was evaluated using both baseline and performance data.

Although the nonlinear trajectory model provided the best descriptive fit to the data, it was found to be a poor predictive model because of difficulties in prediction of the exponential parameter K. None of the other three baseline models appears to be substantially superior to the other two on the basis of the predictive goodness of fit statistics. In particular, the BMM2 model, which corresponds to the model used in the currently implemented SPMS, was comparable in accuracy of prediction to the other two models.

For the performance models, the milestone model was substantially better than the trajectory model in predicting course completion times. The trajectory model appears to need performance data on an additional block to achieve the degree of accuracy in prediction displayed by the milestone model.

In comparing the results for the baseline milestone model with those for the performance milestone model, it is evident that prediction of student progress can be made more accurately from initial performance data than from baseline data. This result is consistent with what was found in several other studies included in the literature review (Wagner et al., 1973; Yeager and Kissel, 1969; Wang, 1968), namely, that the best predictors of student progress are those that are most related to the course content. A measure of actual performance on an initial segment of a course, then, will be a good predictor of student progress if the course is relatively homogeneous in the the types of skills that are necessary for learning the contents.

For the Inventory Management Course, the accuracy of prediction of course completion time can be improved by using initial performance data. For example, when the block elapsed time on the first block is used to predict course completion time, the standard error of estimation is 334 minutes less than the standard error of estimation using the baseline data. Using the first two block elapsed times as predictors, the standard error of estimation is 772 minutes less than that derived using the baseline data (see Table 10). Of course, the amount of time remaining in the course is also decreasing at the end of successive blocks. Nevertheless, the increase in precision of prediction of course completion time is appreciable.

Finally, on the basis of the results for the mixed milestone model, it may be concluded that augmenting the performance data with baseline variables as predictors modestly improves the precision of prediction at the end of the first block. However, at the end of subsequent blocks, the baseline variables do not add substantially to the precision of prediction obtained using performance data alone.

## Other Considerations

As was indicated in Section I, the AIS currently has a student progress management component that was implemented in July 1977. In considering recommendations for Phase III of this project, it is valuable to review some of the initial results found for the currently implemented system, as reported by Dallman and Grau (1977). Comparisons of performance data between students who were in the IM course before the implementation of the SPMS with students who were in the course after the SPMS was implemented indicated:

- The average block elapsed times for the SPMS students were about 10% less than the average block elapsed times for the non-SPMS students.
- There was a definite decrease in block grades and a definite increase in first-attempt failure rates when the SPMS was implemented.

Dallman and Grau's recommendations focus on what course managers, supervisors, and instructors can do to improve the management system. Apparently, the procedures for generating the target completion times, including the prediction equations, were found to be satisfactory.

**Recommendation for Phase III**

Our recommendation for Phase III is to implement a form of the performance milestone model on the AIS to augment the currently implemented baseline model. The currently implemented model provided about as good a predictive fit as the other baseline models that were examined. Therefore, there is no reason to modify the procedure for predicting student performance based on the baseline data alone.

Our evaluation confirms what many researchers have found (Wagner et al., 1973; Wang, 1968; Yeager and Kissel, 1969), namely, that the precision of prediction of student performance increases as a function of the relevance of the predictor variable to the course content. By taking initial performance measures on the course itself, it would be possible to improve prediction of performance on the remaining course material, if the material is reasonably homogeneous throughout the course.

Within the currently specified level of effort, Phase III would consist of the following tasks:

(1) Familiarization with the coding and design of the AIS, with special attention paid to the Adaptive Model and the student progress management system.

(2) Consultation with AFHRL and other personnel at Lowry regarding implementation of the model.

(3) Formulation of specifications for integrating the performance model into the existing AIS, resulting in a detailed integration and design document.

(4) Implementation of the modifications using CAMIL and testing for reliability.

(5) Briefing of course personnel on model output.

(6) Evaluation of the model.

(7) Preparation of the final report.

The performance model as implemented could be utilized in several different ways. At a minimum, updated predictions could be used in the scheduling of the administrative outprocessing activities and in-progress counseling sessions.

The updated predictions could also be used in a modification of the current student management system. For example, rather than students being provided with their targets for the entire course, they could be informed in a stepwise fashion. They would be given their target times on the first block or the first few blocks based on the baseline model. The targets on each subsequent block could then be provided on the completion of the prior block using the performance model predictions.

Under this scheme, provisions would need to be made to preclude increasing target times for students not working up to their capacity. For example, predictions could be set as the lower of those generated from preassessment data and those generated from performance data. Or they could be established as a weighted average of the two predictions.

AFHRL-TR-78-7

An alternative use for the performance model would be in the determination of when progress counseling sessions are necessary. The current SPMS specifies progress counseling sessions in the event that a student falls two days behind his targeted path through the course. This could be replaced by a criterion that is a function of the difference between a student's targeted completion date based on preassessment data, and a student's projected completion date based on his performance. This type of criterion could be more sensitive to lags in performance at the beginning of the course. For example, a progress counseling criterion of a difference of two days between the baseline model and performance model predictions of course completion time would translate into a difference of substantially less than two days between observed and predicted times on the first or second blocks. This approach would be equivalent to setting criteria for differences in learning rates rather than differences in learning time. This approach could lead to earlier detection of students in need of progress counseling and remedial instruction.

# REFERENCES

Anderson, L. W. Time and school learning. Unpublished doctoral dissertation, University of Chicago, 1973.

Anderson, L. W. An empirical investigation of individual differences in time to learn. *Journal of Educational Psychology*, 1976, 68 (2), 226–233.

Arlin, M. Learning rate and learning rate variance under mastery learning conditions. Unpublished doctoral dissertation, University of Chicago, 1973.

Block, J. H. The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished doctoral dissertation, University of Chicago, 1970.

Bloom, B. S. Learning for mastery. *Evaluation Comment*, 1968, 1 (2).

Bloom, B. S. *Human characteristics and school learning*. New York: McGraw-Hill, 1976.

Bond, N. A., Jr. Motivating the student on CAI technical courses. U.S.C. Behavioral Technology Labs, 1971.

Carroll, J. A model of school learning. *Teachers College Record*, 1963, 64, 723–733.

Chant, V. G., & Luenberger, D. G. A mathematical theory of instruction: instructor/learner interaction and instruction pacing. *Journal of Mathematical Psychology*, May 1974, 11 (2), 132–158.

Cogswell, J. F. Computers in student appraisal and educational planning. In J. U. Loughary (Ed.), *Man-machine systems in education*. New York: Harper and Row, 1966, pp. 157–167.

Cohen, J. L., & Fishbein, M. *A field test of the PLATO IV system for company commander behavioral change training*. Report no. NATRAEQUIPCEN-74-C-0095-1. Final Report for period August 1975 – December 1975. Naval Training Equipment Center, Orlando, Florida, July 1976.

Colton, F. V. Effects of giving students data on task completion time in a college media course. *Audio Visual Communication Review*, 1974, 22, 279–294.

Cooley, W. W., & Glaser, R. An information and management system for individually prescribed instruction. In R. C. Atkinson & H. Wilson (Eds.), *Computer-assisted instruction: A book of readings*. New York: Academic Press, 1969, pp. 95–117.

Cooley, W. W., & Lohnes, P. R. *Evaluation research in education: Theory, principles, and practice*. New York: John Wiley & Sons, 1976.

Cronbach, L. J., & Snow, R. E. *Aptitudes and instructional methods*. New York: Irvington Publishers, 1977.

Dallman, B. Personal Communication, 1977.

Dallman, B., & Grau, R. Personal Communication, 1977.

Farr, M. J. Computer-assisted instruction activities in naval research. *Computers and Automation*, January 1973, 10–13.

Gibbons, M. What is individualized instruction? *Interchange*, 1970, 1 (2), 28–52.

AFHRL-TR-78-7

Gordon, J. M., Jr. Individualizing instruction. Unpublished manuscript. Detroit: Wayne State University, 1970.

Hicklin, W. J. A model for mastery learning based on dynamic equilibrium theory. *Journal of Mathematical Psychology*, February 1976, 13 (1), 79–88.

Jamison, D., Suppes, P., & Wells, S. The effectiveness of alternative instructional media: A survey. *Reviews of Educational Research*, 1974, 44, 1–67.

Johnson, K. A., Salop, R. A., & Harding, L. G. The effect of incentives on self-paced instruction. Technical Bulletin STB73-2. Naval Personnel and Training Research Laboratory, 1972.

Katz, A. Personnel reactions to incentives, naval conditions and experiences (PRINCE): Demographic and background information, expectations, attitudes, values and motivations of new recruits. WRR 72-2. Naval Personnel Research and Development Laboratory, Washington Navy Yard, Washington, D. C., September 1971.

Klimoski, R. J., Raben, C. S., Haccoun, R. R., & Gilmore, D. An annotated bibliography on social reinforcement: Evaluative abstracts of research and theory. AFHRL-TR-74-9 (II), AD-A001-119. Air Force Human Resources Laboratory, Technical Training Division, Lowry Air Force Base, Colorado, August 1974.

Larsen, I., Markosian, L. Z., & Suppes, P. Performance models of undergraduate students on computer-assisted instruction in elementary logic. *Instructional Science*, 1977, 1–21.

Malone, T., Suppes, P., Zanotti, M., Macken, E., Kinerva, L. Projecting student trajectories in a computer-assisted instruction curriculum. Unpublished manuscript, Computer Curriculum Corporation, Palo Alto, California, 1977.

Myers, D. L. The effects of predictive solutions on training time and post-training performance for control systems with human operators. Paper prepared for the National Technical Information Service, Springfield, Va., 1969.

Packard, R. B. Models of individualized instruction: The search for a measure. *Educational Technology*, 1972, 12, 11–13.

Pritchard, R. D., Von Bergen, C., Jr., & DeLeo, P. J. *Incentive motivation techniques evaluation in Air Force technical training*. AFHRL-TR-74-24, AD-A005 302. Air Force Human Resources Laboratory, Technical Training Division, Lowry Air Force Base, Colorado, November 1974.

Raben, C. S., Wood, M. T., Klimoski, R. J., & Hakel, M. D. *Social reinforcement: A review of the literature*. AFHRL-TR-74-9 (I), AD-A001 118. Air Force Human Resources Laboratory, Technical Training Division, Lowry Air Force Base, Colorado, August 1974.

Rosenbaum, P. S., On mediation and control in language learning tasks. IBM Research, TC2580, Yorktown Heights, New York, 1969.

Sass, R. E. A computer-based instructional management program for classroom use. Paper no. 13. Learning Research and Development Center, University of Pittsburgh, 1971.

Schwen, T. M. Learner analysis: Some process and content concerns. *Audio Visual Communication Review*, 1973, 21 (1), 44–72.

Self, J. A. Student models in computer aided instruction. *International Journal of Man-Machine Studies*, 1974, 6, 261–276.

Shrable, K., & Sassenrath, J. M. Effects of achievement, motivation and test anxiety on performance in programmed instruction. *American Educational Research Journal*, 1970, 7, 209–220.

Sloan, N. E. Human resources in the guidance programs: Students in helping roles. Paper prepared for the National Technical Information Service, Springfield, Va., 1970.

Spielberger, C. D., O'Neil, H. F., & Hansen, D. N. Anxiety, drive theory, and computer-assisted learning. In B. A. Maher (Ed.), *Progress in experimental personality research: VI*. New York: Academic Press, 1972.

Suppes, P. Modern learning theory and the elementary school curriculum. *American Educational Research Journal*, 1964, 4, 79–93.

Suppes, P., Fletcher, J. D., & Zannoti, M. Models of individual trajectories in computer-assisted instruction for deaf students. Tech. Report No. 214. Institute for Mathematical Studies in the Social Sciences, Stanford, California, 1973.

Suppes, P., Fletcher, J. D., & Zanotti, M. Performance models of American Indian students on computer-assisted instruction in elementary mathematics. *Instructional Science*, 1975, 4, 303–313.

Suppes, P., Fletcher, J. D., & Zanotti, M. Models of individual trajectories in computer-assisted instruction for deaf students. *Journal of Educational Psychology*, 1976, 68(2), 117–127.

Suppes, P., Macken, E., & Zanotti, M. The role of global psychological models in instructional technology. To appear in R. Glaser (Ed.), *Advances in instructional psychology*. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Teel, D. A. A comparison of methods utilizing the contract approach in teaching beginning electricity-electronics fundamentals to college students. Unpublished doctoral dissertation, Texas A&M University, 1967.

Tobias, S., & Abramson, T. Interaction among anxiety, stress, response mode and familiarity of subject matter on achievement from programmed instruction. *Journal of Educational Pyschology*, 1971, 62, 357–364.

Wagner, H., Behringer, R. D., & Pattie, C. L. Individualized course completion time predictions: Development of instruments and techniques. Technical Report 73-25. Human Resources Research Organization, Alexandria, Va., 1973.

Wang, M. C. An investigation of selected procedures for measuring and predicting rate of learning in classrooms operating under a program of individualized instruction. Unpublished doctoral dissertation, University of Pittsburgh, 1968.

Wang, M. C. Maximizing the effective use of school time by teachers and students. Paper no. 23. Learning Research and Development Center, University of Pittsburgh, 1976.

Wang, M. C. Use of the canonical correlation analysis in an investigation of pupils' rate of learning in school. *Journal of Educational Research*, September 1970, 64(1), 35–45.

Wang, M., & Lindvall, C. M. An exploratory investigation of the Carroll learning model and the Bloom strategy for mastery learning. Working Paper No. 61. Learning Research and Development Center, University of Pittsburgh, 1970.

AFHRL-TR-78-7

Washburne, C. W. Educational measurement as a key to individualizing instruction and promotion. *Journal of Educational Research*, 1922, 5, 195–206.

Wood, M. T., Hakel, M. D., Del Gaizo, E. R., & Klimoski, R. J. *Identification and analysis of social incentives in Air Force technical training.* AFHRL-TR-75-10, AD-A017 871. Air Force Human Resources Laboratory, Technical Training Division, Lowry Air Force Base, Colorado, October 1975.

Yeager, J. L., & Kissel, M. An investigation of the relationship between selected student characteristics and time required to achieve unit master. Working Paper No. 46. Learning Research and Development Center, University of Pittsburgh, 1969.

Yeager, J. L., & Lindvall C. M. Evaluating an instructional innovation through the observation of pupil activities. *The High School Journal*, 1968, 51 (6), 248–253.

44

AFHRL-TR-78-7

# Appendix A
# LITERATURE REVIEW

The literature review covered two areas relevant to the management of student progress: (1) mathematical and conceptual models of student progress, and (2) incentives and intervention strategies to help improve student motivation or study methods. The emphasis was on the review of models of student progress since the primary thrust of the project was on the formulation, evaluation, and ultimate implementation of such a model. The review of intervention strategies was intended to provide insight into how the model of student progress could be used in the context of the Advanced Instructional System.

## Description of the Literature Search

The literature search used both computational and manual methods to identify references that were potentially relevant to the project. Computerized searches were made through the following files:

- Smithsonian Science Information Exchange (SSIE) – current projects by definition.
- Resources in Education/Current Index to Journals in Education (ERIC) – 1966 through May 1977.
- Psychological Abstracts – 1967 through April 1977.
- Computer and Control Abstracts (INSPEC) – 1970 through May 1977.
- National Technical Information Service (NTIS) – 1964 through Issue 11, 1977.
- Quarterly Bibliography of Computers and Data Processing – 1968 through 1976.
- Technical Abstract Bulletin – 1969 through May 1977.

The search strategy was designed to match the concepts of computer-assisted and computer-managed instruction with the concepts of adaptive control, optimization, mathematical models, and related terms. An attempt was made to restrict citations for applications of computers to instruction in the military and technical domains, but this proved to be too limiting. Therefore, the domain of instruction was not used as a delimiter in the computer search in order to include as many potentially relevant sources as possible.

The manual search included examination of recent issues of selected journals, review of bibliographies in references already identified, identification of references from conversations with experts in the field, and requests made to librarians supervising specialized collections.

The titles and available abstracts of all the references produced from both the computerized and manual searches were scanned for relevance. From thousands of references, about 400 remained after the initial scan. About 40 of these were then selected as the most relevant for this particular project. These 40, plus the references listed in the Request for Proposal, were then reviewed in detail in preparation for the literature review that follows. The review is organized into four sections: perspective on individualized instruction, models of student progress, incentives and intervention strategies, and conclusion.

## Perspective on Individualized Instruction

The Advanced Instructional System at Lowry AFB is based on a strategy of individualization of instruction. Cooley and Glaser (1969) define individualized education as "adapting instructional practices to individual requirements." In the case of the Advanced Instructional System, individualization is achieved by the self-pacing of instruction and by providing alternative modes of instruction on selected lessons.

AFHRL-TR-78-7

Gibbons (1970) traces the development of individualized instruction to before the turn of the century, where it originated primarily as a reaction to the conventional age-graded educational system. Sass (1971) cites the work of Washburne (1922) as an early attempt at individualization of instruction. Washburne's idea was to reverse the constraints on time and achievement in an educational program. Rather than having fixed blocks of time for instruction and allowing achievement to vary, Washburne advocated that an educational program be designed so that all students attain a specified level of achievement. In more recent years, Carroll (1963) and Bloom (1968, 1976) have extended and refined this approach in their work on mastery learning.

At present, individualization of instruction is one of the most dominant concepts in educational innovation and reform. This emphasis is due in part to the increasing application of computer technology to education, which has made it possible to process the large amount of data necessary to truly individualize instruction and provide immediate feedback and guidance to the student (Suppes, 1964; Cooley and Glaser, 1969).

As the concept of self-paced instruction has been increasingly accepted, much research has gone into examining rates of progress. Some of this work has been directed toward determining what factors are related to rate of progress in a curriculum (Wang, 1968, 1970; Yeager and Kissel, 1969; Wang and Lindvall, 1970). Related work has been concerned with modeling and predicting rate of progress (Wagner, Behringer, and Pattie, 1973; Suppes, Fletcher, and Zanotti, 1975, 1976; Malone et al., 1977). Other work has been concerned with alternative strategies for optimizing rate of progress (Anderson, 1976; Wang, 1976).

In the next section, we turn to describing the models of student progress that have been developed and summarizing their descriptive and predictive adequacy. We then briefly review some of the most relevant literature on incentives and intervention strategies.

## Review of Models of Student Progress

The references regarding models of student progress may generally be categorized by their degree of specificity. A number of studies have taken a conceptual approach to the development of models of student progress, the object being the development of a framework for a theory of instruction. In other cases, the authors specify axioms or assumptions regarding the processing of information and derive specific functional forms for the relationship between achievement and elapsed time. In the former category are included work by Carroll (1963), Bloom (1976), and Cooley and Lohnes (1976). In the latter category are included work by Suppes and his associates (1975, 1976), Chant and Luenberger (1974), and Hicklin (1976).

A distinction that has been made by Suppes, Macken, and Zanotti (1977) is that between studies at the microscopic level, concerning learning of specific types of material under specific types of reinforcement schedules, and studies at the global level, where the focus shifts from protocols of responses on specific trials to mean performance over substantial periods of time. The psychological literature is replete with studies at the microscopic level of detail. However, these studies are of little use in the endeavor to develop a model of student progress appropriate to the AIS. Therefore, only the models at the global level that would appear to have some use in the current project are included in the literature review.

The three mathematical models are first described; this is followed by a description of several conceptual models; finally, the results of prior evaluations of mathematical models are described.

## Mathematical Models of Student Progress

Each of the three global models of student progress are based on assumptions regarding how students process information. The three models are: the trajectory model (Suppes et al., 1976), the dynamic equilibrium model (Hicklin, 1976), and the generalized Thurstone model (Chant and Luenberger, 1974).

*The Trajectory Model*

Suppes, Fletcher, and Zanotti (1976) begin with five assumptions regarding processing of information. Let $y(t)$ = the position of the student in the course and $\dot{y}(t) = \frac{dy(t)}{dt}$ = the student's rate of progress through through the course; $A(t)$ = cumulative amount of information introduced in the course up to time and $\dot{A}(t)$ = the rate of introduction of information; and $s(t)$ = the student's rate of processing or sampling information. The five axioms may then be expressed as follows:

Axiom 1: $s(t) = k_1 \dot{A}(t)/A(t)$ for some constant $k_1$.

Axiom 2: Upon introduction of a new piece of information at time t, for a small interval of time h, $s(t+h) = s(t) - [s(t) - s(\infty)]\, s(t)$.

Axiom 3: The probability that a new piece of information is introduced for a given student at time t is independent of t and the previous introduction of information.

Axiom 4: $y(t) = k_2 A(t)$ for some constant $k_2$.

Axiom 5: $\dot{y}(t) = k_3 \dot{A}(t)$ for some constant $k_3$.

In their discussion of the axioms, Suppes et al. state that they are least satisfied with Axiom 2 because of "the absence of a more fundamental qualitative characterization of the rate assumption expressed in this axiom." They felt that the other four axioms have a "natural intuitive content that does not require explicit discussion."

The basic equation for the trajectory model,

$$y(t) = bt^k + c$$

was derived from the five axioms. It was stressed that this relationship between course position and elapsed time was stochastic rather than deterministic; that is, it represented what would occur on the average for a given student. The parameters b, c, and k were meant to be estimated separately by student rather than as a function of group data.

*The Dynamic Equilibrium Model*

Hicklin's (1976) dynamic equilibrium models are based on the assumption that at any particular time t there are N units of material. Let

$N_1(t)$ = the amount of material yet to be assimilated

$N_2(t)$ = the status of the individual

$N_3(t)$ = the amount of material in the lost category.

Then $N = N_1(t) + N_2(t) + N_3(t)$.

His basic assumption is that during a time interval $\Delta t$, $N_2(t)$ will increase in proportion to the amount of material to be assimilated, $N_1(t)$, and will decrease in proportion to the student's current status, $N_2(t)$. Under these assumptions and the initial condition that $N_2(0) = 0$, the "basic differential equation of dynamic equilibrium theory" can be derived:

$$\frac{dN_2(t)}{dt} = k_1 N_1(t) - k_2 N_2(t) \text{ for some constants } k_1 \text{ and } k_2.$$

Under alternative assumptions regarding the values of $k_1$ and $k_2$ and what happens to lost material, Hicklin derives three models:

Case I: Under the assumption that $k_2 = 0$, $N_3(0) = 0$, and $N_2(0) = 0$,

$$N_2(t) = N(1 - e^{-k_2 t}).$$

Case II: Under the assumption that lost material reverts to the unassimilated state, $k_2 \neq 0$, and $N_2(0) = 0$,

$$N_2(t) = N\left(\frac{k_1}{k_1 + k_2}\right)\left(1 - e^{-(k_1 + k_2)t}\right)$$

Case III: Under the assumption that $k_2 \neq 0$ and $N_2(0) = 0$,

$$N_2(t) = N\left(\frac{k_2}{k_1 - k_2}\right)\left(e^{-k_2 t} - e^{-k_1 t}\right)$$

*Generalized Thurstone Model*

The model of Instruction/Learner Interaction proposed by Chant and Luenberger (1974) was a generalization of Thurstone's model. Thurstone's model (1930) was based on assumptions regarding th relationship of the state of the learner, $p(t)$, and the number of potential successful acts, $s(t)$, and fail acts, $e(t)$, in the learner's repertoire at time $t$. The function $p(t)$ represents the fraction of total learn and by definition:

$$p(t) = \frac{s(t)}{s(t) + e(t)}$$

Assuming that

$$\frac{ds(t)}{dt} = kp(t),$$

$$de(t) = -k[1 - p(t)], \text{ and}$$

$$m = s(t)\, e(t), \text{ where } m \text{ is a constant,}$$

AFHRL-TR-78-7

the basic differential equation of Thurstone's model can be derived:

$$\frac{dp(t)}{dt} = \frac{2k}{m}\left(p(t)[1 - p(t)]\right)^{3/2}.$$

The function $p(t)$ is asymptotic to $p = 0$ as t decreases and to $p = 1$ as t increases. It is symmetric about $p = 1/2$.

Chant and Luenberger generalize the Thurstone model by proposing the following differential equation to specify the state of the learner:

$$\frac{dp(t)}{dt} = u(t)\, ag[p(t)]\ ,$$

where the function g is assumed to be continuous and to approach zero as its argument approaches zero or one. The function $u(t)$ is called the instructional input variable and is assumed to represent the effect of instruction. The function $g(p)$ is called the characteristic learning function and represents characteristics of the learner and the material to be learned. The variable a in the formula is constant for each individual and is intended to represent the student's aptitude.

*Conceptual Models of Student Progress*

In contrast to the detailed mathematical models described in the previous section, the conceptual models provide schemes for examining the variables relevant to school learning. As such, they incorporate more factors than do the mathematical models, but they lack the specificity of the mathematical models. The relevance of the conceptual models to the current study is in providing perspective regarding the role of "time to learn" in past research. Wagner et al. (1973) attribute the use of "time to learn" as a critical variable in modern educational and training research to a model of school learning developed by John Carroll (1963). According to Carroll's model, the degree of learning a given task is a function of the amount of time spent on learning the task and the amount of time needed to learn the task. Thus Carroll's model embodies Washburne's idea of regarding time as a variable rather than as a given quantity. "Time spent," in Carroll's scheme, depends on opportunity, perseverance, and aptitude; "time needed" depends on aptitude, ability of the student to understand instruction, and quality of instruction.

"Opportunity" and "quality of instruction" are attributes of the educational environment. The former is measured by the amount of time allowed for learning; the latter is defined with respect to the efficiency of instruction and is assessed by the degree to which the amount of time needed to learn is minimized.

"Perseverance," "aptitude," and "ability to understand instruction" are attributes of the individual student. "Perseverance" is related to a student's willingness to spend the time necessary for learning the task; "aptitude," in Carroll's scheme, is defined as "the amount of time needed to learn the task under optimal instructional conditions"; and "ability to understand instruction" was considered to be a factor dependent on general intelligence and verbal ability.

The importance of Carroll's model was in regarding time as a major variable in predicting the degree of learning. Thus aptitude was defined with regard to the time necessary to master a task rather than the level of mastery within a given time.

Carroll's model has served as a basis for paradigms developed by Cooley and Lohnes (1976) and by Bloom (1976). The Cooley and Lohnes model was intended to provide a theoretical framework for evaluative inquiry. As such, it was oriented toward assessment of group rather than individual processes and outcomes. Their model retained the "opportunity" component of the Carroll model, but did not retain the emphasis on "time to learn." For example, "time spent" and "time needed" were not retained as intervening factors.

Bloom's (1976) extension of Carroll's scheme has direct relevance to the AIS at Lowry. The variables in his paradigm consist of three major components: (1) cognitive entry behaviors, determining "the extent to which the student has already learned the basic prerequisites of the learning to be accomplished"; (2) affective entry characteristics, determining "the extent to which the student is (or can be) motivated to engage in the learning process"; and (3) the quality of instruction, which indicates "the extent to which the instruction to be given is appropriate to the learner." All these factors interact on the task to be learned to determine the nature of the learning outcomes: the level and type of achievement, the rate of learning, and affective outcomes.

Bloom's primary thesis with regard to his model is that both student characteristics and quality of instruction can be modified to achieve a higher level of learning for individuals and groups. Quality of instruction can be evaluated by the qualities of cues, participation, and feedback in instruction. Bloom emphasizes the use of feedback and corrective procedures as one way of ensuring a high quality of instruction. Furthermore, Bloom provides evidence that "gives support to a strong inference that quality of instruction has an effect on the learning processes of students as well as on their learning outcomes" (p. 135).

Specifically with respect to learning rate, Bloom states that "when students are provided with the time and help they need to learn and when this produces positive entry characteristics (cognitive and affective), students not only become better able to learn, they also become able to learn with less and less time" (p. 191). Bloom cites results from a number of studies of mastery learning (Block, 1970; Arlin, 1973; Anderson, 1973) as evidence for his claim.

He notes that a student's learning characteristics can be altered positively or negatively at practically any point in a student's history, but that the potential for positive change is highest on the learning tasks that are early in a series. Thus, Bloom's model includes an interaction component between a student's learning characteristics and the quality of instruction. A high quality of instruction, meaning the use of feedback and corrective procedures, early in a sequence of learning tasks can, according to Bloom, improve student efficiency in subsequent tasks and can reduce the variation in learning rates as well.

## Past Evaluations of Mathematical Models of Student Progress

Of the three mathematical models presented earlier, neither Hicklin nor Chant and Luenberger present evaluations of their models with regard to descriptive or predictive adequacy. On the other hand, Suppes and his associates have conducted a number of studies of the trajectory model. (Suppes, Fletcher, and Zanotti, 1973, 1975, 1976; Larsen, Markosian, and Suppes, 1977; and Malone et al., 1977).

In the three references by Suppes, Fletcher, and Zanotti and the reference by Larsen, Markosian, and Suppes, the goodness of fit of the trajectory model is assessed using data collected from a variety of student populations: 297 deaf students on a CAI mathematics curriculum (Suppes et al., 1973, 1976); 69 American Indian children attending a Bureau of Indian Affairs school and participating in a CAI mathematics program (Suppes et al., 1975); and 42 Stanford undergraduate students enrolled in CAI in elementary logic (Larsen et al., 1977).

On all three sets of data, the authors feel the model gives an adequate fit to the data. The most important parameter in fitting the data is the power factor k. Under the assumption that the power factor k is constant across students, the goodness of fit is about the same for a rather broad range of values of k. The fit tends to be substantially improved by allowing k to vary across students.

In the first two studies cited, the distribution of k across students is flat or even U-shaped, showing a great deal of variation across students. In the last study, on the other hand, the distribution is concentrated in a short interval.

For k fixed, the correlation between the other two parameters, b and c, is very low. When all three parameters are estimated, however, high correlations among the three are found. This last result is not surprising since changing the value of k has the effect of changing the scale as well as the shape of the trajectory.

Some results on using the trajectory model for prediction of performance are contained in the article by Larsen et al. (1977) and further results are contained in an unpublished manuscript by Malone et al. (1977). Larsen et al. (1977) examine two forms of the trajectory model to predict course completion time using performance times on initial lessons. One form requires that all three parameters be estimated; the other form assumes the power factor k to be given and fixed across students and only requires individual estimation of the remaining two parameters. For the initial third of the course that consisted of a total of 30 lessons, the model assuming a fixed k provided substantially better course completion time predictions than the model that required k to be estimated for each individual student. For the remainder of the course both models performed about as well and provided what the authors regarded as good predictions.

In their unpublished manuscript, Malone et al. examine the ability of ten alternative trajectory models to predict final grade placement, given time on a CAI drill and practice program in reading and mathematics. The ten models differ with respect to (1) assumptions regarding the power factor, k; (2) assumptions regarding use of performance information consisting of grade placement at interim points in the course; (3) assumptions regarding use of initial grade placement; and (4) assumptions regarding how to estimate the learning rate parameter, b. Data from approximately 3,000 students in third through sixth grades in the Fort Worth Independent School District were used in the analysis. The standard error of the difference between observed and predicted grade placement is used as the goodness of prediction measure. The authors find that the simplest two of the ten models gives the best predictions. These are models that assume that gain in grade placement is linear in time (or the square root of time) on the system, rising from the last observed grade placement at a rate that is estimated from the population average. The author's explanation of the result that simpler models give better predictions is that the prediction was for a point outside the range of observations.

> The more parameters that are available to fit the observations, the more sensitive the curve is to small random fluctuations in the data, and therefore the more radically it can be wrong outside the range of the data. (Malone et al., 1977)

*Other Studies of Student Progress*

Other studies of student progress have been based on correlational and regression analyses to identify factors explaining student rate of learning or to generate prediction equations. A regression analysis may be considered to be a mathematical model in the sense of hypothesizing a linear relationship between the rate of learning or the course completion time and certain independent variables. Such an approach corresponds to the milestone model specified in the proposal and evaluated in the body of this report.

A number of studies conducted by the Learning Research and Development Center (LRDC) at the University of Pittsburgh (Wang, 1968; Wang, 1970; Wang and Lindvall, 1970) have examined the relationship between rate of learning and such variables as pupil aptitude and achievement. The studies analyzed several sets of data collected on elementary school students who were participating in the Individually Prescribed Instruction Project (IPI) conducted by LRDC. Four alternative rates of progress were formulated. The rates expressed progress in terms of point gain on tests, number of pages worked, and number of skills learned per unit of time. Independent variables included measures of aptitude, academic achievement, and prior classroom performance.

Correlations between the rate measures and independent variables were generally very small, the largest correlations being in the range between .2 and .4. The results of multiple regression analyses and canonical correlation analyses indicated that there was some relationship between rates and the independent variables. For example, Wang and Lindvall got multiple correlation coefficients in the range between .34 and .64. However, the regression coefficients were not consistent across data sets. Wang (1970) concluded from the inconsistency in the results that "rate of learning is specific to a given task and is not a general factor characterizing student performance in all learning situations."

Another study conducted at the LRDC and reported by Yeager and Kissel (1969) confirms the importance of using variables related to the task in attempting to predict completion times. Data on between 63 and 69 elementary-school students in eight different units of a mathematics program were used. Days needed to master the unit was the dependent variable and five independent variables were selected based on hypotheses concerning the process by which a teacher might develop a student's prescription. The study found that between 52% and 71% of the variance in completion times on given learning units could be accounted for by the five variables selected for study. A student's unit pretest score and the number of skills he would have to master in a given unit were the two best predictors of time to master a given unit. Age was a consistent, although less strong, predictor. IQ and number of units previously mastered were found to be relatively poor predictors. Thus, variables that were the most closely related to learning the required task were the best predictors of completion times.

A study of Wagner, Behringer, and Pattie (1973) on individualized course completion time predictions was very similar to the present study. The objective of the study was to accurately predict each student's course completion date prior to graduation, for a U.S. Army Stock Control and Accounting Specialist course that was being converted to an individualized curriculum at the time of the study.

Their literature review led to two conclusions that they used in their approach to the problem: (1) measures of aptitude directly relevant to the course are better predictors of completion time than general aptitude measures: and (2) the best prediction equation of course completion time would be linear in the independent variables.

In the first phase of the study, available predictor variables consisting of scores on the Armed Forces Qualification Test (AFQT) and the Army Classification Battery (ACB) of tests were correlated with time to completion and performance scores on sections of the conventional course that had been self-paced. With between 61 and 77 students included in four separate analyses, the highest correlation with time to criterion was –.54 achieved by the Arithmetic Reasoning test in the ACB.

During the second and third phases of the study, a test battery was developed that measured skills and knowledge relevant to the specific course. Prediction equations were developed using a stepwise regression analysis with instructional time and total course time as dependent variables. One of the major findings was that by grouping students according to mode of instruction, either audiovisual (AV) or programmed instruction text (PI), a substantial improvement in prediction was achieved. The multiple correlations corresponding to the final prediction equations using only baseline variables were .65 and .74 for the AV and PI students, respectively, with 52 and 81 students, respectively. When within-course performance times were included in the prediction equations, the multiple regressions increased to about .85 for both groups. The baseline variables that entered both sets of equations were the score on the AFQT and scores on several course-specific tests.

**Review of Incentives and Intervention Strategies**

A significant aspect of Air Force technical training is the overall motivation on the part of trainees to pursue their course of study in an efficient and dedicated manner. Many trainees will be adequately motivated by a desire to learn; the old maxim, "learning is its own reward," appears to have a real-world basis. For these trainees, improving the quality of instructional material, instructors, and educational procedures is the key to minimizing their course completion times. For trainees without some minimum of motivation, the training course will take longer and be less effective, even if the instructional material and instructors are of high quality. Therefore, improving motivation is an important aspect of producing more efficient training.

The literature on this general topic is very large, much too large to permit a comprehensive review. Fortunately, however, the Air Force sponsored several recent research projects focused on the specific problems of motivation related to the Lowry AIS courses (Pritchard, Von Bergen, and DeLeo, 1974; Klimoski, Raben, Haccoun, and Gilmore, 1974; Raben, Wood, Klimoski and Hakel, 1974, Wood, Hakel, Del Gaizo, and Klimoski, 1975). These projects include comprehensive reviews of the literature, analysis of attempts to use incentives to enhance motivation of trainees, and general conceptual analyses of motivation in technical training and education. Since numerous studies have found that the best approach to predicting learning rates and to conducting aptitude-treatment analysis is one tailored to the particular instructional setting (Cronbach and Snow, 1977; Wagner, Behringer, and Pattie, 1973; Packard, 1972; Wang, 1968; Yeager and Kissel, 1969), it seems reasonable that prediction of responses to incentives needs to be considered in a specific setting as well, and that the Air Force reports deserve special emphasis.

One of the Air Force reports is simply an annotated biliography of 234 references (Klimoski et al., 1974). The companion report is an analytical review of the literature (Raben et al., 1974), the major conclusion of which is that social reinforcement is related to a large number of "moderating" variables in an extremely complicated manner. This makes it impossible, at the present time, to predict the effects of social reinforcement in a training situation. However, it is worthwhile to experiment with social reinforcement because of its relatively low cost and its effectiveness with at least some individuals.

One empirical study (Pritchard et al., 1974) was an experiment using three different incentive motivation systems: high-feasibility (i.e., cheap and easy to implement) incentives based on performance (magnitude of block scores), high-feasibility incentives based on effort (behavior in the course), and high- and low-feasibility incentives based on effort. Pritchard et al. made a fairly comprehensive review of the literature and interviewed trainees, instructional staff, and administrators in two technical courses at Chanute AFB to identify relevant incentives, implemented the three incentive systems, and analyzed the effects of incentives on trainee performance. The major conclusion was that only the incentive system including low-feasibility (i.e., expensive or difficult to arrange) incentives was cost-effective. The study suggests the following guidelines:

(1) Incentives must be fairly powerful: "Every attempt should be made to use incentives such as choice of assignment, promotion, and extra leave" (p. 214).

(2) Incentives are not cost-effective in courses where students are already performing near capacity.

(3) Self-paced courses are most appropriate for incentive techniques.

(4) Frequent reinforcement should be scheduled: opportunities to earn points should occur at least once a day.

(5) Both authority figures with whom the student comes in daily contact and his peers should provide positive social reinforcement for high performance.

Johnson, Salop, and Harding (1972) report two analogous studies of student responses to incentives and disincentives in an aviation mechanical fundamentals course utilizing CMI materials. They found an average time saving of 11% over controls when a low-feasibility incentive — choice of Service Rating course after graduation — was used.

Prediction equations developed from Navy Basic Test Battery scores and times from previous classes were used to predict completion times for each of the experimental students. They were told that the order in which they could choose from among available Service Rating courses would depend on the ratio of their predicted completion times to their actual completion times. They were also given the disincentive of a Saturday morning study session if they lagged too far behind their predicted progress rate, but this was never applied because no student fell very far behind his predicted rate.

Both the experimental group and the control group were given the incentive of afternoons and evenings off for whatever days remained between the day they completed the course and the day the course officially ended. Leaves away from the base were not granted. The time allotted to the course was the same as that allotted to the same course taught by traditional methods; the aim of the incentive program was evidently to minimize the number of CMI students who took longer than the allotted course time, because no provision was made for assigning students to their initial duty station immediately if they finished the course early.

The incentive students (experimentals) did not differ significantly from controls in scores on final tests when scores on the Arithmetic Reasoning Test (their best single predictor) were used as covariates to reduce within-group variability. They also did not differ markedly in attitude toward the course and toward the quality of their work. They did differ significantly in course completion time, averaging 11% less time than the control group. Choice of (or at least having some control over) one's training specialization and location of duty assignment thus seems to be a very effective incentive.

Johnson et al. do not discuss how well their predictive equations fit actual student times in past courses. They also do not discuss the circumstances and motivation of students taking the Navy Basic Test Battery. Since it would be advantageous to a student to do poorly on the aptitude tests, given the incentive criterion used in the course, thought should be given to motivating every student to do his best on the preassessment tests as well as in the course. If this is not possible, perhaps preinduction academic performance should be used to weight preassessment scores, or a progress criterion weighted to some degree by all past students' rates should be used.

The last of the highly relevant Air Force reports (Wood et al., 1975) concerned the identification of incentives and analysis of them in terms of feasibility, attractiveness, and other characteristics. A list of incentives was developed through interviews, literature review, and group meetings. This list was evaluated and refined through survey procedures with trainees and instructors.

The more attractive incentives involved an effect directly on the trainee, for example, some choice in duty assignment or immediate promotion upon graduation. These tended to be quite costly or low in administrative feasibility, but their potential strength is supported by Katz's (1971) finding that many Navy recruits were highly motivated by upward social mobility when they signed up, and by Johnson, Salop, and Harding's finding that choice of specialization course and choice of initial duty assignment (even a choice between East Coast and West Coast) were the most valued of the available incentives.

Wood et al. found that Black trainees were more likely to prefer recognition-type incentives and White trainees were more likely to prefer control- and future-career-oriented incentives. They conclude with a proposal to experiment with four incentive systems:

(1) Incentives administered by the instructor based on trainee performance.

(2) Incentives administered by the instructor and the class as a group based on individual performance.

(3) Incentives administered by the instructor to the individual based on the performance of the class as a group.

(4) Incentives administered by the instructor and the class based on performance of the class.

These four systems vary along the need-related or "dynamic" dimension as defined by Bond (1971), and do not provide for comparisons or study of interaction of task-related ("intrinsic") incentives and external motivators (rewards). Thus this experiment does not permit evaluation of rewards that combine need-related incentives with external motivation, e.g., promotion in rank for honor students. Nor does it address Bond's evidence that effective external motivators reduce task-related, intrinsic satisfaction. Lowry course managers concerned with cost-effectiveness will want to weigh the greater power of various external motivators against the lower cost of task-related and some need-related incentives.

Among the many studies of traits that might respond to need-related incentives are those of Spielberger et al. regarding anxiety in students. Spielberger, O'Neil, and Hansen (1972) found that students experiencing high anxiety states tend to make more errors and perform less well on creative tasks than students who do not feel threatened. Spielberger et al. cite studies which found that computer instruction lessens stress on anxious students and enables them to perform better. However, an incentive system that included disincentives or competition for greatly desired rewards, particularly socially oriented rewards, might raise their anxiety levels to a point where their progress might be hampered. Even low-anxiety trait students, who may possibly perform better under slight pressure (Spielberger et al., 1972), may do poorly if they are low in ability and "the task is taken seriously" (Cronbach and Snow, 1977, p. 398).

Stress might be particularly likely if course materials are inadequate. According to Jamison, Suppes, and Wells (1974) a study made by Shrable and Sassenrath (1970) found "that an easy program with short steps is better suited to persons who are low on need for achievement and high on fear of failure or test anxiety, and that a difficult program with long steps is preferable for those with a high need for achievement and low fear of failure." However, Tobias and Abramson (1971) failed to replicate this anxiety finding. Cronbach and Snow (1977) summarize a number of experimental studies on this topic by saying that it is not yet "under control". All that can really be said about anxiety and need for achievement is that they may affect student progress differentially when a particular incentive system is introduced, and they should be monitored in case the incentive system would work better if tailored to this aspect of a trainee's personality.

### Innovations To Stimulate Responsbility

Another way to view the problem of motivation in a self-paced course is to consider how to stimulate self-responsibility for one's own learning in a student who is not used to having any latitude in his rate. Gordon (1970) suggests a *gradual* transfer from the teacher to the students, of responsibility for what, how, and how fast students learn.

AFHRL-TR-78-7

Other techniques that may promote more responsibility in students in a self-paced learning environment have been explored by several researchers. One of these is "micro teaching" or "peer tutoring" in which a student who has mastered a skill assists a peer in learning it. Colton (1974) found that the peer tutor was able to do this at no cost to his own performance, although he may not have included time as a consideration. Sloan (1970) found that college students successfully counseled other college students on academic, social, and personal adjustment problems. This could be extended to having a student counsel another one in techniques to increase his rate of progress. In one of the Lowry incentive studies, the chance to tutor a classmate was viewed as one of the more desirable high-feasibility incentives. Of course, a strong incentive system might work against the willingness of faster students to serve as tutors, unless it were structured to foster helpfulness. (The same reasoning would suggest that cheating might become a problem with a strong incentive system. ) From a cost-effectiveness point of view, it would be necessary to test whether enough improvement in student times occurred to offset the cost of the program. Cohen and Fishbein (1976) report success in training company commanders to have different behavior intentions using CAI, and this technique (a CAI guidance program) could help the student identify and alter his attitudes toward learning. Cogswell (1966) counseled students regarding their performances via computer. A similar program could coach the student in techniques for making rapid progress in the course. Such programs could conceivably replace at least some of the counseling sessions specified in the current SPMS for students who fall behind.

A type of intervention strategy that was not included in the Air Force studies is giving the student information on the average or top performance times of past students on a section of a course. This could provide a standard by which the trainee could gauge his own learning rate, and could be given to him either instead of, or along with, the prediction charts based on their presassessment scores that are presently used. Colton (1974) gave students in one section of his self-paced college media course the average completion times of past students on each unit, telling them to treat the information as "a possible guide to determine how efficiently they were using their time" (p. 284). The experimental section averaged less time than the control section on 18 of the 22 tasks; however, they averaged more time than the pilot group whose average times they had been told. The format, aims, subject matter, and students in this class were different from those in the Lowry courses, and it was not a tightly controlled experiment, so comparability is uncertain. In a study of naval personnel learning to use complex control systems, Myers (1969) found that giving students information on post-training performance times of course graduates definitely improved their speed. Again, circumstances were different from the Lowry courses, and transferability is uncertain.

Teel (1967) reporting the use of a contract approach in an electricity-electronics course found that some students responded well to this treatment and some did not, feeling that they needed more guidance in acquiring fundamentals. Face-to-face meetings setting up contracts on a smaller scale, e.g., overcoming a lack of a prerequisite skill or completing a block of lessons by a certain date – might help some Lowry students to manage their time better. It might possibly improve class times if contracts were used for all students rather than only as a counseling tool for those who fall behind.

Other techniques that have been used include team learning and varied presentation of material, where the student chooses how he wishes to be taught a given section of the course.

*Methods to Improve Course Material*

Finally, it should be noted that good course material can stimulate an interest in learning and maximize progress, and improving instructional materials could conceivably lead to significant time savings. This can be done rigorously, using a computer model of the state of knowledge of the student at any point in the course, as is being attempted by Self (1974). It could also be done (through experimentation) by defining an aptitude-treatment decision network to evaluate student progress at many checkpoints and provide the learning conditions most suited to his current state of knowledge, ability, and personality (Schwen, 1973). Farr (1973) describes three sophisticated CAI programs that combine thorough analysis of the subject's structure with flexible presentation. These approaches are expensive and may not be achievable with some types of course material. A more practical approach to improving the course might be to elicit detailed feedback from students who encounter difficulties in the course and from people who work closely with students. Also, records of student times on small segments of the course could be analyzed to pinpoint bottlenecks. Perhaps counseling sessions could include discussions of what changes in the course might help student progress. This emphasis in counseling sessions might improve the counselors' perceptions of their role and make them more likely to follow procedures. Also, typical gaps in the entering students' background could be identified – perhaps with "mini-lesson" tests such as those implemented by Wagner et al. (1973) – and suitable precourse remediation provided; experimentation would determine whether such remediation improved course completion times. Anderson (1976) found in a study of 90 eighth-grade students that "A group of students enter a particular learning sequence with unequal amounts of relevant prior learning," but "by complementing inequality in learner characteristics with inequality in instructional time and help in the early units, we can approach student equality in later units, not only in the achievement level attained, but also in the amount of on-task time needed to attain the criterion level" (p. 233). Structured observation of student behavior in the classroom has been used by Spielberger et al. (1972) to identify portions of a course that produce anxiety in students, and by Yeager and Lindvall (1968) to evaluate instructional innovations. Finally, course material can be examined to see whether it follows sound educational principles such as eliciting active behavior from the student (Suppes, 1964), providing quick feedback remediation with "enough information for the student to diagnose his own shortcomings" (Rosenbaum 1969, p. 3), and avoiding ambiguity, excessive repetition, and requiring unrelated skills.

*Conclusions*

Based on the review of the literature on mathmatical models of student progress and on motivation and intervention strategies, the following conclusions were made:

- There are very few global mathematical models of student progress.
- The mathematical models that have been found to be most successful at prediction have been those with relatively simple structure.
- The accuracy of prediction of student progress increases as a function of the relevance of the predictor variables to the learning task.
- Powerful external incentives improve student performance; intrinsic (task-related) or social incentives have less effect, but are generally easier to implement because of lower cost.
- Some powerful external incentives may displace and others augment the effect of intrinsic or social incentives when both are present.
- Response to incentives varies by the demographic characteristics of the students, such as age and socioeconomic status, as well as by aptitude and prior knowledge. It may also vary depending on measurable personality traits.
- Aside from incentives, rate of learning can be increased by improving instructional material and the quality of instruction, and by fostering self-responsibility through innovative approaches.

AFHRL-TR-78-7

## Appendix B
## SUMMARY STATISTICS

For purposes of reference, the summary statistics for the data included in the analysis are shown in Table B-1 and B-2. Table B-1 shows the means and standard deviations for the baseline variables cumulative block elapsed times, and selected cumulative lesson elapsed times. Table B-2 gives the correlations among the variables included in the analyses.

AFHRL-TR-78-7

Table B-1

MEANS AND STANDARD DEVIATIONS FOR VARIABLES
INCLUDED IN THE EVALUATION
(n = 368)*

| Variable Type: | Variable Title | Mean | S.D. |
|---|---|---|---|
| Baseline | Sex | (79% male) | |
| | Highest School Year Completed | 12.4 | .9 |
| | Age at Course Entry | 20.9 | 2.6 |
| | Reading Vocab. General Scale | 13.3 | 5.1 |
| | Reading Vocab. Scientific Scale | 8.0 | 2.9 |
| | Reading Vocab. Total Scale | 21.3 | 7.4 |
| | Pre-Course State Curiosity | 64.1 | 8.8 |
| | Pre-Course State Anxiety | 37.4 | 8.9 |
| | Trait Curiosity | 25.5 | 5.8 |
| | Trait Anxiety | 36.4 | 8.8 |
| | Internal-External Scale | 14.6 | 4.3 |
| | Text Anxiety | 30.0 | 7.4 |
| | Prefer. for Audio Mode | 10.7 | 2.6 |
| | Prefer. for Visual Mode | 6.8 | 1.8 |
| | Prefer. for Printed Mode | 4.9 | 1.4 |
| | Experience with Self-Pacing | 5.5 | 1.4 |
| | Experience with Convent. Instru. | 6.8 | 1.5 |
| | IM/MF Reading Subscale 1 | 4.3 | 1.6 |
| | IM/MF Reading Subscale 2 | 6.1 | 2.1 |
| | IM/MF Reading Total | 10.3 | 3.2 |
| | IM/MF Logical Reasoning Scale | 20.3 | 7.1 |
| | Concealed Figures Scale | 5.8 | 2.7 |
| | Memory for Numbers Backward Scale | 18.6 | 2.8 |
| | Memory for Numbers Total Scale | 35.3 | 5.0 |
| Cumulative Actual Block Elapsed Times** | CABELT 1 | 1219 | 439 |
| | CABELT 2 | 2915 | 875 |
| | CABELT 3 | 5210 | 1449 |
| | CABELT 4 | 6601 | 1726 |
| | CABELT 5 | 8218 | 2038 |
| | CABELT 6 | 9405 | 2193 |
| Cumulative Lesson Elapsed Times at the End of Each Block | CET9 (Block 1) | 1033 | 336 |
| | CET19 (Block 2) | 2378 | 677 |
| | CET31 (Block 3) | 4025 | 1073 |
| | CET40 (Block 4) | 5063 | 1311 |
| | CET52 (Block 5) | 6250 | 1504 |
| | CET61 (Block 6) | 7186 | 1640 |

*Students in course version 1 with all block elapsed times reliable and a course entry date in 1977.

**Sum of block elapsed times minus the sum of absence times.

Table B-2

CORRELATIONS BETWEEN BASELINE VARIABLES AND CUMULATIVE TIME VARIABLES AND AMONG CUMULATIVE TIME VARIABLES
(n = 368)

| Variable Label | CABELT1 | CABELT2 | CABELT3 | CABELT4 | CABELT5 | CABELT6 |
|---|---|---|---|---|---|---|
| SEX | .28 | .27 | .27 | .27 | .26 | .25 |
| HIYEAR | .01 | .01 | .01 | .00 | .00 | .01 |
| ENTAGE | .08 | .08 | .06 | .05 | .02 | .03 |
| RVOCGN | –.22 | –.27 | –.27 | –.27 | –.27 | –.25 |
| RVOCSC | –.20 | –.24 | –.26 | –.26 | –.27 | –.25 |
| RVOCTL | –.23 | –.28 | –.29 | –.29 | –.29 | –.27 |
| STCUR | .07 | .05 | .02 | .03 | .01 | .02 |
| STANX | .08 | .09 | .11 | .10 | .10 | .10 |
| TRCUR | .09 | .08 | .07 | .07 | .06 | .06 |
| TRANX | .03 | .00 | .02 | .03 | .04 | .04 |
| IESCL | –.04 | –.13 | –.17 | –.17 | –.18 | –.17 |
| TSTANX | .13 | .17 | .17 | .18 | .19 | .18 |
| PREFA | .00 | .00 | .01 | .01 | .01 | .00 |
| PREFV | .01 | .03 | .04 | .03 | .03 | .03 |
| PREFP | –.04 | .05 | .06 | .06 | .06 | .06 |
| EXPSP | –.08 | .10 | .11 | .12 | .11 | .12 |
| EXPCI | –.18 | –.25 | –.27 | –.28 | –.29 | –.29 |
| READS1 | –.24 | –.30 | –.30 | –.31 | –.32 | –.32 |
| READS2 | –.28 | –.33 | –.33 | –.32 | –.31 | –.30 |
| READST | –.31 | –.37 | –.37 | –.36 | –.37 | –.35 |
| LOGREA | –.23 | –.34 | –.31 | –.31 | –.30 | –.30 |
| CONFIG | –.05 | –.10 | –.10 | –.11 | –.10 | –.11 |
| MEMNB | .05 | –.09 | –.16 | –.15 | –.14 | –.13 |
| MEMNT | .02 | –.14 | –.20 | –.18 | –.17 | –.15 |
| CABELT1 | — | .87 | .76 | .75 | .72 | .71 |
| CABELT2 | .87 | — | .93 | .91 | .88 | .86 |
| CABELT3 | .76 | .93 | — | .99 | .96 | .94 |
| CABELT4 | .75 | .91 | .99 | — | .98 | .97 |
| CABELT5 | .72 | .88 | .96 | .98 | — | .99 |
| CABELT6 | .71 | .86 | .94 | .97 | .99 | — |
| CET9 | .80 | .80 | .73 | .72 | .70 | .69 |
| CET19 | .73 | .84 | .80 | .79 | .78 | .76 |
| CET31 | .69 | .80 | .82 | .83 | .80 | .79 |
| CET40 | .68 | .78 | .81 | .83 | .81 | .80 |
| CET52 | .68 | .77 | .80 | .83 | .83 | .83 |
| CET61 | .66 | .76 | .79 | .82 | .83 | .84 |

AFHRL-TR-78-7

Table B-2 (Continued)

| Variable Label | CET9 | CET19 | CET31 | CET40 | CET52 | CET61 |
|---|---|---|---|---|---|---|
| SEX | .27 | .22 | .23 | .23 | .22 | .21 |
| HIYEAR | .00 | .00 | -.02 | -.03 | .00 | .01 |
| ENTAGE | .06 | .04 | .05 | .05 | .05 | .05 |
| RVOCGN | -.32 | -.29 | -2.7 | -.25 | -.23 | -.21 |
| RVOCSC | -.25 | -.25 | -.22 | -.20 | -.19 | -.17 |
| RVOCTL | -.32 | -.30 | -.28 | -.26 | -.23 | -.22 |
| STCUR | .06 | .05 | .06 | .05 | .04 | .05 |
| STANX | .11 | .12 | .10 | .09 | .07 | .06 |
| TRCUR | .10 | .08 | .09 | .09 | .10 | .09 |
| TRANX | .01 | .01 | .00 | -.01 | -.02 | -.02 |
| IESCL | -.09 | -.13 | -.11 | -.10 | -.08 | -.06 |
| TSTANX | .19 | .22 | .19 | .17 | .15 | .12 |
| PREFA | .04 | .02 | .00 | .01 | .01 | .01 |
| PREFV | .02 | -.05 | -.06 | -.04 | -.05 | -.03 |
| PREFP | -.09 | -.09 | -.09 | -.08 | -.07 | -.06 |
| EXPSP | -.15 | -.15 | -.14 | -.14 | -.13 | -.13 |
| EXPCI | -.28 | -.28 | -.29 | -.28 | -.26 | -.26 |
| READSI | -.30 | -.31 | -.28 | -.28 | -.28 | -.28 |
| READS2 | -.30 | -.29 | -.28 | -.28 | -.26 | -.26 |
| READST | -.35 | -.35 | -.33 | -.33 | -.32 | -.31 |
| LOGREA | -.28 | -.32 | -.30 | -.29 | -.27 | -.26 |
| CONFIG | -.10 | -.09 | -.06 | -.06 | -.03 | -.04 |
| MEMNB | -.09 | -.09 | -.10 | -.10 | -.08 | -.07 |
| MEMNT | -.09 | -.10 | -.12 | -.12 | -.10 | -.09 |
| (For CABELTs (see previous page) | | | | | | |
| CET9 | — | .90 | .82 | .78 | .73 | .71 |
| CET19 | .90 | — | .93 | .89 | .84 | .81 |
| CET31 | .82 | .93 | — | .98 | .94 | .91 |
| CET40 | .78 | .89 | .98 | — | .98 | .95 |
| CET52 | .73 | .84 | .94 | .98 | — | .99 |
| CET61 | .71 | .81 | .91 | .95 | .99 | — |

61



U.S. GOVERNMENT PRINTING OFFICE: 1978-771-122/41