PREDICTING READING ACHIEVEMENT FROM MEASURES AVAILABLE DURING COMPUTER-ASSISTED INSTRUCTION

by

J. O. Campbell, E. J. Lindsay, and R. C. Atkinson

TECHNICAL REPORT NO. 249

January 20, 1975

PSYCHOLOGY AND EDUCATION SERIES

This research was sponsored by National Science Foundation Grant NSF EC-43997.

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Abstract

A high correlation between on-line rate of progress and student achievement on a standardized test was found for a computer-assisted instruction (CAI) program in initial reading. In most cases, CAI measures of progress were better indications of spring test performance than was the pretest given in the fall. Rates of progress in the parts or strands of the CAI program were highly correlated with each other, but certain strands proved to be better predictors of spring test scores than an overall rate measure.

Regression models were developed to relate spring test scores to amount of time spent in CAI; results from these models were in accord with data from an earlier experimental study designed to evaluate the effectiveness of the CAI program. Using a stepwise regression, which included both pretests and rates of progress in the strands, multiple correlations were obtained of .79 for the Cooperative Primary Test and .84 for the Metropolitan Achievement Test.

PREDICTING READING ACHIEVEMENT FROM MEASURES AVAILABLE DURING COMPUTER-ASSISTED INSTRUCTION

J. O. Campbell, E. J. Lindsay, and R. C. Atkinson
Stanford University
Stanford, California 94305

The Stanford computer-assisted instruction (CAI) program in initial reading has been under development at Stanford University over a period of eight years (Atkinson, 1968, 1974; Atkinson, Fletcher, Lindsay, Campbell, and Barr, 1973). A recent experimental study has shown that this method of individualized instruction produces significant gains in reading over what would be expected from classroom instruction alone (Fletcher and Atkinson, 1972). Here we present the results of a correlational study relating on-line measures of progress in CAI to posttest achievement.

Computer-assisted instruction is important in teaching reading because it provides effective individualized instruction. Our interpretation of the literature on teaching children to read is that when instruction is not individualized, method variables account for a small proportion of the variance in reading achievement. Much of our work is aimed at making the teaching sequence sensitive on a moment-to-moment basis to the student's unique history of performance.

Improving individualized instruction requires accurate estimates of the learner's state of knowledge about various classes of items; for example, sight words at a given level of difficulty or specific groups of spelling patterns. In addition, the relationship between a student's state of knowledge at various points in the CAI curriculum and his subsequent performance on a standardized test needs to be specified. This report presents models for predicting a student's achievement on a standardized test from measures taken during CAI in initial reading. The aim of this study was not to evaluate the effectiveness of the program; rather it was to assess the predictive power of components of the CAI program and to develop procedures to identify student strengths and weaknesses.

Description of the Reading Program

The Stanford CAI program is comprised of the following seven parts, called strands.

1. Letter Identification
2. Sight-word Recognition
3. Spelling Patterns
4. Phonics
5. Spelling
6. Word Comprehension
7. Sentence Comprehension

Each strand has been designed to provide practice on a particular set of reading skills. In any session the student may study curriculum items from any or all strands. The amount of time spent in each strand is selected to maximize the student's progression through the curriculum. As shown in Figure 1, entry into a strand is determined by the student's level of achievement in the other strands. The student begins with letter identification; when he has mastered a subset of letters used in the initial words of the sight-word strand, he begins that strand. Entry into the other strands is controlled in a similar fashion. A detailed

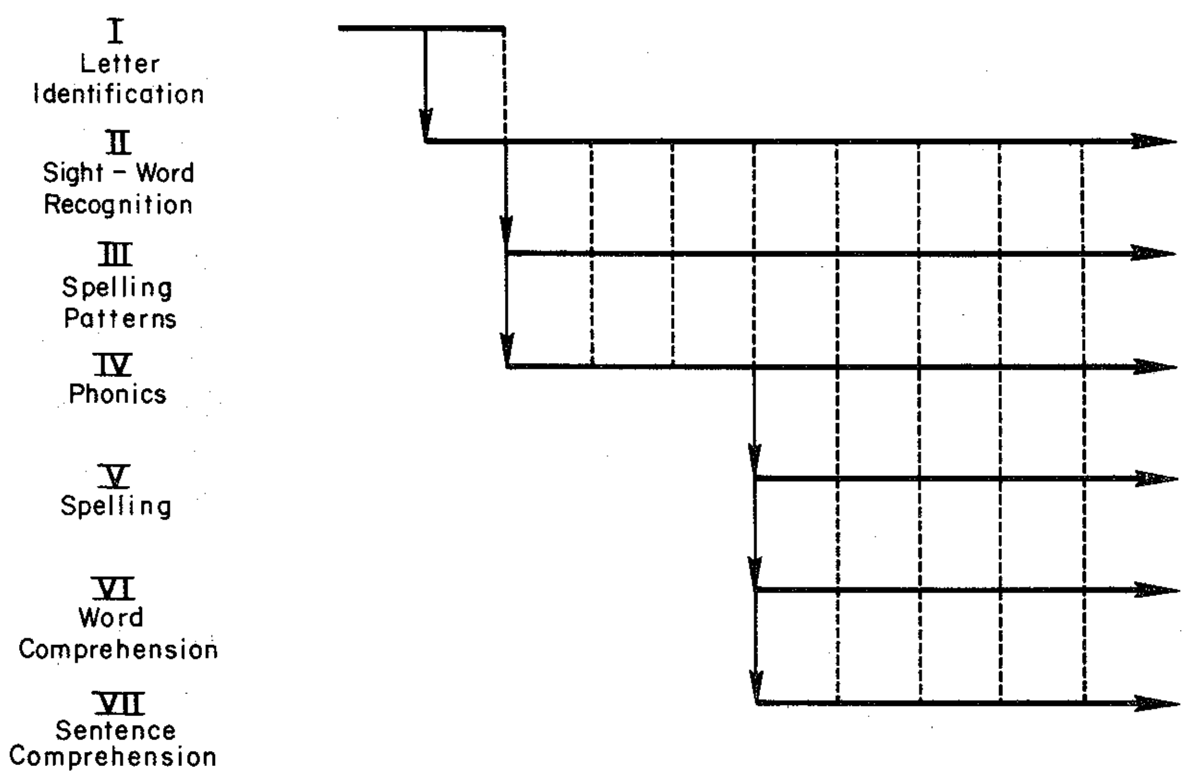

Figure 1. Initial entry points into strands. Vertical dotted lines represent maximal rate contours which control the student's progression in each strand relative to the other strands.

description and formats used in the program are given in Atkinson et al. (1973). Here we summarize the major elements and present a few examples.

In each strand the student studies a curriculum item in several different instructional formats. The instructional procedure varies from one exercise to the next, but in each a curriculum item is presented, a response is elicited from the student, and feedback is given. For example, the recognition exercise in the spelling patterns strand has the following format:

Teletype Display	Audio Message
BIKE LIKE STRIKE	Type STRIKE

Three words with similar spelling patterns are presented on the teletypewriter, followed by an audio presentation of one of the words. If the student types the correct response, a "+" is printed indicating that the student was correct. In addition, the audio may give a reinforcing message such as "great" or "fantastic" depending on the student's overall level of performance. If the student responds incorrectly or exceeds the allotted time, the program prints the correct word and gives audio feedback about the nature of the error.

An example from the phonics strand illustrates the build-a-word exercise: A spoken cue is given ("type 'stuff'") and part of a word is printed (ST-). The student must recognize the correct ending from among three printed alternatives (-UFF -OP -EP). In the word comprehension strand the student is required to select one of three words which fits a given category. The student may, for example, be presented with "CANDY RUN CAR" and asked to type the word that is something to eat. Sentence comprehension is handled by a fill-in-the-blank format where the student is asked to recognize the correct item. An example is, "TED SWAM TO THE ---," with choices, "STAR RAFT RUN." Associated with each exercise is a performance criterion that must be met before the student moves to the next exercise.

Figure 2 illustrates the procedure in each strand for deciding which item the student is to study and in which exercise format it will be presented. The process shown is common to all strands except that some strands have additional provisions for review and pretest. The items the student is to study are sampled from a working pool of items drawn from the master curriculum file, and are presented in one of the exercise formats. The sampling continues until each item in the working pool has been presented. When this occurs, a decision is made to shift the student to another strand, to sign the student off the system depending on the elapsed time, or to replace those items from the working pool which the student has brought to criterion and continue in the current strand.

The equipment used is quite simple from a student's viewpoint: a model KSR-33 teletypewriter and an audio headset. The instructional program is written in SAIL, which is an expanded form of ALGOL, and runs on a PDP-10 computer. Since the program is directed to students who cannot read, spoken communication is necessary. Digitized audio is used; the voice pattern is sampled and the result stored in digital form that can be accessed as needed to reconstitute the spoken word or phrase. The audio system permits fast and essentially random access to any of more than 7,000 items.

When the student logs in at a terminal, his response history is retrieved and the instructional materials are selected for the day. The

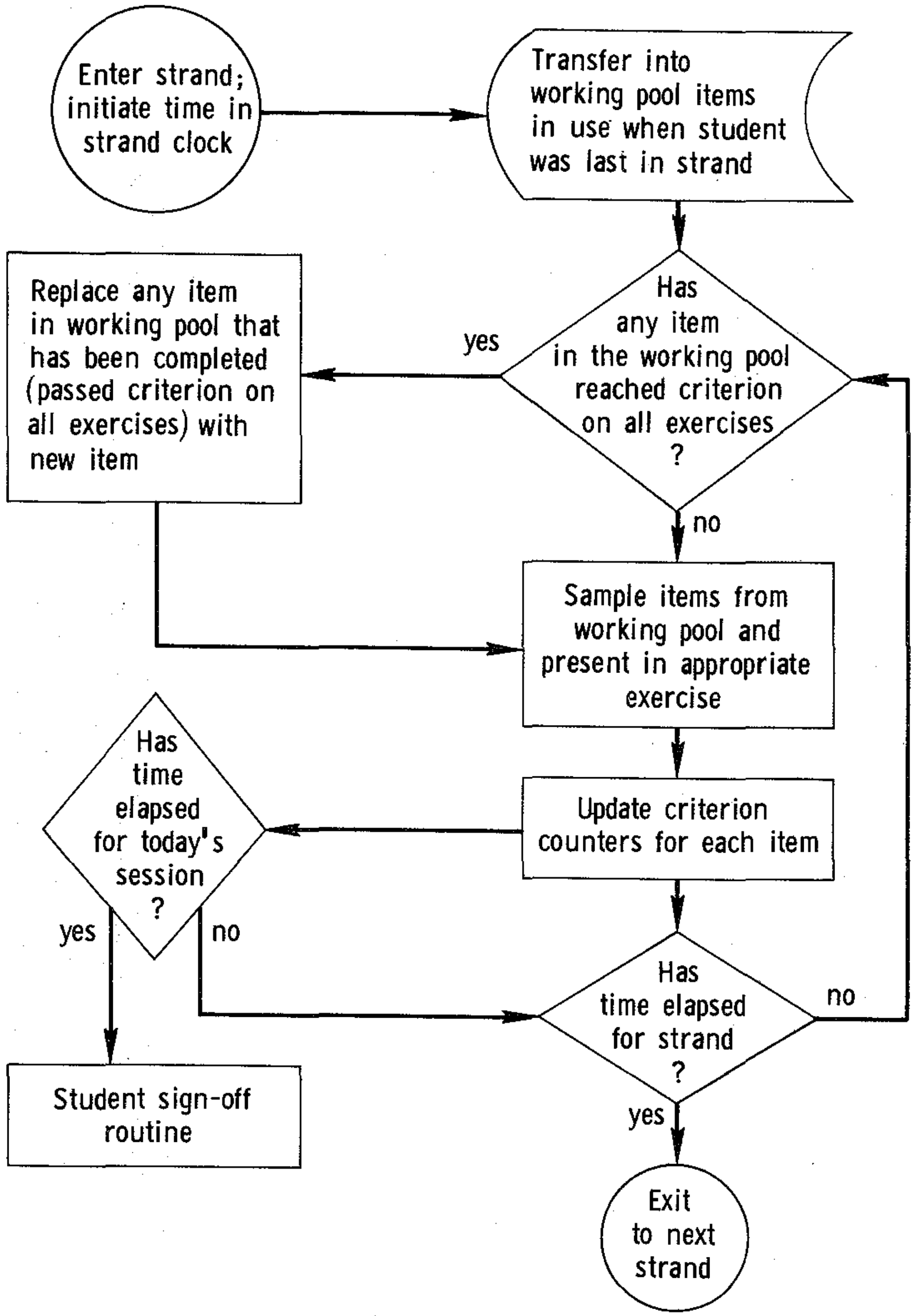

Figure 2. Flow diagram for presentation of curriculum items.

student may study in one or all of the strands; when he finishes, the history record is updated and stored in the computer.

METHOD

Subjects

Second grade students in the Stanford CAI Reading Program were chosen, as part of the Compensatory Reading Project carried out by Educational Testing Service (ETS). The sample of students used in our study was drawn from these second grade classes and involved 69 students (42 boys and 27 girls). All students received 15 minutes of CAI per day in addition to their normal classroom instruction. The students are predominantly from lower income Black families.

Test Instruments

As part of the evaluation, personnel hired by ETS administered the Word Knowledge and Reading sections of the Metropolitan Achievement (MAT) and the Cooperative Primary (COOP) tests in October, 1972, and again in May, 1973. COOP form 12A and Metropolitan Primary I (form G) were used in the fall, while COOP 12B and Metropolitan Primary I (form F) were used in the spring. Scores from the two MAT subtests were combined to yield a total reading score. More difficult forms of these tests are usually used in the spring, but ETS decided, given the sample selected, to use for the spring test parallel forms at the same level as the fall tests.

In this study we evaluate reading gains and develop regression models to predict achievement on the spring tests as a function of rates of progress in the CAI strands and fall test scores. The rate of progress for each student in a strand of the CAI program was computed by dividing the number of items reaching criterion by the time in minutes spent on the strand. Thus, the rate measure represents the average time to reach criterion on an item in a given strand. These rate measures were obtained for all strands, except for the letter strand; this strand provides an introduction to the CAI program and for second grade students is not sensitive to individual differences. An average rate measure over all strands was then obtained for each student. Since rates varied as a function of the strand, a *z*-score was computed for each student on each strand, thereby placing all rates on a common scale. The *z*-scores were then averaged over the six strands to obtain an overall rate measure for each student.

Rates of progress in each strand (rather than items covered) were used because students differed in the time they spent in the program and on any given strand. The mean CAI time per student was 18.9 hours with a standard deviation of 3.8 hours. In summary, the variables used are as follows:

Pretests	Rates in Strands	Posttests
COOP 12A	Spelling	COOP 12B
Metropolitan G Word Knowledge	Word	Metropolitan F Total Reading
	Patterns	
Metropolitan G Reading	Phonics	
Metropolitan G Total Reading	Word Comprehension	
	Sentence Comprehension	
	Average *z*-score	

RESULTS AND DISCUSSION

Achievement

Table 1 summarizes achievement on pre- and posttests and presents percentile equivalents in comparison to national norms. As expected, the student's scaled scores increased during the year on all tests (paired t-tests were significant at the .001 level for all comparisons). Percentile scores based on national norms also increased for all tests. Examination of Table 1 indicates that the pattern of results is similar for boys and girls. The fall tests were quite difficult for these students; the apparent gains may reflect only a better fit of the test to their ability in the spring, and should thus be interpreted cautiously.

Differences in Rates

Table 2 presents rate of progress in each strand. The total group was divided by sex and also into the top, middle, and bottom thirds based on fall scores on the COOP. In this sample the girls progressed faster than the boys, and the top group progressed more quickly than the middle and bottom groups. Note, however, that there is little if any difference between the rates for the middle and bottom groups. This lack of difference in rates was also reflected in the spring test scores for these two groups. Even though the bottom and middle groups differed significantly on fall test scores, their spring scores were nearly identical (138 vs. 137 for the middle and bottom groups on the spring COOP and 40 vs. 39, respectively, on the MAT Total Reading). This result is comparable to other findings (Atkinson, 1968) indicating that CAI effects the largest relative gains for students at the low end of the distribution.

Table 1

Scaled Score, Standard Deviation, and National Percentile for the Cooperative Primary (COOP) and Metropolitan Achievement (MAT) Tests

Test	Fall			Spring		
	Percentile	Mean	SD	Percentile[1]	Mean	SD
Total Group						
COOP	7	133.1	8.0	27	140.3	6.0
MAT Word	12	33.9	12.4	16	47.1	12.8
MAT Reading	10	32.1	10.9	20	44.0	11.5
MAT Total	10	31.7	10.8	14	44.1	11.3
Boys						
COOP	7	132.9	8.6	19	139.3	5.8
MAT Word	12	33.7	12.0	14	45.6	13.3
MAT Reading	8	30.4	10.5	18	42.5	11.7
MAT Total	8	30.5	10.8	12	42.7	12.1
Girls						
COOP	7	133.5	7.3	34	141.9	6.0
MAT Word	12	34.1	13.1	22	49.3	11.8
MAT Reading	18	34.9	11.1	23	46.5	11.0
MAT Total	12	33.4	10.7	20	46.4	9.8

[1] National spring percentiles are based on test forms other than those used for this sample. Scaled scores are basically equivalent across forms and levels of the tests, but spring percentiles should be interpreted as approximate.

Table 2

Mean Rates of Progress in Strands[1]

Strand	Total	Boys	Girls	Top Third	Middle Third	Bottom Third
			Group			
Spelling	.40	.39	.41	.68	.25	.25
SD	(.33)	(.36)	(.29)	(.32)	(.25)	(.20)
Sight Words	.71	.68	.75	.89	.63	.61
SD	(.29)	(.30)	(.27)	(.27)	(.27)	(.26)
Patterns	.71	.67	.77	.98	.58	.57
SD	(.32)	(.34)	(.28)	(.27)	(.29)	(.21)
Phonics	.79	.77	.83	1.00	.69	.69
SD	(.29)	(.33)	(.23)	(.19)	(.29)	(.28)
Word Comp.	.69	.65	.75	1.01	.50	.55
SD	(.40)	(.43)	(.35)	(.32)	(.31)	(.36)
Sentence Comp.	.63	.59	.70	.91	.49	.50
SD	(.32)	(.33)	(.29)	(.26)	(.24)	(.26)

[1]Rates are in items completed per minute; standard deviations for these measures are given in parentheses.

Correlation of Rates and Pretests with Spring Achievement

Table 3 presents correlations of rates with spring COOP and spring MAT total scores. Rates in most of the strands are more highly correlated with spring achievement than are the pretests. That is, the on-line rate measures are a better predictor of student ability than is a parallel form of the test given in the fall. All of the rates are more highly correlated with achievement on the spring Metropolitan than on the COOP tests. Note that the rates of progress in the spelling and sentence comprehension strands are more highly correlated with spring achievement than is the average. These strands are relatively difficult; while the student encounters them after the others, most students reached them well before the spring.

Table 4 presents intercorrelations among variables. The rates are highly correlated with each other (.81 to .91), indicating that students who move rapidly in one instructional strand (for example, on spelling patterns) tend to move rapidly in other strands (for example, sight words and comprehension). This undoubtedly reflects both an aspect of student ability and similarities in the strands of the program. In most cases rates correlate more highly with spring than fall achievement. Moreover, correlation with spring achievement is fairly consistent across strands.

Linear stepwise regressions were used to develop separate models for posttest achievement on the COOP and MAT Total Reading tests. Table 5 presents the regression models, multiple $\underline{R}$, and the step at which each variable entered, together with the $\underline{F}$ to enter (Draper and Smith, 1966). A low cutoff of $\underline{F} = .01$ was used in order to include most of the variables

Table 3

Correlation of Pretests and Strand Rates with Spring COOP and MAT Scores

	Spring Test	
	COOP	MAT
Spelling	.69	.76
Sight Words	.52	.58
Patterns	.64	.71
Phonics	.61	.68
Word Comprehension	.68	.72
Sentence Comprehension	.73	.77
Average Rate	.69	.75
Fall COOP	.55	.55
Fall MAT Word	.60	.64
Fall MAT Reading	.59	.52
Fall MAT Total	.65	.64

Table 4

Intercorrelation of Strand Rates and Fall Test Scores

	Variable[1]										
	1	2	3	4	5	6	7	8	9	10	11
1 Spelling	1.00	.81	.91	.83	.88	.86	.94	.51	.59	.43	.56
2 Sight Words		1.00	.82	.85	.84	.82	.91	.37	.49	.33	.47
3 Patterns			1.00	.87	.86	.85	.94	.45	.56	.39	.54
4 Phonics				1.00	.91	.89	.95	.36	.49	.33	.47
5 Word Comprehension					1.00	.89	.96	.41	.52	.40	.52
6 Sentence Comprehension						1.00	.94	.48	.60	.50	.61
7 Average Rate							1.00	.46	.58	.42	.56
8 Fall COOP								1.00	.81	.56	.78
9 Fall MAT Word									1.00	.55	.89
10 Fall MAT Reading										1.00	.84
11 Fall MAT Total											1.00

[1]Column numbers correspond to numbered variables in rows.

Table 5

Regression Equations for Predicting Spring COOP and MAT Total Reading

COOP[1]

Step number	Variable entered	Regression coefficient	Multiple R	F to enter
1	Sentence Comprehension rate	10.69	.73	75.24
2	Fall COOP	.16	.76	7.85
3	Sight Word rate	-6.97	.77	2.47
4	Word Comprehension rate	5.27	.79	3.66
5	Spelling rate	3.28	.79	1.01
6	Phonics rate	-3.15	.79	.49

Intercept is 114.86. Standard Error of Estimate = 3.83.

MAT Total Reading[2]

Step number	Variable entered	Regression coefficient	Multiple R	F to enter
1	Sentence Comprehension rate	13.00	.77	98.54
2	Fall MAT Total	.09	.80	9.00
3	Spelling rate	14.46	.82	5.63
4	Sight Word rate	-12.38	.83	4.27
5	Word Comprehension rate	2.83	.83	.66
6	Fall MAT Word	.13	.83	.36
7	Phonics rate	5.06	.84	.19
8	Patterns rate	-3.10	.84	.19
9	Fall MAT Reading	.05	.84	.05

Intercept is 25.86. Standard Error of Estimate = 6.65.

[1]The rate measures for spelling patterns and average rate did not enter the regression equation with cutoffs at F = .01 and Tolerance Level = .01. Low cutoffs were used in order to include most of the variables for comparison.

[2]Under the cutoffs above, average rate did not enter the regression equation.

for comparison; even with this cutoff, the average rate measure did not enter the regression for either test. The increase in the multiple *R* as each variable is added is an indication of greater predictive power. After the sentence comprehension rate and fall test are used, little is gained by adding further variables. The multiple correlations obtained are relatively high and approach the reliability of the tests themselves. The sentence-comprehension rate by itself accounts for most of the variance in the posttest data; rather than looking to the average rate as a simple measure, sentence comprehension rate by itself can in effect serve in place of a reading test. Note also that the sight-word rate carries a negative regression coefficient for both the COOP and MAT tests, although it is positively correlated with both spring tests.

At a higher *F* cutoff, say $F = 3.5$, the results in Table 5 indicate that only sentence-comprehension rate and fall COOP would be included in predicting spring COOP. Similarly, only sentence-comprehension rate, fall MAT, spelling rate, and sight-word rate would be included in predicting MAT. The resulting multiple *R*'s are .76 for the spring COOP and .83 for the spring MAT; these compare favorably with the multiple *R*'s of .79 and .84 obtained with the low cutoff.

Relation of CAI Time to Spring Achievement

We have also estimated the parameters of a linear equation relating spring test scores to time on the CAI program. The equations for the COOP and MAT tests are as follows:

$$\text{COOP} = 134.18 + .32(\text{hours of CAI})$$

$$\text{MAT} = 30.97 + .70(\text{hours of CAI})$$

The equations are based on correlational data and should not be interpreted as suggesting a cause and effect relationship; however, they are in accord with earlier results based on experimental evidence (Fletcher and Atkinson, 1972).

CONCLUSION

This study has yielded several useful results. First, the high correlation of rates with spring COOP scores indicates that progress in the CAI program is highly related to a student's reading ability as measured by standardized achievement tests. Second, the high intercorrelations among CAI rates suggests that the several strands of the CAI program may be tapping the same skill, or that skills in one strand of the program are highly related to skills in others. However, the average rate score was not the best predictor of posttest scores; after the first rate measure went into the regression equation, the partial correlation of the average rate was so low that it did not enter the model under the cutoffs established. Third, regression models for relating test scores to instructional time yielded slope measures of .32 and .70 for the COOP and MAT, respectively. These slope parameters indicate the gain that can be expected with each hour of CAI, and can be used in formulating optimal policies for allocating instructional time among students (e.g., see Atkinson, 1972). They can also be used to estimate the amount of time needed for a student or group of students to reach a given level of reading performance. Finally, using entering achievement scores and rate measures from several strands, we obtained multiple $\underline{R}$'s of .79 for the COOP and .84 for MAT Total Reading.

Evidence from prior experimental work has shown that the CAI reading program is effective (Fletcher and Atkinson, 1972). This study has developed regression models for predicting posttest achievement from measures available during instruction. It is interesting to note that the rate measure associated with the sentence-comprehension strand is highly correlated with posttest scores (.73 for the COOP and .77 for the Metropolitan; see Table 3). This single statistic proves to be almost as good a predictor of achievement as the multiple R's. In experimenting with one or another version of the CAI program, one could use this measure by itself as a crude but continuous monitor of the effectiveness of an experimental manipulation. To the extent that such a measure is a valid predictor of posttest performance, we can reduce the effort and time involved in assessing a particular experimental manipulation.

References

Atkinson, R. C. Computerized instruction and the learning process. *American Psychologist*, 1968, *23*, 225-39.

Atkinson, R. C. Ingredients for a theory of instruction. *American Psychologist*, 1972, *27*, 921-931.

Atkinson, R. C. Teaching children to read using a computer. *American Psychologist*, 1974, *29*, 169-178.

Atkinson, R. C., Fletcher, J. D., Lindsay, E. J., Campbell, J. O., and Barr, A. Computer-assisted instruction in initial reading. *Educational Technology*, 1973, *13*, 27-37.

Chant, V. G., and Atkinson, R. C. Optimal allocation of instructional effort to interrelated learning strands. *Journal of Mathematical Psychology*, 1973, *10*, 1-25.

Draper, N. R., and Smith, H. *Applied regression analysis*. New York: John Wiley, 1966.

Fletcher, J. D., and Atkinson, R. C. An evaluation of the Stanford CAI program in initial reading. *Journal of Educational Psychology*, 1972, *63*, 597-602.