A Field Test of Computer-Assisted Instruction in First Grade Mathematics

FLOYD G. DELON*

FOR the past two years the South Central Region Educational Laboratory (SCREL) served as independent evaluator for the computer-assisted instruction project in mathematics conducted in McComb, Mississippi. This project represented the first grade segment of a multi-grade program implemented cooperatively by the McComb Public Schools and the Institute for Mathematical Studies in the Social Sciences, Stanford University. The programmed materials used were those designed by Patrick Suppes and the Institute staff. Financing for the project was derived from Title III of the Elementary and Secondary Education Act of 1965.

Although the recent literature contains numerous reports on computer-assisted instruction in general and the Suppes programs in particular, few of these provide research data relative to cognitive behavioral changes in children. The laboratory sought to identify such changes in samples of disadvantaged children. It is the purpose of this article to report the evaluator's findings in the McComb project.

Statement of the Problem

This study was designed to assess both the immediate (end of year) and the residual (one year follow-up) effects of computer-assisted instruction in mathematics presented to classes of first grade culturally disadvantaged pupils. More specifically, answers were sought to the following questions: (a) What effects (if any) did the treatment have on achievement in mathematics and in reading? (b) What effects (if any) did the treatment have on measured intelligence? (c) If the treatment produced significantly greater results than the regular instructional program alone, were these treatment differences maintained through the following year?

Setting and Target Population

McComb is located in Pike County some 90 miles south of Jackson, Mississippi. The population of McComb is about 12,500, and the McComb separate school district serves an attendance area of approximately 17,000 persons. More than 4,000 students attend school in the district, of which 60 percent are Caucasian and 40 percent Negro.

The experimental project was housed in an elementary school located near the city's edge. In 1967-69, this school enrolled 359 Negro elementary school children, 85 percent of whom are members of families with annual incomes under $2,000. Comparison classes were selected from a neighboring school in the same district serving other children from the same population. Nearly all children from this rural non-farm area are

* *Floyd G. Delon, Associate Dean and Professor of Education, University of Missouri-Columbia; formerly Director, South Central Region Educational Laboratory, Little Rock, Arkansas*

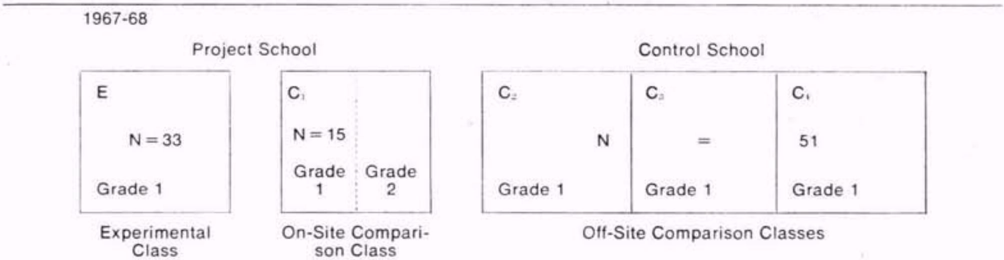

Figure 1. The Experimental Design, 1967-68

considered educationally and economically disadvantaged as evidenced by family income levels and standardized test scores. However, the primary criterion used for selection of subjects was eligibility of the schools for ESEA Title I support.

Experimental Design

The evaluation plan for the 1967-68 school year utilized a pretest—post-test experimental design with an on-site and multiple off-site comparison groups. The on-site group most nearly met the requirements for a true experimental control since children were randomly assigned within the project school to the experimental and the control classes. Also, teachers rotated from room to room such that the same teacher taught the same subject to each class. The only obvious extraneous variable was a split classroom (grades 1 and 2) situation involving this comparison group. The off-site comparison groups in the control school were all self-contained first grade classrooms utilizing identical instructional materials excepting the computer-assisted instruction. Within this school, pupils were randomly assigned among classrooms. The experimental design is diagrammed in Figure 1.

Figure 2 illustrates the experimental design for the 1968-69 school year. The design for the new experimental group and its comparison groups is similar to that of the previous year. However, the on-site control group was a full class of first grade pupils and the number of off-site comparison groups was limited to one.

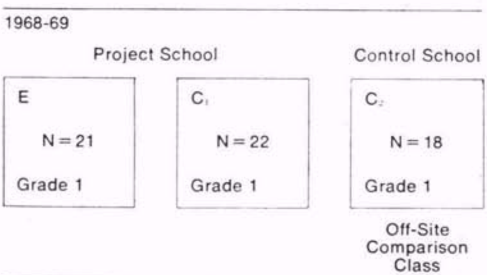

Figure 2. The Experimental Design, 1968-69

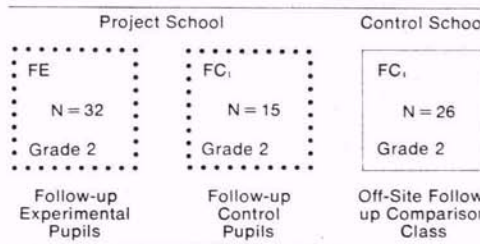

Figure 3. The Follow-up Study

In addition, a follow-up study was conducted of those pupils who received computer-assisted instruction during the initial year of the project. Pupils were assigned randomly to the two second grade classrooms, thus mixing experimental and control pupils of the previous year. Again the pretest—post-test design was used, this time with no intervening treatments. The dotted lines in Figure 3 illustrating this portion of the design indicate that these groups were not intact classes.

The Treatment

The Suppes program of computer-assisted instruction in mathematics comprised the treatment in this study. The description of this program, which appears in numerous reports, is briefly summarized in this section.

After an orientation session of four to six weeks, daily lessons referred to as "drill and practice" are administered to each pupil through computer-based, remote-controlled terminals. These lessons, approximately 10 minutes in length, are designed to review topics in arithmetic fundamentals formally introduced in the classrooms two or more weeks previously.

More specifically, the pupil must correctly type his assigned number and name on the computer terminal keyboard to initiate a lesson. Once the lesson begins, each problem is completely typed out under computer control, including a blank for any required response. A correct response is reinforced by the appearance of the next exercise. If a response is incorrect the words "no, try again" are typed out and the exercise itself is retyped. If the error is repeated, the message "no, the answer is ______" appears. Finally, if a third error is made the correct answer is given again and a new exercise is presented.

The year's work is divided into concept blocks each containing about seven days' work. The first lesson in each concept block is a pretest and the last lesson is a post-test. The pretest and post-test referred to here should not be confused with those mentioned as a part of the experimental design. Programs were constructed at five levels of difficulty, and after each lesson the pupil is assigned or reassigned to an appropriate level. Provision is made for review as a part of each day's work. Review constitutes approximately 30 percent of the daily lesson.

Instruments

Criterion measures were obtained exclusively through the administration of standardized tests. These widely used tests are listed below along with the purposes for which they were administered.

1. California Achievement Tests (CAT). Achievement was operationally defined as the scores in the arithmetic and reading sections of this test. Each section yields three scores: mathematical reasoning, mathematical fundamentals, and total mathematics; reading vocabulary, reading comprehension, and total reading respectively. However, for this study only the "total mathematics" and the "total reading" scores were analyzed.

Reading achievement was selected as a criterion variable because of the unique role this skill plays in the use of computer-assisted instruction by first grade pupils. Since all information transmitted by the computer must be read by the pupil, the orientation period prior to actual work at the terminal focuses on the rapid development of the required level of reading proficiency. It was anticipated that: (a) subjects studying computer-assisted instruction in mathematics might be motivated to improve their reading skills, and (b) that practice in reading messages from the computer might influence achievement in reading.

2. Wechsler Intelligence Scale for Children (WISC) and Otis-Lennon (O-L). Intelligence was operationally defined as performance on these two tests. These data were collected for two purposes: (a) to assess any effect of the treatment on measured intelligence; and (b) to verify the effectiveness of the random assignment in equating groups during the second year replication. Because the subjects were disadvantaged children it was assumed that the WISC might be sensitive enough to detect changes not apparent in group test results. However, after the initial year results failed to verify this assumption, the O-L was used exclusively for the 1968-69 post-test.

Analysis of Data

Because of the purposes cited in the previous section, the achievement data and the intelligence data were treated somewhat differently. In 1967-68, T tests were used to examine the achievement difference in mean

Continued on page 177

Continued from page 172

gain scores (post-test minus pretest) between the experimental group (E) and the control group (C_1) and the pooled mean gain scores for the off-site controls ($C_2 + C_3 + C_4$). The analysis of variance technique was used to test the difference between group means for the intelligence test data collected using the WISC as a post-test only.

The data for the 1968-69 first grade classes were analyzed using nearly identical procedures. Again, T tests were used to analyze data measured in terms of "total mathematics" and "total reading" scores. Also, analysis of variance was used with the WISC pretest data and with the O-L pretest and post-test data. The follow-up study was limited to achievement data and T tests were used to compare the mean gain difference between the second grade pupils who had and those who had not experienced computer-assisted instruction during the previous year.

Findings

The findings are presented as the results of tests of three sets of hypotheses. These sets, with component hypotheses, stated in null form, correspond to the three aspects of the problem statement.

Hypothesis Ia: There will be no significant difference in mean gain achievement scores in mathematics between the experimental group and the on-site comparison group.

Hypothesis Ib: There will be no significant difference in mean gain achievement scores in mathematics between the experimental group and the off-site comparison group.

For the 1967-68 school year, each of the null hypotheses was rejected for mathematics achievement, as the mean gain for the experimental group (E) was significantly greater (.01 level) than for each of the comparison groups (C_1 and $C_2 + C_3 + C_4$).

Hypothesis Ic: There will be no significant difference in mean gain achievement scores in reading between the experimental group and the on-site comparison group.

Hypothesis Id: There will be no significant difference in mean gain achievement scores in reading between the experimental group and the off-site comparison groups.

The mean gain in reading did not differ significantly between the experimental group and the on-site comparison group; therefore, Hypothesis Ic was accepted. On the other hand, Hypothesis Id was rejected because the mean gain of the experimental group (E) was significantly greater (.01 level) than the pooled mean gains of the off-site comparison groups ($C_2 + C_3 + C_4$). The results of the tests of Hypothesis Ia, Ib, Ic, and Id are presented in Table 1.

The four null hypotheses were again tested utilizing data collected by replicating the initial experiment. For the 1968-69 school year, these hypotheses were rejected for both mathematics and reading. As indicated in Table 2, the experimental group (E) gained significantly more (.01 level) in achievement than did either the on-site (C_1) or the off-site (C_2) comparison group.

Hypothesis IIa: There will be no significant differences among mean pretest intelligence test scores for the experimental group and the on-site and off-site comparison groups.

Hypothesis IIb: There will be no significant differences among mean post-test comparisons

		Total Mathematics							Total Reading						
Group	N	Pre-test Mean	σ	Post-test Mean	σ	Mean Gain	t	p	Pre-test Mean	σ	Post-test Mean	σ	Mean Gain	t	p
E	33	28.28	19.24	53.40	24.12	25.12			42.84	20.61	62.81	21.16	19.97		
C_1	15	20.29	15.10	33.61	18.50	12.78	2.79	.01	42.07	19.85	59.93	20.85	17.86	.89	N.S.
$C_2+C_3+C_4$	51	23.19	16.73	39.61	17.83	16.42	2.83	.01	38.77	21.13	48.32	19.42	9.55	3.62	.01

Table 1. T-test Results for CAT Mean Gain Differences, 1967-68 Academic Year

Group	N	Total Mathematics: Pre-test Mean	σ	Post-test Mean	σ	Mean Gain	t	p	Total Reading: Pre-test Mean	σ	Post-test Mean	σ	Mean Gain	t	p
E	21	3.48	2.86	48.95	22.96	45.47			19.48	6.99	57.10	12.67	37.62		
C_1	22	2.41	2.52	28.14	18.31	25.73	3.04	.01	15.32	5.34	30.45	15.04	15.13	6.51	.01
C_4	18	4.00	3.68	27.75	10.65	24.07	4.30	.01	16.85	7.30	44.15	13.15	27.30	3.24	.01

Table 2. T-test Results for CAT Mean Gain Differences, 1968-69 Academic Year

for the experimental group and the on-site and off-site comparison groups.

Both hypotheses IIa and IIb were accepted for all pretest and post-test comparisons. As indicated in Table 3, differences among the post-test WISC mean scores were not significant for the 1967-68 school year.

Source	d.f.	MS	F	p
Between Classes	2	28662.24	1.302	N.S.
Error	77	22007.89		
Total	79			

Table 3. Analysis of Variance of Mean Total IQ Scores at the End of the 1967-68 Academic Year

For the 1968-69 school year, Hypothesis IIa was again accepted for both the pretest analyses which included separate comparisons of WISC and O-L mean scores. Hypothesis IIb was again accepted for the post-test analysis which involved only the O-L mean score comparisons. These results are presented in Tables 4 and 5.

Source	d.f.	MS	F	p
Between Classrooms	2	274.6949	1.78	N.S.
Error	65	153.5170		
Total	67			

Table 4. Analysis of Variance of Mean WISC Total IQ Scores at the Beginning of the 1968-69 Academic Year

Hypothesis IIIa: There will be no significant difference in the mean gain achievement scores in mathematics between the follow-up experimental group and the on-site follow-up comparison group.

Hypothesis IIIb: There will be no significant difference in the mean gain achievement scores in mathematics between the follow-up experimental group and the off-site follow-up comparison group.

In the year following the termination of the treatment, the follow-up experimental group (FE) did not differ significantly from the on-site follow-up group (FC_1). Hypothesis IIIa was therefore accepted. On the other hand, the off-site comparison group (FC_4) gained significantly more (.01 level) than did the follow-up experimental group. Consequently, Hypothesis IIIb was rejected.

Hypothesis IIIc: There will be no significant difference in mean gain achievement scores in reading between the follow-up experimental group and the on-site follow-up comparison group.

Hypothesis IIId: There will be no significant difference in mean gain achievement scores in reading between the follow-up experimental group and the off-site follow-up comparison group.

In reading achievement, also, the mean gain for the follow-up experimental group (FE) did not differ significantly from the on-site follow-up comparison group (FC_1). Therefore, Hypothesis IIIc was accepted. The mean gains of FE, however, continued to be significantly greater than the off-site comparison group (FC_4) at the .05 level, and Hypothesis IIId was rejected.

Testing	Source	d.f.	MS	F	p
Beginning of Year	Between Classrooms	2	104.6042	1.12	N.S.
	Error	54	93.7364		
	Total	56			
End of Year	Between Classrooms	2	120.6899	6.84	N.S.
	Error	54	144.4623		
	Total	56			

Table 5. Analysis of Variance of Mean Otis-Lennon IQ Scores at the Beginning and End of the 1968-69 Academic Year

Discussion and Conclusions

The present study was designed to provide answers to three basic research questions. These questions were used as the bases for organizing this section.

What effects (if any) did the treatment have on achievement in mathematics and reading? The findings support the conclusion that the use of computer-assisted instruction can increase the mathematics achievement of culturally disadvantaged pupils. Both the initial study in 1967-68 and the replication in 1968-69 revealed superior gains by the experimental group during the treatment period.

The data on reading achievement must be interpreted with added caution. Superior gains in three out of four of the comparisons suggest that computer-assisted instruction in mathematics did produce significantly higher gains in first grade reading achievement. However, no significant difference was found in one instance in which the same teacher taught reading to both the experimental and the comparison group. This would suggest that the obtained differences are due in part to the teacher variable.

While no attempt was made to compare the 1967-68 and the 1968-69 groups statistically, the post-test scores would appear to be within the same range. The higher pretest scores for the 1967-68 groups can be explained by delay in testing which occurred during the initial years. These pretests were administered six to eight weeks later than those of the subsequent year.

What effects (if any) did the treatment have on measured intelligence? Data collected using both individual and group intelligence tests revealed no significant differences among groups. Thus, as indicated by pretest comparisons, the randomization procedures appear to have equated groups with respect to intellectual ability. On the other hand, the post-test comparisons showed no significant differences in gains in measured intelligence as a result of the treatment.

Were treatment differences maintained through the following year? The follow-up study of the initial experimental and control groups indicated that the experimental group failed to maintain superior gains in mathematics achievement after the termination of computer-assisted instruction. In fact, all differences in mean gain favored the comparison groups. If this trend is representative, it may be concluded that when the treatment is ended after grade, the differences in mathematics achievement as measured by the California Achievement Test tend to disappear.

The findings in the follow-up of reading achievement were consistent with the results of the treatment year. Significantly greater gains by both the follow-up experimental group and the on-site follow-up comparison group (taught by the same teacher) again suggest that these differences may be attributable as much to the teacher variable as to computer-assisted instruction.

Since no opportunity was available in this setting to follow up pupils who continued to receive computer-assisted instruction, it is impossible to draw conclusions concerning the accrued effects of such continued use. This additional information is needed by school officials who are contemplating the use of this particular system of programmed instruction.

In summary, on the basis of evidence presented in this report, it can be concluded that: (a) there are measurable benefits in the

		Total Mathematics							Total Reading						
Group	N	Pre-test Mean	σ	Post-test Mean	σ	Mean Gain	t	p	Pre-test Mean	σ	Post-test Mean	σ	Mean Gain	t	p
FE	32	47.22	14.62	60.34	17.04	13.12			35.97	10.92	53.38	15.14	17.41		
FC_1	15	35.40	21.71	51.87	22.38	16.47	.91	N.S.	31.00	8.20	49.00	16.68	18.00	−.18	N.S.
FC_4	26	30.92	17.95	58.73	12.95	27.81	−5.32	.01	35.65	10.80	46.85	14.45	11.20	2.45	.05

Table 6. T-test Results for CAT Mean Gain Differences, Second Grade Follow-up of 1967-68 Subjects

form of increased achievement by disadvantaged first grade pupils resulting from the use of computer-assisted instruction in mathematics, and (b) some or all of these benefits may be lost if this method is discontinued after the first year.

References

Robert T. Filep. "What We Know So Far About Computer-Assisted Instruction." *Nation's Schools* 80: 62; October 1967.

Max Jerman and Patrick Suppes. "A Workshop on Computer-Assisted Instruction in Elementary Mathematics." *Arithmetic Teacher* 16: 193; March 1969.

Morris Norfleet and Leonard Burkett. "CAI in Eastern Kentucky." *Contemporary Education* 40: 290; April 1969.

Marian J. Patterson. "An Observation of Computer-Assisted Instruction on Underachieving Culturally Deprived Students." *Journal of Secondary Education* 44: 187; April 1969.

Patrick Suppes. "Accelerated Program in Elementary School Mathematics." *Psychology in the Schools* 13: 294; October 1966.

Patrick Suppes. "Computer-Assisted Instruction." *Nation's Schools* 82: 52; October 1968.

Pauline Sears and David H. Feldman. *Changes in Young Children's Classroom Behavior After a Year of Computer-Assisted Instruction: An Exploratory Study.* Stanford, California: Stanford University, Stanford Center for Research and Development in Teaching, Research Memorandum No. 31, May 1968. □