DOCUMENT RESUME

ED 084 871 EM 011 695

AUTHOR	Fletcher, J. D.; Suppes, Patrick
TITLE	Computer-Assisted Instruction in Mathematics and Language Arts for the Deaf. Final Report.
INSTITUTION	Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C. Div. of Research.
BUREAU NO	14-2280
PUB DATE	Aug 73
GRANT	OEG-0-70-4797-607
NOTE	45p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Algebra; Aurally Handicapped; *Computer Assisted Instruction; Deaf Children; *Deaf Education; *Deaf Research; Educational Research; *Elementary School Mathematics; English Instruction; *Language Arts; Logic

ABSTRACT

A three-year project supported research, development and evaluation of computer-assisted instruction (CAI) for hearing impaired, or deaf, children. Over 4,000 students from 15 schools for the deaf in five states participated in the effort. Although students received CAI in algebra, logic, computer programing and basic English, the skill subjects of elementary school mathematics and language arts were emphasized. Experimentation supported by the project ranged from practical evaluation studies of the specific curriculums presented to general, theoretical studies of the use of language by deaf students. The project demonstrated that CAI can significantly benefit deaf students, that CAI can support serious research in deaf education, and that CAI is economically practicable. A general aim of the project was to initiate large-scale use of CAI in schools for the deaf, and the available evidence indicates that this objective was successfully achieved. (Author)

ED 084871

Final Report

Project No. 14-2280
Grant No. OEG-0-70-4797-607

J. D. Fletcher and Patrick Suppes
Institute for Mathematical Studies in the Social Sciences
Stanford University
Stanford, California 94305

COMPUTER-ASSISTED INSTRUCTION IN MATHEMATICS AND LANGUAGE ARTS FOR THE DEAF

August 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

Bureau of Education for the Handicapped
(Division of Research)

SO 011695

ERIC

FILMED FROM BEST AVAILABLE COPY

COMPUTER-ASSISTED INSTRUCTION IN MATHEMATICS AND LANGUAGE ARTS FOR THE DEAF

J. D. Fletcher and P. Suppes

ABSTRACT

This project supported development, evaluation, and research of computer-assisted instruction (CAI) for hearing-impaired, or 'deaf', students. More than 4,000 students from 15 schools for the deaf in five different states participated during the three-year term of the project. Although students received CAI in algebra, logic, computer programming, and basic English, the skill subjects of elementary-school mathematics and language arts were emphasized by the project. Experimentation supported by the project ranged from practical evaluation studies of the specific curriculums presented to general, theoretical studies of the use of language by deaf students.

The project demonstrated that CAI can significantly benefit deaf students, that it can support serious research in deaf education, and that it is economically practicable. A general aim of the project was to initiate large-scale use of CAI in schools for the deaf, and available evidence indicates that this aim was successfully met.

ERIC Full Text Provided by ERIC

ED 084871

Final Report

Project No. 14-2280
Grant No. OEG-0-70-4797-607

Computer-assisted Instruction in Mathematics and Language Arts for the Deaf

J. D. Fletcher and Patrick Suppes
Stanford University
Stanford, California 94305

August 1973

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Education for the Handicapped
(Division of Research)

PREFACE

We acknowledge the essential work performed by Jamesine Friend in the initial stages of this project. Much of what was accomplished by the project depended directly on the foundation she created.

TABLE OF CONTENTS

		Page
Preface		i
List of Tables		iv
I.	Introduction	1
	A. The Stanford CAI System	1
	B. 1970-71	2
	C. 1971-72	3
	D. 1972-73	3
	E. Summary of CAI Curriculum	6
II.	Methods	13
	A. Experiment I: Mathematics Strands	13
	B. Experiment II: Mathematics Strands	15
	C. Experiment III: GP Measures	16
	D. Arithmetic Skills of Deaf and Hearing Students	17
	E. Problem Solving Experiment	17
	F. Language Arts Experiment	20
	G. Language Arts Item Analysis	21
	H. Test Development	22
	I. Economics and Technology of the Network	23
	J. Grammatical and Semantic Analysis of English Used by the Deaf	23
	K. Surveys	24
III.	Results	25
	A. Experiment I: Mathematics Strands	25
	B. Experiment II: Mathematics Strands	26
	C. Experiment III: GP Measures	27
	D. Arithmetic Skills of Deaf and Hearing Students	27
	E. Problem Solving Experiment	29
	F. Language Arts Experiment	29
	G. Language Arts Item Analysis	30
	H. Test Development	31

Page

I. Economics and Technology of the Network 31

J. Grammatical and Semantic Analysis of English Used by the Deaf 32

K. Surveys 33

IV. Publications 34

V. Conclusions 36

VI. Recommendations 37

VII. Bibliography

LIST OF TABLES

Table		Page
1	Participating Schools in 1971-72	4
2	Participat.ng Schools in 1972-73	5
3	Institute CAI Curriculums Used by Participating Schools for the Deaf, 1971-72	7
4	Institute CAI Curriculums Used by Participating Schools for the Deaf, 1972-73	8
5	Grade Level Spanned by Each Strand in the Elementary Mathematics Program	9
6	Matrix of Simple Correlation Coefficients for GP Scores Obtained by 44 Subjects on Posttreatment SAT, MSAT, and Mathematics Strands Measures	28

I. INTRODUCTION

This project supported research and development of computer-assisted instruction (CAI) for hearing-impaired, or 'deaf', students. The term of the project was July 1, 1970 to June 30, 1973. CAI curriculums developed by the Institute for Mathematical Studies in the Social Sciences (IMSSS) at Stanford University were used by more than 1,000 deaf students during the 1970-71 school year and by more than 2,000 deaf students during the 1971-72 and 1972-73 school years. The project generated significant interest and support among educators of the deaf throughout the country, and served as a major impetus for the installation of CAI in schools for the deaf across the country. General descriptions of the project were provided by Fletcher, Jamison, Searle, and Smith (1973) and by Suppes (1971, 1972a, 1972b).

A. The Stanford CAI System

The central processor for the Institute's computer system is a Digital Equipment Corporation PDP-10. In addition to 256K of core memory, short-term storage of programs and student information is provided by sixteen 180,000,000-bit disk modules. Long-term storage of student response data is provided by magnetic tape. About 280,000,000 bits of information can be stored by the system on one magnetic tape. Communication with remote student terminals in participating schools is provided by private telephone lines. High-speed data transmission (generally 2400 or 4800 baud) and time-division multiplexing are used to communicate with clusters of 16 or more student terminals. In 1972-73 more than 180 terminals were connected to the Institute system. About 125 of these terminals could be used simultaneously with no appreciable detriment in the system's speed of response. Any curriculum or other program could be run at any time on any student terminal.

The student terminals are 'KSR Model 33' teletypewriters. These teletypewriters communicate with the central computer system at a rate of about 10 characters per second (110 baud). They provide no audio, visual, or graphic capability, but their cost is about one-tenth of terminals that do. Despite their limitations, these inexpensive terminals permitted development of CAI that has produced dramatic gains in pedagogical achievement for hearing students as Suppes and Morningstar (1970, 1972), Fletcher and Atkinson (1972), and others have reported. For that matter, Jamison, Fletcher, Suppes, and Atkinson (1972) have argued that for cost-effectiveness, CAI, using satellite communication and teletypewriters, is a superior method for providing compensatory education.

In a typical school, there is one room containing 8 to 15 student terminals. One person, who is chosen by the school to be the CAI terminal proctor, is in charge of the equipment and of supervising students in the terminal room. Usually accompanied by their classroom teacher, the students enter and sit at any free terminal. The student starts the instruction by pressing the start key to signal that he is positioned at the terminal and is ready for attention. The program responds by typing

HI

PLEASE TYPE YOUR NUMBER AND NAME.

and the student responds accordingly.

Each student receives a unique number when he enrolls for CAI, so the request for the first name is merely an additional safeguard to ensure correct identification. A student can be, and usually is, enrolled for several available CAI courses. He uses the same number for all courses and types a one-letter identifier to indicate which course he is requesting. In the following example, the student is using S, the identifier for the elementary mathematics strands course. (Student responses are underlined in the example.)

HI

PLEASE TYPE YOUR NUMBER AND NAME.

S3456 MARY

The program scans the file of registered students, finds this student's last name, and types it. Unless he types special instructions, the student will then be placed in his sequence of lessons exactly where he left off.

HI

PLEASE TYPE YOUR NUMBER AND NAME.

S3456 MARY SMITH

JOB 10 ON TT5013 FRI FEB 2 73 8:46AM-PDT

HELLO MARY

HERE IS SESSION 46

16 - 9 = ___

B. 1970-71

The original proposal for this project was prepared in cooperation with three schools for the deaf and one school district that had already indicated great interest in having CAI curriculum for their deaf students. These were the California School for the Deaf in Berkeley, California; the Kendall School for the Deaf in Washington, D.C.; the San Jose Unified School District in San Jose, California, which has day classes for the deaf in four schools; and the Texas School for the Deaf in Austin, Texas.

Students at all of these schools regularly used Stanford CAI courses during the 1970-71 school year. In addition, the Model Secondary School for the Deaf, Washington, D.C., and the Palo Alto Unified School District, Palo Alto, California, used funding from other sources to participate in the project during 1970-71. The Palo Alto School District included the Stanford terminals as part of a larger experiment using CAI for deaf students. This project was documented by Jackson (1972).

During 1970-71 there were 60 terminals in operation at the participating schools. Over 1,000 deaf students used one or more of the Stanford CAI courses.

C. 1971-72

All of the schools that participated in the project during the 1970-71 school year continued during the 1971-72 school year, and several new schools joined the project. The schedule originally proposed that one new school with 15 terminals would be added during the 1971-72 school year. However, so many schools for the deaf expressed interest in CAI that it was offered on a 50-50 cost-sharing basis, thereby allowing more schools to join the network. Two residential schools, the Florida State School for the Deaf and the Blind, St. Augustine, Florida, and the Oklahoma School for the Deaf, Sulphur, Oklahoma, were added. Also, five new locations were added in Texas under the state's day school program. Houston Independent County-Wide Day School, Houston, Bexar County Day School for the Deaf, San Antonio, Dallas County Day School for the Deaf, Dallas, Tarrant County Day School for the Deaf, Fort Worth, and Beaumont Bi-County Wide Day School for the Deaf, Beaumont, all joined the network. The telephone line costs for the Texas county-wide day schools, the Texas School for the Deaf, and the Oklahoma School for the Deaf were reduced by creating a circuit that combined both time-division and frequency-division multiplexing in a single network.

All schools for the hearing impaired that participated in the project during the 1971-72 school year, including number of CAI terminals installed and percentage of financial support provided by the project, are listed in Table 1.

D. 1972-73

Despite a number of requests from schools across the country, no new project-supported schools were added during 1972-73. The emphasis during the year was on consolidating and evaluating the existing program and on studying the economics and technology of the CAI network. Participating schools, including number of CAI terminals installed and percentage of financial support provided by this project, are listed in Table 2.

Table 1

Participating Schools in 1971-72

School	Terminals	Support*
California School for the Deaf at Berkeley	15	100
Florida State School for the Deaf and the Blind	8	50
Kendall School for the Deaf	12	100
Model Secondary School for the Deaf	8	0
Oklahoma School for the Deaf	10	50
Palo Alto Unified School District (February 10, 1972-June 10, 1972)	2	0
San Jose Unified School District		
Bachrodt Elementary School	2	100
Hester Elementary School	2	100
Hoover Junior High School	1	100
San Jose High School	1	100
Texas County-Wide Day Schools		
Houston Independent County-Wide Day School	4	50
Bexar County Day School for the Deaf (San Antonio)	2	50
Dallas County Day School for the Deaf	3	50
Tarrant County Day School for the Deaf (Fort Worth)	2	50
Beaumont Bi-County Wide Day School for the Deaf	1	50
Texas School for the Deaf	15	100
Total	88	

*Percentage of financial support of CAI terminals provided by OE funds to IMSSS.

Table 2

Participating Schools in 1972-73

School	Terminals	Support*
California School for the Deaf at Berkeley	16	100
Florida State School for the Deaf and the Blind	8	50
Kendall School for the Deaf	12	100
Model Secondary School for the Deaf	5	0
Oklahoma School for the Deaf	10	50
San Jose Unified School District		
Hester Elementary School	3	100
San Jose High School	3	100
Texas County-Wide Day Schools		
Montrose School (Houston, Texas)	4	50
Bexas County Day School for the Deaf (San Antonio)	2	50
John B. Hood Junior High School (Dallas, Texas)	1	50
Skyline High School (Dallas, Texas)	2	50
Tarrant County Day School for the Deaf (Fort Worth)	2	50
Beaumont Bi-County Wide Day School for the Deaf	1	50
Texas School for the Deaf	16	100
Total	85	

*Percentage of financial support of CAI terminals provided by OE funds to IMSSS.

E. Summary of CAI Curriculums

All CAI curriculums developed by IMSSS were available to students in the participating schools for the deaf. Some of these curriculums, such as reading (grades K-3), French, and Russian, were inappropriate because they require audio. However, most IMSSS curriculums, even though not specifically designed for hearing-impaired students, were used successfully by the participating schools. Table 3 and Table 4 list all IMSSS curriculums used in this project with the numbers of hearing-impaired students enrolled for them in 1971-72 and in 1972-73, respectively. Brief descriptions of these curriculums follow. Those most relevant to this project were mathematics strands, arithmetic word problem solving, and language arts.

1. Elementary mathematics strands. The objectives of the curriculum were (1) to provide supplementary individualized instruction in elementary mathematics at a level of difficulty appropriate to each student's level of achievement, (2) to allow acceleration in any concept area in which a student demonstrates proficiency and repeated drill in areas of deficiency, and (3) to provide a daily profile report of each student's progress through the curriculum.

A strand is a series of problems of the same operational type (e.g., number concepts, addition, subtraction, fractions) arranged sequentially in equivalence classes according to their relative difficulty. The 14 strands in the program and the grade levels spanned by each strand are presented in Table 5.

A student in the strands program works on fewer than 14 strands; the actual number depends on his grade level and performance. The strands approach provides a high degree of individualization because each student's lesson is prepared for him daily by the computer, the lessons are presented as mixed drills at a level of difficulty in each strand determined by the student's prior performance, and the student moves up each strand at his own pace.

Details of the curriculum were documented by Suppes (1971), Suppes, Goldberg, Kanz, Searle, and Stauffer (1971), and by Suppes, Searle, and Lorton (1973).

2. Arithmetic word problem solving. This course gives students experience in solving arithmetic word problems. The course emphasizes methods of solution; the student constructs a well-formed algebraic expression, and the computer carries out the actual computation. The essential learning experience in solving arithmetic word problems is translating the English text into an algebraic expression. When such problems are presented in textbooks the student is often bogged down in computation and the correctness of his solution may be masked by careless errors in arithmetic. Having the computer do the computations allows the student to concentrate on the more fundamental aspect of expressing the problem as a well-formed algebraic expression.

Table 3

Institute CAI Curriculums Used by Participating

Schools for the Deaf, 1971-72

Curriculum	Number of students
Elementary Mathematics (Strands)	2146
Arithmetic Word Problem Solving	107
Language Arts	1071
Algebra	83
Basic English	165
Computer Programming in AID	93
Computer Programming in BASIC	124
Logic and Algebra	216
Total Students 2279	

Table 4

Institute CAI Curriculums Used by Participating Schools for the Deaf, 1972-73

Curriculum	Number of students
Elementary Mathematics (Strands)	1793
Arithmetic Word Problem Solving	520
Language Arts	1058
Basic English	64
Computer Programming in AID	34
Logic and Algebra	77
Total Students 2113	

Table 5

Grade Level Spanned by Each Strand in the Elementary Mathematics Program

Content	Grade level
Number Concepts	1.0-7.9
Horizontal Addition	1.0-3.9
Horizontal Subtraction	1.0-3.4
Vertical Addition	1.0-5.9
Vertical Subtraction	1.5-5.9
Equations	1.5-7.9
Measurement	1.5-7.9
Horizontal Multiplication	2.5-5.4
Laws of Arithmetic	3.0-7.9
Vertical Multiplication	3.5-7.9
Division	3.5-7.9
Fractions	3.5-7.9
Decimals	4.0-7.9
Negative Numbers	6.0-7.9

The student learns a few simple commands and uses them in a responsive dialogue with the instructional program to specify which computations are to be carried out. The student is free to experiment with the computer calculator made available to him, and he may choose any combination of steps to produce an answer. The program permits this freedom by storing a solution string for each problem and calculating the correct answer from the variables generated for the problem presentation by applying them to the solution string. The student's answer is evaluated only when he instructs the computer to do so. The text of the word problems is stored by the computer, and the numbers used in each problem are randomly generated for each presentation. Although all students see the same problem statement, each student receives a unique set of numbers from which he must construct an answer. Thus, dimensions of problem difficulty can be investigated independently of the numbers used in computing solutions.

The curriculum was described in more detail by Fletcher, Jamison, Searle, and Smith (1973), and by Searle, Lorton, and Suppes (1973).

3. Language arts. The language difficulties of hearing-impaired students were carefully considered in developing the Language Arts curriculum. The curriculum was designed to stress the structure of English, with particular emphasis on the roles of syntax and inflection and on the meaning of function words. An inductive rather than a deductive strategy is used. The course does not explicitly state 'rules' of English usage; it presents items illustrating aspects of standard English usage singly and in combination. Incidental learning of basic sentence patterns is enhanced by presenting curriculum items in complete sentences. Fewer than one-tenth of the exercises present the student with single words or isolated phrases. Incidental learning is also enhanced by requiring many constructed rather than multiple-choice responses.

There are four general course objectives. Students are to:

(1) Recognize specified grammatical categories;

(2) Recognize and supply various forms of given grammatical structures;

(3) Select appropriate grammatical units to complete a specified structure; and

(4) Perform specified transformations on grammatical structures.

The curriculum is divided into 218 lessons of 20-30 exercises. Separate topics are presented in separate lessons and often there is a sequence of lessons on a single topic. The lessons are ordered to provide a cumulative basis of concepts building upon one another. Several lessons are intended to review topics presented in preceding lessons.

The course was described in more detail by Fletcher and Beard (1973), by Fletcher, Jamison, Searle, and Smith (1973), and by Fletcher and Stauffer (1973).

4. Algebra. Algebra is a self-contained tutorial course for secondary school students. It teaches the basic properties of arithmetic operations, simplifying and solving equations, and the properties of inequalities. Each lesson contains instruction, printed homework, corrections of homework problems on the computer and a quiz.

5. Basic English. This course provides secondary school students with practice in problem areas of standard English usage. The objective of the course is to diagnose and correct the twelve most common usage errors: run-on sentences, sentence fragments, incorrect principal parts of the verb, confusion of adjectives and adverbs, lack of agreement between subject and verb, lack of agreement between pronoun and antecedent, incorrect case of pronouns, vague or indefinite pronomial reference, dangling elements, misplaced modifiers, errors in comparative forms of adjectives and adverbs, and double negatives. Further documentation was provided by Suppes, Goldberg, Kanz, Searle, and Stauffer (1971).

6. Computer Programming in AID. This is a self-contained, tutorial program that teaches the use of Algebraic Interpretive Dialogue (AID), a high-level algebraic language. The course was designed for junior college students and requires at least one year of algebra background. The course consists of 50 lessons, about one hour in length, plus summaries, reviews, tests, and extra-credit problems. Further documentation was provided by Friend (1973), Friend and Atkinson (1971), and by Suppes, Goldberg, Kanz, Searle, and Stauffer (1971).

7. Computer Programming in BASIC. This is a self-contained course for secondary students. It provides an introduction to programming for students without a knowledge of algebra. The course consists of 50 lessons, about one hour in length, plus summaries, reviews, and self tests. Further documentation was provided by Suppes, Goldberg, Kanz, Searle, and Stauffer (1971).

8. Logic and Algebra. This course is designed for secondary school students. The first year introduces numerical and sentential variables, formation of algebraic terms and sentences, and truth conditions of simple sentences. The second year of the course is concerned with the foundations of algebra. From a small set of axioms and rules of inference the properties of the field of rational numbers are developed. Further documentation was provided by Goldberg (1971), Goldberg and Suppes (1972), Suppes (1971), and by Suppes, Goldberg, Kanz, Searle, and Stauffer (1971).

9. Games. Additionally, there are several 'games' currently being used by all the participating schools for the deaf. Although they are called games and are entertaining to use, they have redeeming pedagogical value.

Bagels. This game is properly called Pico-Fermi-Bagels. The program creates a 3-digit number at random, and the student-player is to guess it. If one of the digits he guesses is correct but in the wrong position, he is told PICO. If one of the digits is correct and in the correct place, he is told FERMI. If none of the digits is correct, he is told BAGELS. He has 20 chances to guess the number.

Hangman. Hangman is the familiar game that most American children play in elementary school. In the IMSSS version, a definition or hint for a word is given, and the student-player is given six chances to guess the word by giving the letters that belong in it. The figure being hanged is 'drawn' by the teletypewriter. Three vocabulary lists are available for the game. Before beginning, the student-player must select the easy, medium, or hard list. Hangman is also available in a Spanish version.

Poster. Poster creates a poster by taking in any number of lines of text and enlarging the characters in each line to fill up the (8-inch) width of the teletypewriter paper.

Spell. In Spell, a word is typed and the student-player must indicate if it is spelled correctly or incorrectly. If the word is incorrect, the student-player must supply the correct spelling.

Spanish. In Spanish, a word is given in Spanish or English, and the student-player must translate it to English or Spanish, respectively. All instructions and hints are given in Spanish.

II. METHODS

A. Experiment I: Mathematics Strands

The purpose of the experiment was to measure the effect of varying numbers of math strands sessions on arithmetic computation grade placement (GP) measured by the strands curriculum and by an on-line, computer administered version of the Stanford Achievement Test (SAT) Arithmetic Computation subscale. This on-line version of the SAT was called the Modified SAT or MSAT. Construction and administration of the MSAT was detailed by Suppes, Fletcher, Zanotti, Lorton, and Searle (1973). Each student was allowed to take only a specified number of math sessions at the terminal. All other sign-ons were spent working language arts lessons.

Three hundred eighty-five students from among those who were taking both CAI math strands and CAI language arts, whose average GP on strands was between 2.4 and 5.9, and who had taken at least 15 math strands sessions, began the experiment. The students selected were assigned at random to five experimental groups that differed in the maximum number of math strands sessions they permitted during the experimental period of approximately 70 school days. Treatment groups 1, 2, 3, 4, 5 were assigned 10, 30, 70, 100, and 130 sessions, respectively.

Session limits were imposed on a calendar basis so that students with low numbers of sessions received them distributed throughout the experimental period. A participating student had no control over the type of lesson, math strands or language arts, he received. Whether he signed on for math strands or language arts a student was given a math strands lesson if he was eligible for one. Otherwise, he received a language arts lesson.

The number of math sessions a student received was monitored daily. The assistance of teachers and proctors was sought to help students achieve the number of math sessions they were assigned. Teachers were urged not to give compensatory off-line work to students assigned to low numbers of on-line sessions, and, in general, not to alter the classroom work of any student because of his participation in the experiment.

The MSAT was administered in January at the beginning of the experimental period and again in May immediately after the experiment ended.

One-way, fixed-effects analysis of variance was used to test for overall effect of the treatment groups. Additionally, it was important to investigate the relationship of posttreatment scores to pretreatment scores and the number of math strands sessions given. Five models of this relationship were tested. In all five models, T_{i1} denotes the pretreatment score of student i, T_{i2} denotes the posttreatment score of student i, and N_i denotes number of math strands sessions taken by student i.

Model I, Linear

$$E(T_{i2}) = a_0 + a_1T_{i1} + a_2N_i \ .$$

In this model, the effect of pretreatment score and number of sessions on posttreatment performance was assumed to be linear.

Model II, Linear with interaction.

$$E(T_{i2}) = a_0 + a_1T_{i1} + a_2N_i + a_3T_{i1}N_i \ .$$

In Model II, a linear effect of pretreatment score and number of sessions was assumed, but a linear effect from the interaction of pretreatment score and sessions was also postulated.

Model III, Cobb-Douglas.

$$E(\ell n\ T_{i2}) = a_0 + a_1\ \ell n\ T_{i1} + a_2\ \ell n\ N_i \ .$$

Model III was based on a formulation of the Cobb-Doublas type (from econometrics), namely,

$$T_{i2} = a_0T_{i1}^{a_1}\ N_i^{a_2} \ .$$

This model was multiplicative and assumed 'weighted interaction' in that a_1 and a_2 indicate the relative importance of pretreatment score and number of sessions, respectively, in accounting for change in posttreatment scores.

Model IV, Log quadratic.

$$E(T_{i2}) = a_0 + a_1T_{i1} + a_2\ \ell n\ N_i + a_3(\ell n\ N_i)^2 + a_4(\ell n\ N_i)^3 \ .$$

In Model IV, the effect of the pretreatment score was assumed to be linear. The effect of number of sessions was assumed to be logarithmic, rather than linear. In order to explore this logarithmic assumption fully, second- and third-order terms in $\ell n\ N_i$ were included.

Model V, Exponential.

$$E(\ell n\ T_{i2}) = a_0 + a_1N_iT_{i1} \ .$$

Model V was based on an exponential formulation, namely,

$$T_{i2} = a_0 e^{a_1 N_i T_{i1}} .$$

In this model, the effect of number of sessions and pretreatment score may be strictly increasing or strictly decreasing, depending on the sign of a_1. Pretreatment score and number of sessions were assumed to interact.

Additional details are provided in Suppes, Fletcher, Zanotti, Lorton, and Searle (1973).

B. Experiment II: Mathematics Strands

The purpose of this experiment was to demonstrate the utility of predictive-control integrated within CAI. In the experiment, a performance goal, defined in terms of GP, for progress over a predetermined time period was set for each student. At regular intervals during the experiment each student's performance history was examined to determine if he needed reassignment to more or fewer daily CAI sessions in order to reach his goal. At the end of the experiment, GP achieved by each student was compared with his goal.

More specifically, GP goals in elementary-school mathematics were set for students taking the mathematics strands CAI curriculum. These students were then assigned to one, two, or three daily sessions for each of six two-week periods. At the end of each period, each student's progress toward his goal was evaluated and reassignment of daily sessions was made when necessary.

Two measures of elementary-school mathematics GP were used: average GP achieved in the mathematics strands curriculum and GP achieved on the MSAT. Strands average GP was monitored daily, and the MSAT was given as a pre- and posttreatment measure.

Subjects for this experiment were chosen from the entire population of students who were enrolled in one of three residential schools for the deaf in California, Texas, and Florida and who were receiving daily CAI sessions in elementary mathematics from IMSSS in 1971-72. Three hundred fifty-five students from this population whose average GP was between 2.0 and 5.9, who had received more than 20 mathematics strands sessions, and who were not assigned to any other strands evaluation experiment participated as subjects.

Two performance goals were set for each subject. One goal was 'externally' derived, and one goal was 'internally' derived. Because the experiment period of 12 weeks was about one-third of a school year, the external goal for each student was defined as a gain of .33 in GP. The predictive-control aspects of the experiment did not apply to the external GP goal.

It should be noted that a GP gain of .33 over 12 weeks of school is an overly-optimistic projection for students from this population. Gentile and Di Francesca (1969) surveyed the academic test performance of hearing-impaired students. From their data it is clear that in this population an improvement in GP of .33 measured by the SAT Arithmetic Computation subtest is far more typical of an entire school year than of a 12-week period.

The internal GP goal was more individualized than the external GP goal in that it was uniquely determined from each student's performance history, and the predictive-control aspects of this experiment were applied to these goals. In setting the internal performance goals, average GP change per session was determined for each student by examining his 20th-40th strands sessions. Average GP change per session for any student who had not received 40 strands sessions was determined by examining his 20th to his latest CAI session. The internal GP goal for each student was then determined by extrapolation from these initial observations and from a linear model of his progress. Number of sessions taken and the internal GP goals were then used to integrate predictive-control techniques within the mathematics strands curriculum. Details of the method are given in Suppes, Fletcher, and Zanotti (1973).

C. Experiment III: GP Measures

Comparisons of the achievement measures used in this study with each other and with standardized tests are of natural interest. Generally, when an educator speaks of grade placement he has a standardized test in mind. Because neither the MSAT nor GP measured by the strands curriculum is a common measure, it was decided to estimate the concurrent validity of the MSAT GP and strands average GP by comparing them with each other and with paper and pencil administrations of the SAT.

Sixty subjects were drawn at random from among all subjects participating in Experiment I. Selection of the subjects was stratified so that four were chosen for each of the 15 MSAT form (Primary II, Intermediate I, Intermediate II) by treatment group (10, 40, 70, 100, 130 sessions) cells. Two of the four subjects were chosen at random and assigned to Group I; the remaining two were assigned to Group II. There were then 30 subjects (two from each form by treatment cell times 15 cells) assigned to Group I and 30 assigned to Group II. Group I received paper and pencil administration of the SAT Arithmetic Computation Subtest (SAT-COMP), Form W, before receiving the pretreatment MSAT, and Group II received the SAT-COMP after receiving the pretreatment MSAT. The roles of Group I and Group II were reversed for the posttreatment measure. Group II received the SAT-COMP before the posttreatment MSAT, and Group I received the SAT-COMP after the posttreatment MSAT. Pre- and posttreatment strands GP scores were also recorded for all the subjects. Additional details are given in Suppes, Fletcher, Zanotti, Lorton, and Searle (1973).

D. Arithmetic Skills of Deaf and Hearing Students

Comparisons of deaf and hearing students' performance were of natural interest in this project and these comparisons were made using performance data from the mathematics strands curriculum. This work was not undertaken in the context of an evaluation experiment, instead it drew on the extensive data base composed of information automatically recorded by the strands program during student sessions.

Most results from evaluation experiments reported here and elsewhere are analyzed using linear regression models. These models are adequate for many applications, but, however accurately they predict response probabilities, they do not postulate specific algorithmic processes that students use in solving problems. Finite automaton models are natural theoretical tools for describing algorithmic processes, but they have no place for a probabilistic theory of error, and a natural step is to use probabilistic automata in place of finite deterministic automata. However, automaton models do not reflect obvious commonalities that exist among the algorithmic tasks of arithmetic that students are asked to solve. For this reason, Suppes and Flannery (1973) used register machine models to compare the performances of deaf and hearing students in the mathematics strands curriculum.

The basic data for this work were the mean percentages correct for each equivalence class in the strands curriculum. It should be noted that before a student can reach a given equivalence class in a given strand he must master the previous equivalence class leading up to it independent of such external factors as grade in school, chronological age, hearing loss, and minority group status. In a genuine sense, the strands performance data provide a very broad basis for comparing different groups of students with common preparation and prior performance. These comparisons were discussed in detail by Suppes and Flannery (1973).

E. Problem Solving Experiment

The purpose of the work with arithmetic word problems was to achieve optimization of learning rates in the context of individual differences. There were three distinguishable aspects to this effort: identification of appropriate dimensions of word problem difficulty, identification of optimal error rates, and development of a model of problem difficulty for individual students.

Considering group data only and letting $\bar{p}_i$ be the observed proportion of correct responses on problem i for a population of students, a simple linear model of the form,

$$\bar{p}_i = \sum_{j=1}^{k} a_j x_{ij} + a_0 ,$$

can predict proportions of correct responses to new problems from the a_j weightings already estimated for the entire student population and assigned to the k dimensions of problem difficulty. These models were used in the following steps:

Step 1. A pilot study was conducted using a set of 65 problems. These were presented to students at a terminal, using response formats somewhat different from those of the problem solving course, but with the same constraints; that is, the computer carried out the calculations.

Step 2. The results of this study were examined and a set of variables thought to relate to problem difficulty were defined. A stepwise multiple regression analysis was used to determine the contribution of each of these variables to the multiple regression coefficient. Five variables were found to account for 60 percent of the variability, and the contribution of each remaining variable was less than 1 percent. These variables were, in the order in which they entered the regression: OPERS, the minimum number of arithmetic operations required for solution; CONVR, whether the solution required conversion of units with the unit of conversion absent from the problem statement; LENGT, number of words in the problem statement; DIV, whether the solution required a division; and VCLUE, whether the problem contained a verbal clue for an operation.

Step 3. A set of problems was written and edited, and coded using the five variables identified in the pilot study.

Step 4. A multiple regression analysis was carried out on the pilot study data, using only these five variables. The regression coefficients obtained were used to predict the probability correct for the 700 problems written for the main curriculum, and the problems were ordered according to predicted probability correct. The predicted probability correct ranged from .95 to .07.

The 1971-72 performance data were used to repeat the entire process of identifying variables that contribute to the multiple regression coefficient, predicting probabilities of correct answers, and arranging the problems in order of increasing difficulty. This analysis was performed for deaf and hearing students separately in order to compare the separate orderings of problems and weightings of the dimensions of difficulty.

Assuming that the problems are appropriately ordered in difficulty for any student, we can position him in the curriculum so that the proportion of correct responses he makes remains fairly stable and so that his progress in the curriculum is optimal. The following constant error rate model from Suppes (1967) stabilizes the proportion of correct responses about the value assigned to r, which is assumed to be the optimal rate of correct responding for a student.

$$x_{n+1} = \begin{cases} x_n + \epsilon & \text{if } p_s(c) > r_s, \quad x_n + \epsilon \leq 1 \\ x_n & \text{if } p_s(c) > r_s, \quad 1 < x_n + \epsilon \\ x_n & \text{if } p_s(c) \leq r_s, \quad x_n < \epsilon \\ x_n - \epsilon & \text{if } p_s(c) \leq r_s, \quad \epsilon \leq x_n \end{cases}$$

where x_n is the position of student s in the curriculum,

ϵ is a small, positive constant used to define the step size of student s' walk through the curriculum,

$p_s(c)$ is the proportion of correct answers achieved by student s over some previously used portion of the curriculum,

r_s is the optimum rate of correct answers for student s.

Although discussed earlier by Suppes (1967) many will recognize aspects of 'tailored testing' as discussed by Lord (1971a, 1971b) in this procedure. Lord, however, uses Birnbaum's 3-parameter logistic model of item difficulty which is strictly empirically derived and, therefore, cannot be applied to predicting difficulty of new problems that have not been attempted by a large number of students.

The comparison of dimensions that contribute significantly to models of problem difficulty for deaf and hearing students as well as the comparison of weightings given these dimensions is of natural interest. Also of interest are the weightings given these dimensions in models of problem difficulty derived for individual students.

It should be noted that because we have to find one set of a_j to apply across all problems for each individual, we can determine the relative importance of the k a_j variables for each individual from the magnitude of these values. These values may not vary in any significant manner from those developed for the population. In this case the calculation of parameters for models of problem difficulty for individuals is a waste of time. This question can be investigated by testing the null hypothesis that a linear relation holds between the set of a_j calculated for any individual student and the a_j calculated for the entire population of students. Specifically, we can test the null hypothesis that there exists an $\alpha_s \neq 0$ and β_s such that the k-tuple of parameters calculated for the model of problem difficulty over the entire population

$$(a_0, a_1, \ldots, a_k) = \alpha_s(a_{0s}, a_{1s}, \ldots, a_{ks}) + \beta_s .$$

Changes of the a_{js} over time (or learning) and different values of r will also be of interest.

The work with arithmetic word problems is documented in Searle, Lorton, and Suppes (1973).

F. Language Arts Experiment

The purpose of the experiment was to measure the effect of varying numbers of language arts sessions on posttest scores. Each student was allowed to take only a specified number of language arts sessions. All other sign-ons were spent working math strands sessions. The experiment was analogous to the math strands Experiment I described by Suppes, Fletcher, Zanotti, Lorton, and Searle (1973).

Two hundred thirty students from among those who were taking both CAI math strands and CAI language arts in 1972-73 were selected for the experiment, and were assigned at random to one of five experimental groups that differed in the maximum number of 10-minute language arts sessions they permitted. Students assigned to Groups 1, 2, 3, 4, and 5 were permitted 20, 45, 70, 95, and 120 sessions, respectively. The subjects were selected from students in the California School for the Deaf, Berkeley, California, the Oklahoma School for the Deaf, Sulphur, Oklahoma, and the Texas School for the Deaf, Austin, Texas. Random assignment of these subjects to the five treatment groups was stratified so that roughly the same number of students from each school could be assigned to each of the treatment groups. When the experiment began, 45 students were assigned to Group 1, 46 were assigned to Group 2, 46 were assigned to Group 3, 47 were assigned to Group 4, and 46 were assigned to Group 5. One-way, fixed effects analysis of variance and five models of student progress were used to investigate student performance at the end of the 80 school-day experiment period. The five models of student progress investigated were the same as those used in mathematics strands Experiment I.

The assistance of teachers and proctors was sought to help students achieve the number of language arts sessions they were assigned. Teachers were urged not to give compensatory off-line work to those students assigned to low numbers of on-line sessions, and, in general, not to alter the classroom work of any student because of his participation in the experiment.

The language arts experiment was documented by Fletcher and Beard (1973).

G. Language Arts Item Analysis

The intent of this analysis was to discover useful dimensions of difficulty that affected performance on language arts items taken by deaf students. Three different item classifications were used. Items were classified by subdivisions of the four course objectives listed earlier, by the required exercise task, and by the format of the correct answer.

There were four dimensions of classification by exercise task.

(a) Instructions given or no instructions given. This dimension distinguishes exercises that occur early in lessons for which the instructions are printed or repeated, from exercises that occur later in lessons when it was assumed the student had them well in mind.

(b) Instance (number) or instance (text) or concept. This dimension distinguishes exercises in which the student must answer with an instance of a concept from exercises in which the student must answer with a concept based on a given instance. When concepts are answers they are always abbreviated. Some instances were numbered so the student could reply only with the number(s) associated with the text of the instance--instance (number)--rather than with the actual text of the instances--instance (text).

(c) Recognition or construction (explicit basis) or construction (implicit basis). This dimension distinguishes exercises in which the answer is printed in the exercise display--recognition--from exercises in which the answer does not appear in the display--construction. The construction (explicit basis) and construction (implicit basis) dimensions distinguish between degrees of explicitness in the exercise directions. In construction (explicit basis) a form, but not the correct form, of the correct answer text is given; in construction (implicit basis) no form of the correct answer is given explicitly.

(d) Usage or definition. This dimension distinguishes exercises in which the answer is to be derived on the basis of an implicit rule of usage taught inductively in the curriculum from exercises in which the answer is to be derived from the definition of a grammatical category.

Given 2 times 3 times 3 times 2 possibilities, there would be 36 categories under this task classification scheme if it were not for the following combinations that do not occur:

there are no concept-construction tasks;
there are no concept-usage tasks;
and there are no instance (number)-construction tasks.

Eighteen categories are left plus one category labeled "Giveaway."

There were two dimensions of classification based on format of correct answers.

(a) Word or letter or number or abbreviation. There is some 'nesting' under this dimension: word is subclassified as 1, 2, 3, or 4 word strings; letter is classified as 1, 2, or 3 letter strings; and number is classified as 1, 2, 3, 4, or 5 number strings. Abbreviations present a problem in that they could reasonably be classified as single letters, multiple letters, or single words. It was decided that they would confuse the single letter, multiple letter, or single word results, and they were treated separately in the exercise task data analyses.

(b) Sequence or no sequence. In some instances, the sequence of a multiple word, multiple letter, or multiple number response is important; in some instances sequence is not important. This dimension distinguishes between these instances.

Classified in this way there were 17 correct answer formats that occurred in the language arts curriculum.

A more detailed description of the language arts item analysis is documented in Fletcher and Beard (1973).

H. Test Development

Two tests were developed specifically for this project: the computer-administered MSAT, and a language arts test (LAT). Construction and validation of the MSAT are discussed in Suppes, Fletcher, Zanotti, Lorton, and Searle (1973), and construction and validation of the LAT are discussed in Fletcher and Beard (1973). These practical developments evolved naturally into some theoretical work on test development.

Jensema (1973a) studied a simple method for estimating parameters for the Birnbaum 3-parameter logistic mental-test model. The accuracy of the method was investigated using Monte-Carlo data, and data from six vocabulary tests were then used to demonstrate the usefulness of the method in prescreening items. Jensema (1973b) also used Monte-Carlo data and four different item banks to study termination of Bayesian tailored testing by two different methods: according to the number of items administered and according to the magnitude of the standard error of estimate. The estimate of ability obtained through tailoring was compared with the known ability of each Monte-Carlo 'examinee' as each standard error of estimate level was reached and as each item was administered.

I. Economics and Technology of the Network

The primary effort under this project was to develop and evaluate CAI curriculums used by deaf students. However, development and evaluation are insufficient in themselves to have a practical payoff for substantial numbers of students. For this reason, we investigated the operational implementation of CAI. A central aspect of this implementation concerned the basic economics of CAI--its cost, performance, and degree of substitutability for other inputs into education of the deaf.

The basic cost assumption was that for $300 per month a teletypewriter terminal can be maintained in a typical school. This cost includes amortization of capital costs, use of the central computer system, communications, proctoring, and supplies. It does not include expenditures for classroom space. A further assumption was that for 20 days per month an average of 25 student sessions per day are given at each student terminal. Thus, 500 sessions per terminal per month were assumed to cost $300, or $.60 per session. The number of sessions per terminal per day obtained by different schools varies widely, and with effective scheduling it is feasible to obtain many more than 25 sessions per terminal per day. Many schools for the deaf obtained utilization rates in the range of 35-40 sessions per terminal per day, suggesting the possibility of substantially lower costs per session than $.60. Also, a 6-hour school day was assumed; the residential schools for the deaf used their terminals 8-10 hours per day, further increasing the number of sessions per terminal per day and further decreasing the cost per session.

The decision of whether to provide CAI and how much CAI to provide depends not only on cost per session but on two other critical factors. First is the performance of CAI in raising student achievement. Second is the issue of what must be given up in order to have CAI. Given that budgets are inevitably constrained, the more CAI an administrator provides his students, the less he can provide of something else. A requirement of good administration is to make these trade-offs explicitly, both in terms of their cost and of their performance.

The study of the economics and technology of the CAI network maintained for the schools that participated in this project was described in more detail by Fletcher, Jamison, Searle, and Smith (1973), and by Ball and Jamison (1972).

J. Grammatical and Semantic Analysis of English Used by the Deaf

There were three aspects of this work: collection and analysis of a corpus of writing samples; study of the manual alphabet; study of prelingually deaf adolescents as non-native users of English.

An important step in examining the language of hearing-impaired students was the collection and analysis of a corpus of writing samples. An intensive analysis of a small sample of the written language of deaf students obtained from Kendall School for the Deaf in Washington, D.C.

and the California School for the Deaf in Berkeley, California, was undertaken. This corpus was edited and divided into 1,311 sentences. Informally, these sentences were analogous to the complete thoughts of classical grammar. Nearly all of the sentences in the writing samples were terminated by conventional punctuation so the original corpus was modified very little. This corpus was analyzed with respect to sentence length, vocabulary, and the grammar of its noun phrases. This work was detailed by Fletcher, Jamison, Searle, and Smith (1973).

Perceptual confusions in learning and perceiving letters of the manual alphabet were studied in depth. Letters of the manual alphabet were displayed as characters on a computer-graphics terminal, and perceptual confusions among these characters were investigated by close examination of the response profiles recorded for subjects learning to 'read' the manual alphabet and by multidimensional scaling techniques applied to perceptual confusions that arose when the letters of the manual alphabet were rapidly displayed. It should be noted that the precise control over display of these characters and the comprehensive and accurate recording of subject response data permitted by computer-graphics presentations established a unique experimental situation for studying perception of the manual alphabet. This work is documented by Weyer (1973).

One explanation for the difficulties that deaf students experience with standard English is that they learn English as a second language and that American Sign Language is their native language. This explanation has occurred recently in the literature of deaf education as a speculation, but systematic, empirical studies of this explanation were lacking. Pursuant to this regard the Test of English as a Foreign Language (TOEFL) published by Educational Testing Service (1970) was administered to 13 prelingually deaf children of hearing parents (HP) and to 13 prelingually deaf children of deaf parents (DP). Mean age in the HP group was 219 months and mean age in the DP group was 206 months. The TOEFL scores of the HP and DP subjects were compared with each other and with the scores of the foreign student population used to standardize the TOEFL. Scores achieved by the HP and DP subjects on the Paragraph Meaning and Language subtests of the SAT were also compared. This work was described by Charrow and Fletcher (1973).

K. Surveys

Three extensive surveys of the cognitive literature on deafness were completed as an essential adjunct to the project. Suppes (1972b) surveyed the cognition of blind, deaf, and educable mentally retarded children in three areas: language and language development; concept formation and abstraction; and elementary mathematical skills. Bonvillian and Charrow (1972) reviewed the psycholinguistic implications of deafness with particular emphasis on language acquisition and the use of sign language by the deaf. The Bonvillian and Charrow review was considerably expanded by Bonvillian, Charrow, and Nelson (1973) to include careful review of educational achievements of the deaf and available educational programs for the deaf with particular emphasis on the relative effectiveness of oral and manual communication techniques.

III. RESULTS

A. Experiment I: Mathematics Strands

Suppes, Fletcher, Zanotti, Lorton, and Searle (1973) reported that complete data were obtained for 60 students in group 1, 62 students in group 2, 60 students in group 3, 60 students in group 4, and 70 students in group 5. The number of sessions taken fell short of the number assigned in groups 3, 4, and 5, primarily because of difficulties in scheduling extra CAI sessions in the schools. However, the groups remained sufficiently distinct to warrant proceeding with analysis of variance which used posttreatment MSAT scores and average GP of the mathematics strands as dependent measures. In order to make comparisons across all three MSAT battery scores, SAT scales were used to convert MSAT raw scores to GP scores. Analyses of variance were performed on pretreatment measures as well as posttreatment measures to check for any bias in the assignment of students to treatment groups. The F ratio of 9.088 for the strands posttreatment GP was significant ($F_{.99}(4,307) = 3.48$), and the average GP improvement for the 10-sessions group 1 was only .15 compared with .96 for group 5. The F ratio of 1.404 for the MSAT scores was nonsignificant ($F_{.95}(4,307) = 2.45$), but the average GP improvement for the 10-sessions group 1 was .42 compared with .76 for group 5.

Parameters for the five models were generated twice, once using mathematics strands average GP as pretreatment and posttreatment achievement measures and once using MSAT GP scores. The linear model with interaction accounted for more of the variance in the dependent variable (posttreatment average GP) than did any of the other models, but despite the inclusion of a term for the interaction of number of sessions with pretreatment GP, it represented only a slight improvement over the simple linear model. Assuming $N_i = 120$ or slightly less than one session per day for a school year and taking $a_2 = .0123$ from the linear model, we can project $T_{i2} - T_{i1} = 1.48$. That is to say, if a student from this population takes about one strands session per day for an entire school year, we can expect his strands average GP to increase by about a year and a half. Data presented later show that strands average GP underestimated both GP measured by paper and pencil administrations of the SAT and GP measured by the MSAT. This improvement of 1.48 can be compared with an expected GP increase over a school year of .3 to .4 in the SAT computation subtest for hearing-impaired students receiving ordinary instruction (Gentile & DiFrancesca, 1969).

Among the models and parameters using MSAT GP as pretreatment and posttreatment measures, the multiplicative model from econometrics that assumed weighted interaction of number of sessions with pretreatment GP accounted for more of the variance in the posttreatment measure than did any other model, but, as with strands average GP, it represented only a slight improvement over Model I, the simple linear model. Again, assuming $N_i = 120$ and taking $a_2 = .0084$ from the linear model, we can project $T_{i2} - T_{i1} = 1.01$. That is to say, if a student from this population takes about one strands session per day for a school year of 120 net days, we can expect his MSAT GP to increase by about one year. Roughly, we can expect an increase of .1 in MSAT GP for every 12 sessions taken.

Suppes et al. concluded that the mathematics strands CAI curriculum can lead to substantial increases in mathematics computation GP when used by hearing-impaired students. The increases are sufficient to bring the students to GP gains expected of normal-hearing students. Moreover, these gains can be achieved by students working intensely for only a few minutes a day in a supplementary drill-and-practice program. The time spent at a computer terminal by each student ranged from 6 to 10 minutes for each session.

In addition, Suppes et al. concluded that a simple linear model of student achievement gives a good account of the posttreatment distribution of GP measured either by the MSAT or by the strands GP. The investigation of other models, including models with interaction terms, did not lead to any substantial improvement in accounting for posttreatment GP variance. The results of the analysis, including the application of the linear model, indicate that greater numbers of CAI sessions are beneficial for all students, across all levels of pretreatment achievement.

B. Experiment II: Mathematics Strands

Suppes, Fletcher, and Zanotti (1973) reported that complete data were obtained for 297 subjects. The subjects took far fewer mathematics strands CAI sessions than expected. In designing the experiment, it was assumed that each subject could take as many as 150 sessions in the 12-week experiment period and that each subject would take a minimum of 50 sessions. As it turned out, 159 subjects (54%) took less than 50 sessions. Despite these low numbers of accumulated sessions, 274 subjects (92%) exceeded their external goals measured by strands average GP and 191 subjects (64%) exceeded their external goals measured by the MSAT GP.

The biweekly predictions in the experiment were made in terms of a linear model because the extensive analysis required to fit a more sophisticated model had not taken place. Suppes, Fletcher, and Zanotti reported that models of the following form predicted student progress in the experiment with a high degree of precision:

$$GP_i = a_i + b_i S_i^c , \qquad c \leq 1 ,$$

where GP_i is the grade placement of student i,

S_i is the number of sessions taken by student i,

a_i and b_i are unique parameters of the model calculated for student i.

The two instances from these models used were:

$$GP_i = a_i + b_i S_i^{1/3} \qquad (1)$$

$$GP_i = a_i + b_i \ell n(S_i) \qquad (2)$$

The precision of these models indicated the great promise of predictive-control integrated within CAI curriculums. The standard errors of estimate for 90% of the subjects ranged from .013 to .114 for (1) and from .016 to .131 for (2).

Suppes, et al. emphasized the two aspects of individualization achieved by their approach. Even though the amount of CAI time given to an individual student may be highly individualized, the goal originally set for him can be totally unreasonable if it, too, is not tailored to a model of his progress. Using the approach developed in this experiment precise individualization of instruction can be achieved both in the amount of instruction required and in the goal set for each student.

C. Experiment III: GP Measures

Complete data were obtained for 44 students. The loss of 16 subjects was solely due to such random factors as student illness, change of schools, and administrative errors. Suppes, Fletcher, Zanotti, Lorton, and Searle (1973) noted that the SAT consistently gave the highest estimate of GP for this group of students, the MSAT consistently gave the second highest GP estimate, and the strands GP consistently gave the lowest GP estimate. Evidently, both the MSAT and the strands average GP measures underestimated GP measured by paper and pencil administration of the SAT. A matrix of simple correlations for the GP scores obtained by the 44 students on posttreatment SAT, MSAT, and mathematics strands is given in Table 6. These correlations are fairly large, but they are not sufficiently large to identify SAT GP, MSAT GP, and strands GP as parallel measures as defined by Lord and Novick (1968).

D. Arithmetic Skills of Deaf and Hearing Students

Two major conclusions were reported by Suppes and Flannery (1973). The first was that objective features of the curriculum, for example, whether a vertical addition problem has a carry or not, dominated the ease or difficulty of exercises in much the same way for both deaf and

Table 6

Matrix of Simple Correlation Coefficients for GP Scores Obtained by 44 Subjects on Posttreatment SAT, MSAT, and Mathematics Strands Measures

	SAT post-treatment	MSAT post-treatment	Strands post-treatment
SAT post-treatment	1.000	.827	.794
MSAT post-treatment		1.000	.807
Strands post-treatment			1.000

normal-hearing children. This finding leads to the second conclusion, which was more surprising than the first: the performance of the deaf children was almost always slightly better than that of the normal-hearing children. More exactly, of the 761 equivalence classes, summing across all grades and strands for which there were data, the mean percentage correct of the deaf students was higher than that of the normal-hearing students for 673 equivalence classes, and it was the same to two decimals for 22 equivalence classes. These massive data support the thesis that the cognitive performance of deaf children is as good as that of normal-hearing children, when the cognitive task does not directly involve verbal skills. From an educational standpoint, the data suggest that with proper organization of teaching effort, we should be able to obtain results in arithmetic as good for deaf children as we do for average to slightly below-average normal-hearing children.

E. Problem Solving Experiment

Searle, Lorton, and Suppes (1973) were able to account for 72% of the variance in observed proportions of correct answers to the 125 arithmetic word problems worked by their subjects. Using seven independent variables and proportion correct as the dependent variable, they obtained a multiple correlation coefficient of .85 with a standard error of estimate of .27. They concluded that it is possible to account for a substantial portion of variability in student responses to arithmetic word problems using independent variables that describe structural features of the problems. However, their models of problem difficulty were inadequate in that both the predicted proportions of correct answers and the expected amount contributed to problem difficulty by each independent variable lacked precision. On this basis, Searle, Lorton, and Suppes further concluded that their results were situation dependent.

Comparisons of performance by deaf and disadvantaged hearing students showed that both groups found the same problems easy or hard. There was a significant correlation between the rank-order of problems for the two groups (Kendall's rho = .511, $p < .001$), and Searle, et al. concluded that the different handicaps characterizing the two groups of subjects did not produce different performance on the word problems.

F. Language Arts Experiment

Fletcher and Beard (1973) reported that complete data were obtained for 197 subjects. However, 46 of these subjects had received 100 or more sessions in 1971-72 and these subjects were removed from the experiment prior to any data analyses which were then performed on the 151 remaining subjects. In the analysis of variance there were 33, 27, 26, 33, and 32 subjects in treatment groups 1, 2, 3, 4, and 5, respectively. Students in groups 1, 2, 3, 4, and 5 received an average of 22, 46, 69, 88, and 106 sessions, respectively. These averages were less than expected for groups 3, 4, and 5, but the treatment groups appeared sufficiently distinct to proceed with analysis of variance. The F-ratio for this analysis was not statistically significant, indicating that the range of sessions considered did not have a significant effect on posttest scores. The paper and pencil language arts test developed by the project appeared to be reliable and fairly valid. The correlation between pre- and posttest scores on the test was .910 with an F-ratio for significance of regression beyond $p < .01$, and the correlation between posttest scores and number of lessons completed was .645 with an F-ratio for significance of regression beyond $p < .01$.

Models I, II, III, IV, and V accounted for 83%, 83%, 66%, 83%, and 33%, respectively, of posttest score variance. The only model to which a term that included a measure of sessions taken contributed significantly was Model V. In all other models the only significant independent variable was the pretest score. An additional model, Model VI, was investigated. This model was of the form

$$E(T_2) = a_0 + a_1T_1 + a_2N + a_3L$$

where T_2 refers to posttest score,

T_1 refers to pretest score,

N refers to number of sessions taken,

L refers to number of lessons completed,

and a_0, a_1, a_2, and a_3 are parameters of the model. Model VI accounted for 85% of the variance in posttest scores. Both sessions and lessons, in addition to pretest scores, contributed significantly ($p < .01$) to the model. However, the regression coefficient in Model VI for number of sessions taken was negative, indicating an inverse relationship between number of sessions taken and posttest scores when number of lessons completed is taken into account.

Fletcher and Beard concluded that the course is of significant value to students whose ratio of lessons completed to sessions taken is high but of much less value to students whose ratio of lessons completed to sessions taken is low. The relationship between sessions taken and posttest scores was concluded to be more complex than anticipated.

G. Language Arts Item Analysis

Fletcher and Beard (1973) reported several results from their item analysis of the language arts curriculum that are not widely noted in the literature on deafness.

First, the 'directions' lessons were far easier than anticipated given the general impression among deaf educators that deaf students experience difficulty in following directions. Some reasons for this result may be that the directions in these lessons and in the curriculum were easier to follow than those given in classroom instruction, that the directions given in the language arts CAI were more clearly communicated to students than are directions given in classroom instruction, and that deaf students have less difficulty following directions than generally supposed. More research is required to decide among these alternatives.

Second, although pronouns were generally far easier than anticipated, items on possessive pronouns were extremely difficult for the students. Specifically, possessive pronouns that differed in number (his boxes, their box) and/or gender (his sister, her husband) from the nouns they modified were seldom completed correctly.

Third, copulas joining subjects with predicate complements that differed in number from their subjects were very difficult for the students. Copulas for items such as the following:

> The house (is, are) blue and white.
> The girls (seem, seems) lonely.

were seldom completed correctly.

Fourth, the students had very little trouble with contractions with the exception of 'I'm' which was far more difficult for the students than anticipated.

H. Test Development

Jensema (1973a) showed that estimation of guessing parameters for Birnbaum's 3-parameter model must be improved before tailored testing techniques with multiple-choice items can become generally useful. The Monte-Carlo data demonstrated that estimation of item difficulty by the method discussed is accurate, but that estimation of item discrimination can be relied on only for items with difficulties between -1.0 and +1.0. Jensema concluded that the method discussed should be used with caution and must be viewed not as a replacement for maximum likelihood estimation, but as a convenient technique for economically prescreening items.

Jensema (1973b) showed that the standard error of estimate was a good index of reliability regardless of the characteristics of the items presented. He concluded that, generally, Bayesian tailored testing is best terminated by reaching a criterion value for the standard error of estimate. However, if all the items presented have approximately equal discrimination and guessing parameters, the number of items that must be administered to achieve a given level of reliability can be estimated.

I. Economics and Technology of the Network

Fletcher, Jamison, Searle, and Smith (1973) concluded that substantial amounts of CAI are feasible with only minor increases in student-to-staff ratios. The school administrator must determine the cross-over that results from the increasing achievement due to CAI that is counterbalanced by the decreasing achievement due to larger student-to-staff ratios. Ball and Jamison (1972) demonstrated the economic viability of a satellite-based CAI communications network for dispersed populations. Four implementation alternatives were considered by Fletcher, et al. and by Ball and Jamison: first, operational utilization of the IMSSS facility at Stanford with administrative and operational responsibilities borne by Stanford personnel; second, operational utilization of the IMSSS facility with administrative and operational responsibilities borne by a school serving the deaf community; third, implementation of curriculums developed under this project on small stand-alone computer systems located at school sites; and fourth, establishment of a large CAI center for the deaf that would run 500-1500 student terminals simultaneously. Ball and Jamison indicated that there are no absolute rules for choosing among these alternatives but that choices must be made relative to trade-offs that were explicated in their paper.

J. Grammatical and Semantic Analysis of English Used by the Deaf

Fletcher, Jamison, Searle, and Smith (1973) reported an average sentence length of about nine words in the SAMPLE corpus. An interesting theoretical problem connected with the length of utterances in the corpus was to account for the distribution of utterance lengths with a formal model of utterance generation. The three theoretical distributions investigated were the geometric, the poisson, and the negative binomial. Of these, the negative binomial provided the best fit of the data. There were 11,697 words (tokens) in SAMPLE, but only 1,898 different words (types). Fletcher, et al. listed the 100 most frequently occurring words and the grammatical categories appropriate for all entries in the SAMPLE lexicon. Noun-phrases predominated the utterances in SAMPLE and a probabilistic grammar was constructed for the 2,366 noun-phrases in the corpus.

Weyer (1973) reported major clusters of perceptually confused characters in the manual alphabet. The largest cluster (A N S T) was composed of letters that involve a fist and differ from one another in thumb position only. Other letters represented by folded fingers (E M O) composed an adjacent cluster. Letters in the next cluster (B F U) were distinguished by the number of vertical fingers--4, 3, and 2, respectively--displayed. Minor confusion clusters composed of K and V, V and W, R, D, X, and Z, G and H, I and J, and P and Q were also reported. There were practically no confusions involving C, L, and Y. Weyer's study represented the first application of multiple dimensional scaling to the manual alphabet.

Charrow and Fletcher (1973) reported statistically superior performance by DP subjects on three of the four TOEFL subtests (HP and DP subjects' scores on the Reading Comprehension subtest did not differ significantly), on total TOEFL scores, and on the Paragraph Meaning and Language SAT subtests. Parentage (whether the subjects had deaf or hearing parents) accounted for 53% of the variance in total TOEFL score and the indicated importance of parentage was corroborated by stepwise multiple regression. Item-by-item comparisons within the TOEFL subtests for number of responses to the correct answer and to the most likely wrong answer made by a standardization group of hearing, foreign students were undertaken. These comparisons showed the TOEFL performance of DP subjects to be more like that of the standardization group than was the TOEFL performance of HP subjects. Charrow and Fletcher concluded that their results suggest that deaf students learn English as a second language, but that more sensitive measures must be devised to provide more conclusive results.

K. Surveys

Suppes (1972b) drew three broad conclusions from his survey of research on cognition in deaf children. First, language problems are central to the education of handicapped children, but much remains to be learned about the source of their difficulties and about how they may be met. Second, the excellent methodology developed for the learning of discriminations and of simple associations by the handicapped needs to be extended to more complex tasks. Third, more attention needs to be given to estimating the magnitudes of the effects that result from various training procedures and less attention to establishing statistically significant differences among training procedures.

Bonvillian, Charrow, and Nelson (1973) concluded that: first, the deaf are not deficient in intellectual competence, their weaker skills in English and lower educational achievement require other explanations; second, despite deficiencies in processing English, many deaf persons communicate effectively in sign language; and, third, similar linguistic abilities underlie effective use of sign language and spoken language. Bonvillian, et al. strongly recommend increased use of sign language in educating the deaf.

IV. PUBLICATIONS

This project generated a substantial amount of documentation. Much of this documentation took the form of IMSSS technical reports, and these reports will be or already have been submitted either in whole or in part for publication in the professional literature. The purpose of this section is to review publications arising from the project with reference to topics discussed in this report.

General

General discussions of the standards, objectives, and methods that formed a foundation for this project were given by Fletcher and Stauffer (1973), Suppes (1971, 1972a), and Suppes and Morningstar (1970). Suppes, Goldberg, Kanz, Searle, and Stauffer (1971) documented many of the IMSSS curriculums used by this project and many of the operational techniques of CAI. Friend (1971) documented a coding language, INSTRUCT, and an associated coding system that was used to program several curriculums for this project.

Experiment I: Mathematics Strands

Experiment I was reported by Suppes, Fletcher, Zanotti, Lorton, and Searle (1973).

Experiment II: Mathematics Strands

Experiment II was reported by Suppes, Fletcher, and Zanotti (1973).

Experiment III: GP Measures

Experiment III was reported by Suppes, Fletcher, Zanotti, Lorton, and Searle (1973).

Arithmetic Skills of Deaf and Hearing Students

Comparisons of deaf and hearing students' arithmetic skills were reported by Suppes and Flannery (1973). Comparisons of deaf and hearing students' arithmetic word problem solving skills were reported by Searle, Lorton, and Suppes (1973).

Problem Solving Experiment

The experiment on structural variables that affect arithmetic word problem difficulty was reported by Searle, Lorton, and Suppes (1973).

Language Arts Experiment

The language arts evaluation experiment was reported by Fletcher and Beard (1973).

Language Arts Item Analysis

An extensive analysis of the language arts curriculum items was reported by Fletcher and Beard (1973).

Test Development

Theoretical and practical results in test development were reported by Fletcher and Beard (1973), Jensema (1973a, 1973b), and Suppes, Fletcher, Zanotti, Lorton, and Searle (1973).

Economics and Technology of the Network

The economics and technology of the CAI network maintained by this project was discussed by Ball and Jamison (1972) and Sanders and Ball (1972).

Grammatical and Semantic Analysis of Language Used by the Deaf

The analysis of a corpus of writing samples produced by deaf students was reported by Fletcher, Jamison, Searle, and Smith (1973). Perceptual confusions observed in learning and perceiving letters of the manual alphabet was reported by Weyer (1973). A study of English as the second language of deaf students was reported by Charrow and Fletcher (1973).

Surveys

Three extensive surveys of the professional literature on deaf students were prepared by Bonvillian and Charrow (1972), Bonvillian, Charrow, and Nelson (1973), and Suppes (1972b).

V. CONCLUSIONS

We began this project with the conviction that we had a powerful instructional tool at our disposal. Our aims were to demonstrate that CAI could be used to advantage by deaf students, that it could support serious research in deaf education, and that a favorable argument could be made for the economics of CAI. Behind these aims was the general intent of initiating large-scale use of CAI in schools for the deaf. To some extent we successfully met each of these aims.

It seems reasonable to conclude that CAI can be used successfully by deaf students. We did not set out to apply CAI to all of deaf education; we attempted only what we could do well. Curriculum work was concentrated on the skill subjects of mathematics and language arts, and, within these subjects, we emphasized aspects that were most amenable to computer presentation. Under these constraints we achieved favorable results. Certainly, the gains in mathematics computation ability that were two to three times what is expected from classroom instruction and the precision with which grade placement increase could be predicted as a function of CAI sessions are notable.

We also conclude that CAI provides a substantial foundation for research on the problems and processes of deaf education. The range of research undertaken by this project barely represents the diversity of inquiry that can be supported by CAI. The unobtrusive and precise control over experimental conditions made possible by computer presentations as well as the accuracy and speed of computer arithmetic and data retrieval permit a wide spectrum of experimental possibilities that we have only begun to explore.

The major drawback of CAI is its cost. Computers require a sizable commitment of funds both for acquiring capital equipment and for maintaining operations. Fortunately, the steady increase in the quality of available CAI is matched by a steady decrease in its costs. In the mid-1960's, when CAI first became available, it cost about $40 per student contact hour. Currently, CAI offered by the IMSSS system costs $1.50-$2.50 per student contact hour, depending on communication expenses. For the immediate future we can expect continued decreases in the costs and continued increases in the quality of CAI.

The proof of this project is in its impact on deaf education. Specifically, the willingness of the participating schools to support CAI from their own funding sources is the ultimate test of the project's impact. To date 13 of the 15 schools that participated in this project have committed funds to continuing their CAI activity in 1973-74. The two remaining schools have not decided what CAI implementation alternative to adopt. Two schools that received no CAI from this project will be added to those supporting CAI in one network that directly resulted from this project. We expect the growth of CAI in schools for the deaf to continue.

VI. RECOMMENDATIONS

We emphasize that all work under this project properly represents beginnings. There is no recommendation we support with more enthusiasm than the recommendation that our work be taken up and continued. The accomplishments of this project are not end results, and it would be a disappointment if they were considered final.

ERIC Full Text Provided by ERIC

VII. BIBLIOGRAPHY

Ball, J., and Jamison, D. Computer-assisted instruction for dispersed populations: System cost models. Technical Report No. 190, Institute for Mathematical Studies in the Social Sciences, Stanford University, September 15, 1972.

Bonvillian, J. D., and Charrow, V. R. Psycholinguistic implications of deafness: A review. Technical Report No. 188, Institute for Mathematical Studies in the Social Sciences, Stanford University, July 14, 1972.

Bonvillian, J. D., Charrow, V. R., and Nelson, K. E. Psycholinguistic and educational implications of deafness. Human Development, 1973, in press.

Charrow, V. R., and Fletcher, J. D. English as the second language of deaf students. Technical Report No. 208, Institute for Mathematical Studies in the Social Sciences, Stanford University, July 20, 1973.

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

Fletcher, J. D., and Beard, M. H. Computer-assisted instruction in language arts for hearing-impaired students. Technical Report No. 215, Institute for Mathematical Studies in the Social Sciences, Stanford University, September, 1973.

Fletcher, J. D., Jamison, D. T., Searle, B. W., and Smith, R. L. Computer-assisted instruction for the deaf at Stanford University. Annual Report, U.S.O.E., Contract No. OEG-0-70-4797 (607), Institute for Mathematical Studies in the Social Sciences, Stanford University, January 8, 1973.

Fletcher, J. D., and Stauffer, C. M. Learning language by computer. The Volta Review, 1973, 75, 302-311.

Friend, J. E. Instruct coders' manual. Technical Report No. 172, Institute for Mathematical Studies in the Social Sciences, Stanford University, May 1, 1971.

Friend, J. E. Computer-assisted instruction in programming: A curriculum description. Technical Report No. 211, Institute for Mathematical Studies in the Social Sciences, Stanford University, July 31, 1973.

Friend, J. E., and Atkinson, R. C. Computer-assisted instruction in programming: AID. Technical Report No. 164, Institute for Mathematical Studies in the Social Sciences, Stanford University, January 25, 1971.

Gentile, A., and Di Francesca, S. Academic achievement test performance of hearing impaired students. Washington: Office of Demographic Studies, Series D, Number 1, September, 1969.

Goldberg, A. A generalized instructional system for elementary mathematical logic. Technical Report No. 179, Institute for Mathematical Studies in the Social Sciences, Stanford University, October 11, 1971.

Goldberg, A., and Suppes, P. A computer-assisted instruction program for exercises on finding axioms. Technical Report No. 186, Institute for Mathematical Studies in the Social Sciences, Stanford University, June 23, 1972.

Jackson, R. M. Computer-assisted instruction for deaf and hard of hearing children. Report to the Superintendent and the Board of Education, Palo Alto Unified School District, 25 Churchill Avenue, Palo Alto, California, June, 1972.

Jamison, D., Fletcher, J. D., Suppes, P., and Atkinson, R. C. Cost and performance of computer-assisted instruction for education of disadvantaged children. In J. Fromkin and R. Radner (Eds.), Education as an industry. New York: National Bureau of Economic Research, Columbia University Press, 1973 (in press).

Jensema, C. Useful techniques for applying latent trait mental-test theory. Technical Report No. 202, Institute for Mathematical Studies in the Social Sciences, Stanford University, May 9, 1973. (a)

Jensema, C. A note on the reliability of Bayesian tailored testing. Journal of Educational and Psychological Measurement, 1973, in press. (b)

Lord, F. M. Some test theory for tailored testing. In W. H. Holtzman (Ed.), Computer assisted instruction, testing, and guidance. New York: Harper & Row, 1971. (a)

Lord, F. M. Robbins-Monro procedures for tailored testing. Journal of Educational and Psychological Measurement, 1971, 31, 3-31. (b)

Lord, F. M., and Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Sanders, W. R., and Ball, J. R. Logic documentation standard for the Institute for Mathematical Studies in the Social Sciences. Technical Report No. 191, Institute for Mathematical Studies in the Social Sciences, Stanford University, October 4, 1971.

Searle, B. W., Lorton, P. V., and Suppes, P. Structural variables affecting CAI performance on arithmetic word problems of disadvantaged and deaf students. Technical Report No. 213, Institute for Mathematical Studies in the Social Sciences, Stanford University, August, 1973.

Suppes, P. Some theoretical models for mathematics learning. Journal of Research and Development in Education, 1967, 1, 5-22.

Suppes, P. Computer-assisted instruction for deaf students. American Annals of the Deaf, 1971, 116, 500-508.

Suppes, P. Facts and fantasies of education. Technical Report No. 193, Institute for Mathematical Studies in the Social Sciences, Stanford University, October 18, 1972. (a)

Suppes, P. A survey of cognition in handicapped children. Technical Report No. 197, Institute for Mathematical Studies in the Social Sciences, Stanford University, December 29, 1972. (b)

Suppes, P., and Flannery, L. I. A comparison of deaf and normal-hearing students' performance in arithmetic. Institute for Mathematical Studies in the Social Sciences, Stanford University, in press.

Suppes, P., Fletcher, J. D., and Zanotti, M. Models of individual trajectories in computer-assisted instruction for deaf students. Technical Report No. 214, Institute for Mathematical Studies in the Social Sciences, Stanford University, September, 1973.

Suppes, P., Fletcher, J. D., Zanotti, M., Lorton, P. V., and Searle, B. W. Evaluation of computer-assisted instruction in elementary mathematics for hearing-impaired students. Technical Report No. 200, Institute for Mathematical Studies in the Social Sciences, Stanford University, March 17, 1973.

Suppes, P., Goldberg, A., Kanz, G., Searle, B., and Stauffer, C. Teacher's handbook for CAI courses. Technical Report No. 178, Institute for Mathematical Studies in the Social Sciences, Stanford University, September 1, 1971.

Suppes, P., and Morningstar, M. Technological innovations: Computer-assisted instruction and compensatory education. In F. Korten, S. Cook, and J. Lacey (Eds.), Psychology and the problems of society. Washington, D.C.: American Psychological Association, 1970.

Suppes, P., and Morningstar, M. Computer-assisted instruction at Stanford, 1966-68: Data, models, and evaluation of the arithmetic programs. New York: Academic Press, 1972.

Suppes, P., Searle, B., and Lorton, P., Jr. The strands arithmetic CAI program. Institute for Mathematical Studies in the Social Sciences, Stanford University, 1973, in press.

Weyer, S. A. Fingerspelling by computer. Technical Report No. 212, Institute for Mathematical Studies in the Social Sciences, Stanford University, August, 1973.