DOCUMENT RESUME

ED 054 978		SE 012 549
AUTHOR	Jamison, Dean; And Others	
TITLE	Cost and Performance of Computer-Assisted Instruction for Education of Disadvantaged Children.	
INSTITUTION	Stanford Univ., Calif. Graduate School of Business.; Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.	
PUB DATE	Jul 71	
NOTE	59p.	
EDRS PRICE	MF-$0.65 HC-$3.29	
DESCRIPTORS	Achievement Gains; Arithmetic; *Computer Assisted Instruction; *Disadvantaged Youth; *Evaluation; Instructional Media; Program Costs; Reading; *Research	

ABSTRACT

This paper is divided into three sections. The first section describes three programs in computer-assisted instruction (CAI) that have been developed by the Institute for Mathematical Studies in the Social Sciences at Stanford University and have performed well with underachieving children. These programs are in elementary arithmetic, initial reading, and computer programing for high school students. The second section, the major part of this paper, reports a detailed evaluation of these programs. Two criteria for successful performance are examined: simple achievement gain, and reduction of achievement inequality. The final section deals with the problem of making CAI available in rural as well as urban areas, and attempts a realistic assessment of the total costs. An estimate is also made of the increase in student to teacher ratio required to provide CAI without an increase in expenditure per student. (MM)

Graduate School of Business

STANFORD UNIVERSITY

Research Paper No. 20

COST AND PERFORMANCE OF COMPUTER-ASSISTED INSTRUCTION FOR EDUCATION OF DISADVANTAGED CHILDREN*

Dean Jamison
J. Dexter Fletcher
Patrick Suppes
Richard Atkinson

July, 1971

STANFORD UNIVERSITY

Graduate School of Business
Research Paper 20

Institute for Mathematical Studies in the Social Sciences (IMSSS)

Technical Report No. 175

Psychology and Education Series

COST AND PERFORMANCE OF COMPUTER-ASSISTED INSTRUCTION FOR EDUCATION OF DISADVANTAGED CHILDREN*

Dean Jamison[a]
J. Dexter Fletcher[b]
Patrick Suppes[c]
Richard Atkinson[d]

July 1971

*Presented at the National Bureau of Economic Research Conference on Education as an Industry, Chicago, June 1971.

(a) Assistant Professor of Management Science, Graduate School of Business.

(b) Ph.D. Candidate, School of Education.

(c) Professor of Philosophy, Statistics, and Education; Director, IMSSS.

(d) Professor of Psychology and Education; Chairman, Department of Psychology.

All the authors are staff members of the IMSSS.

Cost and Performance of Computer-assisted Instruction for Education of Disadvantaged Children*

by

D. Jamison, J. D. Fletcher, P. Suppes and R. Atkinson

Stanford University, Stanford, California 94305

I. INTRODUCTION

This paper discusses the potential role of computer-assisted instruction (CAI) in providing compensatory education for disadvantaged children. All CAI involves, to one extent or another, the interaction of students with computers. Curriculum material is stored by a computer which is provided with decision procedures for presenting the material to individual students. Typically students work at terminals, usually teletypewriters, which are located at school sites and are connected by telephone lines to a central computer. Using time-sharing techniques, a single computer may serve more than 500 students simultaneously at diverse and remote locations. These advances in time-sharing techniques coupled with reductions in hardware costs and increasing availability of tested curriculum material are beginning to make CAI economically attractive as a source of compensatory education. Pedagogically, the value of CAI is established by its capacity for immediate evaluation of student responses and detailed individualization of treatment based on accurate and rapid retrieval of performance histories.

A number of institutions in the United States have computer-assisted programs underway in varying scales of complexity. Zinn (1969) provides

---

*This research was supported in part by National Science Foundation Grant NSFG-443X to the Institute for Mathematical Studies in the Social Sciences, Stanford University. A. Kelley, S. Michelson, and D. Wiley provided helpful comments on an earlier draft of this paper.

an overview of these efforts. Stanford University's Institute for Mathematical Studies in the Social Sciences (IMSSS) has been engaged in such development efforts for a period of ten years and now operates one of the largest CAI centers in the country. This paper discusses the Institute's efforts to use CAI to provide compensatory education for disadvantaged students. Before turning to these efforts, however, it is worthwhile to place our work in the context of the large national effort in compensatory education that has been financed, primarily, by Title I of the Elementary and Secondary Education Act of 1965.

For a number of years, about one billion dollars has been spent annually by the federal government to provide compensatory education for disadvantaged children in the United States. Unfortunately, much of the available evidence suggests that these federally funded Title I programs have met little success. During the period 1966-68 Piccariello (1969) conducted a large-scale evaluation of Title I-funded reading programs and in more than two instances out of three found no significant achievement differences between children in control groups and children in one of the Title I programs. Further, only slightly more than half of the significant differences obtained were in a positive direction. In his widely discussed paper on I.Q. and scholastic achievement, Jensen (1969) surveyed a large number of studies indicat[illegible] general failure of compensatory education.

Rather than studying the typical compensatory education program, Kiesling (1970) undertook a study of those compensatory education programs that had been most successful in the State of California. Kiesling concluded that there were a number of common elements in all these successful programs, and that one could learn from their success and replicate them. Thus while compensatory education may have been, on the average, unsuccessful in the past, Kiesling feels there is no reason to repeat these failures. Success could be achieved by tailoring future compensatory programs around those that have proven themselves previously. Kiesling presented a number of paradigmatic compensatory programs for both arithmetic and reading and estimated their annual cost per student to be on the order of $200 to $300 per year in addition to the normal school allotment for that student.

A different interpretation from Kiesling's of the failure of compensatory education of that what goes on in schools has little effect on the achievement of students. This view received considerable support in Coleman (1966), and is consistent with the views of Jensen (1969). Coleman concluded that factors within the schools seem to affect achievement much less than do factors outside the schools; these somewhat disheartening conclusions have been subject to rather vigorous debate since their initial publication. A number of recent views of interpreting the data of the Coleman survey may be found in Mood (1970). The general drift of the papers in this book is that schooling is rather more important than one would conclude from the initial Equality of Educational Opportunity report; nevertheless, there is an increasing concensus, since publication of the Report, that input factors in the schooling process seem to have a good deal less effect on the outputs than had been thought previously.

Our own work, however, has led us to more optimistic conclusions concerning the potential capability of the schools to affect scholastic performance. We have found strong and consistent achievement gains by disadvantaged students when they are given CAI over a reasonable fraction of a school year. Thus we are more inclined to accept Kiesling's general conclusions that compensatory education can work than the pessimistic interpretations of the Coleman Report. As Bowles and Levin (1968) pointed out: "The findings of the Report are particularly inappropriate for assessing the likely effects of radical changes in the level and compositions of resources devoted to schooling because the range of variation in most school inputs in this sample is much more limited than the range of policy measures currently under discussion." Our evaluations of CAI provide detailed information about the output effects of a much broader variety of school inputs than the Coleman Report was able to consider.

This paper reports on the performance of three CAI programs that have performed well with underachieving children. Section II of the paper describes those programs--one in elementary arithmetic, one in initial reading and one designed to teach computer programming to high school students. Section III reports on an evaluation of the performance of these programs. We consider two aspects of performance: achievement gain and the degree to which the program enabled disadvantaged students to close the gap between themselves and more advanced students. In order to examine this latter, distributional effect, we rely in part on Gini coefficients derived from Lorenz curve representations of achievement data. We also examine the results in the light of several alternative mathematical formulations of "inequality-aversion". Section IV of the paper provides a detailed discussion of costs. In particular, we examine the problem of making computer-assisted instruction available in rural areas as well as urban ones and attempt a realistic assessment of those costs. Our cost projections are for systems having on the order of 1,000 student terminals; this number of terminals would allow 20,000 to 30,000 students to use the system per day. We compute not only dollar costs but also opportunity costs for using CAI in order to estimate the increase in student to teacher ratios that would be required if CAI were introduced under the constraint that per student expenditures remain constant.

II. DESCRIPTION OF THREE PROGRAMS

A. Arithmetic

Development of computer-assisted drill and practice in elementary-school mathematics (grades 1-6) was begun by the Institute in 1965. The intent of the program is to provide drill and practice in arithmetic skills, especially computation, as an essential supplement to regular classroom instruction. Concepts presented by the CAI program are assumed to have been previously introduced to the students by their classroom teacher.

Curriculum material for each of the six elementary-school grades is arranged sequentially in 20-27 concept blocks that correspond in order and content to the mathematical concepts presented in several textbook series that were surveyed during the development of the curriculum. Each concept block consists of a pretest, five drills divided into five levels of difficulty, and a posttest. The pre- and posttests are comprised of equal numbers of items drawn from each of the five difficulty levels in the drills. Each block contains approximately seven days of activity, one day each for the pre- and posttests and five days for the five drills. As part of each day's drill a student also receives review items drawn from previously completed concept blocks. Review material comprises about a third of a day's drill.

The level of difficulty for the first drill within a block is determined by a student's pretest performance for the block. The level of difficulty for each successive day's drill is determined by the student's performance during the preceding day. If a student's performance on a drill is 80 percent or more correct, his next drill will be one level of difficulty higher; if his performance on a drill is 60 percent or less correct, his next drill will be one level of difficulty lower.

The drill content, then, is the same for all students in a class with only the difficulty levels varying from student to student. The content of the review material, however, is uniquely determined for each student on the basis of his total past performance history. His response history is scanned to determine the previously completed concept block for which his posttest score was lowest, and it is from this block that review exercises are drawn. Material from the review block is included in the first four drills for the current block, and a posttest for the review block is given during the fifth drill. The score on this review posttest replaces the previous posttest score for the review block and determines subsequent review material for the student.

Student terminals for the arithmetic drill and practice are Model-33 teletypewriters without the random audio capability required for the reading program. As in the reading program, these teletypewriters are located at school sites and are connected by telephone lines to the Institute's central computer facility at Stanford University. Students complete a concept block about every 1-1/2 weeks. The program is described extensively in a number of publications including Suppes and Morningstar (1969) and Suppes, Jerman and Brian (1968).

A more highly individualized strand program in arithmetic has been developed over the past several years and is now replacing the program just described. Our performance data in this paper are for the earlier program; a description of the more recent program may be found in Suppes and Morningstar (1970).

B. Reading

CAI in initial reading (grades K-3) has been under development by IMSSS since 1965. The original intent of the reading program was to implement a complete CAI curriculum using cathode-ray tubes (CRT), light pen and typewriter input, slides, and random access audio. These efforts, described in Atkinson (1968), were successful, but prohibitively expensive. Economically and pedagogically, some aspects of initial reading seemed better left to the classroom teacher. Subsequent efforts of the reading project were directed toward the development of a CAI reading curriculum that would supplement, but not replace, classroom reading instruction.

The current reading curriculum requires only the least expensive of teletypewriters and some form of randomly accessible audio. No graphic or photographic capabilities are needed and only upper-case letters are used. Despite these limitations, an early evaluation of the curriculum indicates that it is of significant value (Fletcher and Atkinson, 1971).

The curriculum, more fully described in Atkinson, Fletcher, Chetin and Stauffer (1971), emphasizes phonics instruction. There are two primary reasons for this emphasis. First, it enables the curriculum to be based on a relatively well-defined aspect of reading theory making it more amenable to computer presentation. Second, the phonics emphasis on the regular grapheme-phoneme correspondences (or "spelling patterns") which occurs across all English orthography insures that the program appropriately supplements classroom instruction using any initial reading vocabulary.

Instruction is divided into seven content areas or "strands": 0 - machine readiness; I - letter identification; II - sight-word vocabulary; III - spelling patterns; IV - phonics; V - comprehension categories; and VI - comprehension sentences.

The term strand in the reading program defines a basic component skill of initial reading. Students in the reading program move through each strand in a roughly linear fashion. Branching or progress within strands is criterion dependent; a student proceeds to a new exercise within a strand only after he has attained some (individually specifiable) performance criterion in his current exercise. Branching between the strands is time dependent; a student moves from one strand to take up where he left off in another after a certain (again, individually specifiable) amount of time, regardless of what criterion levels he has reached in the strands. Within each strand there are 2-3 progressively more difficult exercises that are designed to bring students to fairly high levels of performance. The criterion procedure is explained in more detail in Atkinson et al. (1971), but basically it requires two consecutive correct answers for each item.

Entry into each strand is dependent upon a student's performance in earlier strands. For example, the letter-identification strand starts with a subset of letters used in the earliest sight words. When a student in the letter-identification strand exhibits mastery over the set of letters used in the first words of the sight-word strand, he enters that strand. Initial entry into both the phonics and spelling pattern strands is controlled by the student's placement in the sight-word strand. Once he enters a strand, however, his advancement within it is independent of his progress in other strands. On any given day, a student's lesson may draw exercises from one to five different strands.

Most students spend 2 minutes in each strand and the length of their daily sessions is 10 minutes. A student may be stopped at any point in an exercise, either by the maximum-time rule for the strand or by the session time limit; however, sufficient information is saved in his record to assure continuation from precisely the same point in the exercise when he next encounters that strand.

C. Computer Programming

Development of computer-assisted instruction in computer programming was begun by the Institute in 1968 and was initially made available to students at an "inner city" high school in February, 1969. Requisite knowledge of computer languages and systems varies greatly among applications and, for this reason, general concepts of computer operations rather than knowledge of the specific languages or systems used are emphasized in the curriculum. To achieve this generality, the curriculum ranges from problems in assembly-language coding to symbol manipulation and test-processing. The three major components of the curriculum are SIMPER (Simple Instruction Machine for the Purpose of Educational Research), SLOGO (Stanford LOGO), and BASIC. Associated with each component are interpreters, utility routines and curriculum material.

Basically, computers "understand" only binary numbers. These numbers may be either data or executable instructions. A fundamental form of programming is to write code as a series of mnemonics, which bear a one-to-one relationship to the binary number-instructions executable by a machine; this type of coding is called assembly-language programming. The instructions of higher order languages, such as BASIC and SLOGO, do not bear a one-to-one relationship to the instructions executed by a machine and, therefore, obscure the fundamental operations performed by computers during program execution. The intent of SIMPER, therefore, is to make available to students using teletypewriters a small computer that can be programmed in a simple assembly language. The SIMPER computer is, of course, mythical, since giving beginning students such sensitive access to an actual time-sharing computer would be both prohibitively expensive and potentially disastrous.

As simulated, SIMPER is a two-register, fixed-point, single-address machine with a variable size memory. There are 16 operations in its instruction set. To program SIMPER, a student types the pseudo operation "LOC" to tell SIMPER where in its memory to begin program execution, and then enters the assembly-language code that comprises his solution to an assigned problem. During execution of the student's program, SIMPER types the effect of each instruction on its memory and registers. In this way, students hopefully receive special insight into how each instruction operates and how a series of computer instructions is converted into meaningful work.

SLOGO, the Institute's implementation of LOGO, is the second major component of the curriculum. LOGO is a symbol manipulation and string-processing language developed by a major computer utilities company expressly for teaching the principles of computer programming. It is suitable for manipulating data in the form of character strings, as well as for performing arithmetic functions, and its most powerful feature is its capacity for recursive functions. It was thought that the computer applications most characteristic of the employment available to these students would be the inventory control problems that arise in filing and stockroom management, and it is these problems that are stressed in the SLOGO component of the curriculum. Students are taught not only the SLOGO languages, but the data structures needed for applications such as tree searches and string editing.

SIMPER and SLOGO are more fully documented in Lorton and Slimick (1969). They were written for the Institute's PDP-10 computer and were made available to students in the Spring and Fall of 1969. Mixed with the usual, well-documented enthusiasm of all students for CAI was some disappointment among the computer programming students that they were not learning a computer language generally found in industry. For this reason, the ubiquitous BASIC programming language was prepared for the Institute's PDP-10 computer and made available to the students in the spring of 1970.

The BASIC course, as the SIMPER and SLOGO courses before, was designed to permit maximum student control. Most of this control concerned the use of such optimal material as detailed review, overview lessons and self-tests. Students were aware that they would be graded only on homework and tests, and it was emphasized that their course grades would not include wrong answers made in the BASIC teaching program.

The course consists of 50 lessons, each comprised of 20-100 problems and each requiring 1-2 hours to complete. The lessons are organized into blocks of five. Each lesson is followed by a review printout and each block of five lessons is followed by a self-test and overview lesson. Students receive these review printouts, self-tests and overview lessons at their option. Each block is terminated by a short graded test that is evaluated partly by computer and partly by the supervising teacher.

Students are given as much time as needed to answer each problem. Since the curriculum emphasizes tutorial instruction rather than drill material, students may spend several minutes thinking or calculating before entering a response; hence, there is no time limit. Because the subject matter of the course is a formal language which is necessarily unambiguous to a computer, extensive analysis of students' responses is possible and highly individualized remediation can be provided for wrong, partially wrong or simply inefficient solutions to assigned problems. Significantly, individual errors and misconceptions can be corrected by additional instruction and explanation without incorporating unnecessary exposition in the mainstream of the lesson.

III. PERFORMANCE

We conceive compensatory education to have two broad purposes with respect to student achievement. The first is, of course, to increase the student's achievement level over what it would have been without compensatory education. We discuss achievement gains in III. A. The second purpose of compensatory education is to decrease the spread among students or to make the distribution of educational output more nearly equitable. The notion of "equality" in education has received considerable attention in recent years, and we make no attempt to review that literature here; Coleman (1968) provides a useful overview of some of the issues. Michelson (1970) discusses inequality in real inputs in producing achievement and in a later paper--Michelson (1971)-- discusses inequality in financial inputs. Our treatment differs in focusing on output inequality and, methodologically, in utilizing tools recently developed by economists for analyzing distribution of income. Section III.B. discusses our results in this area.

A. Achievement Gain

Gains in arithmetic. During the 1967-68 school year, approximately 1,000 students in California, 1,100 students in Kentucky and 600 students in Mississippi participated in the arithmetic drill-and-practice program. Sufficient data were collected to permit CAI and non-CAI group comparisons for both the California and Mississippi students. The California students were drawn from upper middle-class schools in suburban areas quite uncharacteristic of those for which compensatory education is usually intended. The Mississippi students, on the other hand, were drawn from an economically and culturally deprived rural area and provided an excellent example of the value of CAI as compensatory education.

The Mississippi students (grades 2-6) were given appropriate forms of the Stanford Achievement Test (SAT) in October, 1967. The SAT was administered to the Mississippi first-grade students in February, 1968. All the Mississippi students (grades 1-6) were posttested with the SAT in May, 1968. Twelve different schools were used; eight of these included both CAI and non-CAI students, three included only CAI students, and one included only non-CAI students. Within the CAI group, 1-10 classes were tested at each grade level, and within the non-CAI group, 2-6 classes were tested at each grade level. Achievement gains over the school year were measured by the differences between pre- and posttest grade placements estimated by the SAT computation subscale. Average pretest and posttest grade placements, calculated differences of these averages, t-values for these differences, and degrees of freedom for each grade's CAI and non-CAI students are presented in Table III.1. Significant t-values $(p < .01)$ are starred. The

----

Insert Table III.1 about here

----

performance of the CAI students improved significantly more over the school year than that of the non-CAI students in all but one of the six grades. The largest differences between CAI and non-CAI students occurred in grade 1 where, in only three months, the average increase in grade placement for CAI students was 1.14, compared with .26 for the non-CAI students.

On other subscales of the SAT, the performance of CAI students, measured by improvement in grade placement, was significantly better than that of the non-CAI students on the SAT concepts subscale for grade 3 $(t(76) = 3.01,\ p < .01)$ and for grade 6 $(t(433) = 3.74,\ p < .01)$ and on the SAT application subscale for grade 6 $(t(433) = 4.09,\ p < .01)$. In grade 4, the non-CAI students improved more than the experimental group on the concepts subscale $(t(131) = 2.25,\ p < .05)$.

Appropriate forms of the SAT were administered to all the California students (grades 1-6) in October, 1967 and again in May, 1968. Seven different schools were used. Two of the schools included both CAI and non-CAI students, two included only CAI students and three included only non-CAI students. Within the CAI group 5-9 classes were tested at each grade level, and within the non-CAI group, 6-14 classes were tested at each grade level. Average pretest and posttest grade placements on the SAT computation subscale, calculated differences of these

Table III.1 - Average Grade-placement Scores on the Stanford Achievement Test: Mississippi 1967-68[a]

Pretest		Posttest		Posttest-pretest			Degrees of free-dom
Experi-mental	Con-trol	Experi-mental	Con-trol	Experi-mental	Con-trol		
1.41(52)*	1.19(63)	2.55	1.45	1.13	0.26	9.63**	113
1.99(25)	1.96(54)	3.37	2.80	1.38	0.84	4.85**	77
2.82(22)	2.76(56)	4.85	4.04	2.03	1.26	4.87**	76
2.34(56)	2.45(77)	3.36	3.14	1.02	0.69	2.28	131
3.09(83)	3.71(134)	4.46	4.60	1.37	0.89	3.65**	215
4.82(275)	4.35(160)	6.54	5.49	1.72	1.13	4.89**	433

es in parentheses are numbers of students.

.01

assumptions underlying this test of significance are, first, that the two distributions red are distributed normally and, second, that their variances are equal. Robustness e t-test is discussed by Boneau (1960) and Elashoff (1968) among others.

averages, t-values for these differences and degrees of freedom for each grade's CAI and non-CAI students are presented in Table III.2. As in Table III.1, significant t-values ($p < .01$) are starred. The performance of the CAI students improved significantly more over the

---

Insert Table III.2 about here

---

school year than that of the non-CAI students in grades 2, 3 and 5. On other subscales of the SAT, the CAI students improved significantly more over the school year than did the non-CAI students on the concepts subscale for grade 3 ($t(344) = 4.13$, $p < .01$) and on the application subscale for grade 6 ($t(399) = 2.14$, $p < .05$).

A comparison of the California students with the Mississippi students suggests at least two observations worth noting. First, when significant effects were examined for all six grades, the CAI program was more effective for the Mississippi students than for the California students. Second, changes in performance level for the CAI groups were quite similar in both states, but the non-CAI group changes were very small in Mississippi relative to the non-CAI group changes in California. These observations suggest that CAI may be more effective when students perform well below grade level and are in need of compensatory education, as in the rural Mississippi schools, than when the students receive an adequate education, as in the suburban California schools.

These data do not fully reflect the breadth of educational experience permitted by CAI. Some of the Mississippi students took the Institute's beginning course in mathematical logic and algebra, which had been prepared for bright fourth to eighth grade students whose teachers were not prepared to teach this advanced material. At the end of the 1967-68 school year, two Mississippi Negro boys placed at the top of the first-year mathematical logic students, almost all of whom came from upper middle-class suburban schools.

<u>Gains in reading</u>. The data used in this report were collected during the 1969-70 school year and are also discussed in Fletcher and Atkinson (1971). In November, 1969, 25 pairs of first-grade boys and

Table III.2 - Average Grade-placement Scores on the Stanford Achievement Test: California 1967-68[a]

Grade	Pretest Experimental	Pretest Control	Posttest Experimental	Posttest Control	Posttest-pretest Experimental	Posttest-pretest Control	t	Degrees of freedom
1	1.39(58)*	1.31(259)	2.62	2.51	1.23	1.21	0.20	315
2	2.06(65)	2.16(238)	3.20	2.89	1.14	0.73	4.96**	301
3	3.00(136)	2.85(210)	4.60	3.86	1.60	1.02	6.70**	344
4	3.40(103)	3.49(185)	4.87	5.00	1.46	1.51	-0.41	286
5	4.98(149)	4.44(90)	6.41	5.31	1.43	0.88	4.06**	237
6	5.42(154)	5.70(247)	7.43	7.59	2.01	1.90	0.84	399

*Values in parentheses are numbers of students.

**$p < .01$

[a]The assumptions underlying this test of significance are, first, that the two distributions compared are distributed normally and, second, that their variances are equal. Robustness of the t-test is discussed by Boneau (1960) and Elashoff (1968) among others.

25 pairs of first-grade girls were matched on the basis of the Metropolitan Readiness Test (MET). Matching was achieved so that the MET scores for a matched pair of subjects were no more than two points apart. Moreover, an effort was made to insure that both members of a matched pair had classroom teachers of roughly equivalent ability.

The experimental member of each matched pair of students received 8 to 10 minutes of CAI instruction per school day roughly from the first week in January until the second week in June. The control member of each pair received no CAI instruction. Except for the 8- to 10-minute CAI period, there is no reason to believe that the activities during the school day were any different for the experimental and control subjects.

Four schools within the same school district were used. Two schools provided the CAI students and two different schools provided the non-CAI subjects. The schools were in an economically depressed area eligible for federal compensatory education funds.

Three posttests were administered to all subjects in late May and early June, 1970. Four subtests of the Stanford Achievement Test (SAT), Primary I, Form X, were used: word reading (S/WR), paragraph meaning (S/PM), vocabulary (S/VOC), and word study (S/WS). Second, the California Cooperative Primary Reading Test (COOP), Form 12A (grade 1, spring) was administered. Third, a test (DF) developed at Stanford and tailored to the goals of the CAI reading curriculum was administered individually to all subjects.

During the course of the school year, an equal number of pairs was lost from the female and male groups; complete data were obtained for 22 pairs of boys and 22 pairs of girls.

Means and t-values for differences in SAT, COOP, and DF total scores are presented in Table III.3. In this table t-values are

---

Insert Table III.3 about here

---

displayed in brackets. The t-values calculated are for nonindependent samples, and those that are significant ($p < .01$, one-tailed) are starred.

Table III.3 - Means and t-values[a] for the Stanford Achievement Test (SAT), the California Cooperative Primary Test (COOP), and the CAI Reading Project Test (DF)[b]

	SAT	COOP	DF
CAI	112.7	33.4	64.5
	[4.22*]	[4.04*]	[6.46*]
non-CAI	93.3	35.7	54.8

*$p < .01$, df = 43

[a]in brackets

[b]The assumptions underlying this test of significance are, first, that the two distributions compared are distributed normally and, second, that their variances are equal. Robustness of the t-test is discussed by Boneau (1960) and Elashoff (1968) among others.

The results of these analyses were encouraging. All three indicated a significant difference in favor of the CAI reading subjects. These differences were also important from the standpoint of improvement in estimated grade placement. Table III.4 displays the mean grade placement of the two groups on the SAT and COOP.

Insert Table III.4 about here

Means and t-values for the differences on the four SAT subtests are presented in Table III.5. As in Table III.3 t-values are displayed in brackets; t-values that are significant ($p < .01$,

---

Insert Table III.5 about here

---

one-tailed) are starred.

These SAT subtests revealed some interesting results. Of the four SAT subtests, the S/WS was expected to reflect most clearly the goals of the CAI curriculum; yet greater differences between CAI and non-CAI groups were obtained for both the S/WR and S/PM subtests. Also notable is the lack of any real differences for the S/VOC. One explanation for this result is that the vocabulary subtest measures a pupil's vocabulary independent of his reading skill (Kelley et al., 1964); since the CAI reading curriculum is primarily concerned with reading skill and only incidentally with vocabulary growth, there may have been no reason to expect a discernible effect of the CAI curriculum on the S/VOC. Most notable, however, are the S/PM results. The CAI students performed significantly better on paragraph items than did the non-CAI students, despite the absence of paragraph items in the CAI program and the relative dearth of sentence items. These results for phonics-oriented programs are not unprecedented, as Chall's (1967, pp. 106-107) survey shows. Nonetheless, for a program with so little emphasis on connected discourse, they are surprising.

The effect of CAI on the progress of boys compared with the progress of girls is interesting to note. The Atkinson (1968) finding that boys

Table III.4 - Average Grade Placement on the Stanford Achievement Test (SAT) and the California Cooperative Primary Test (COOP)

	SAT	COOP
CAI	2.3	2.6
non-CAI	1.9	2.1

Table III.5 - Means and t-values[a] for the Word Reading (S/WR), Paragraph Meaning (S/PM), Vocabulary (S/VOC), and Word Study (S/WS) Subtests of the Stanford Achievement Test[b]

	S/WR	S/PM	S/VOC	S/WS
CAI	26.5	23.0	21.6	41.6
	[5.18*]	[4.17*]	[.35]	[3.78*]
non-CAI	20.1	16.3	21.2	35.7

*$p < .01$, $df = 43$

[a] in brackets

[b] The assumptions underlying this test of significance are, first, that the two distributions compared are distributed normally and, second, that their variances are equal. Robustness of the t-test is discussed by Boneau (1960) and Elashoff (1968) among others.

benefit more from CAI instruction than do girls is corroborated by these data. On the SAT the relative improvement for boys exposed to CAI versus those not exposed to CAI is 22 percent; the corresponding figure for girls is 20 percent. On the COOP the percentage improvement due to CAI is 42 for boys and 17 for girls. Finally, on the DF the improvement is 32 percent for boys and 13 percent for girls. Overall, these data suggest that both boys and girls benefit from CAI instruction in reading, but that CAI is relatively more effective for boys. Explanations of this difference are discussed in Atkinson (1968).

Achievement gains in the computer programming course. Eight weeks prior to the end of the 1969-70 school year, students who received CAI instruction in BASIC were given the SAT's mathematical computation and application sections. A control group of students from the same school was given the same test. At semester's end the test was repeated and the following additional data were gathered: (i) verbal achievement scores from the ninth-grade level test of the Equality of Educational Opportunity Survey, and (ii) responses to the socioeconomic status questionnaire of the EEO survey.

Sufficient pre- and posttest scores were obtained for 39 CAI students and 19 non-CAI students. Average pre- and posttest scores for the SAT computation and application subscales, average gains, and t-values for differences in the average gains achieved by CAI and non-CAI students are presented in Table III.6.

Insert Table III.6 about here

The SAT tests were used here in the absence of a standardized achievement test in computer programming; gains in arithmetic achievement are, then, only a proxy for gains in the skills to be taught in the course. Presumably students gained in arithmetic skill because they spent more than the usual time working on quantitative problems.

There was also a good deal of textual output at the teletype that the students needed to read and comprehend, and it was the unanimous

Table III.6 - Arithmetic Achievement for Computer Programming Course[a]

	CAI			Control				
	PRE	POST	GAIN	PRE	POST	GAIN	t	df
SAT computation	7.97	9.11	1.14	7.97	8.41	.44	1.68	55
SAT application	7.74	8.61	.86	8.33	8.38	.05	1.73	55

[a]The assumptions underlying this test of significance are, first, that the two distributions compared are distributed normally and, second, that their variances are equal. Robustness of the t-test is discussed by Boneau (1960) and Elashoff (1968) among others.

subjective impression of the teachers who worked with the students that they were better able to read as a result. However, scores on verbal achievement tests administered at the end of the school year showed virtually no differences between the CAI and control groups in this respect.

In order to identify some of the sources of achievement gain we ran a stepwise linear regression of gain scores (posttest minus pretest) against pretest scores, verbal scores, and various items from the SES questionnaire. The dependent variable was the sum of the gain scores on the computation and applications sections of the test. Table III.7 below lists the independent variables and the coefficients estimated for them.

Insert Table III.7 about here

The results in the table are self-explanatory, but we make two comments in conclusion. First, failure to have had CAI during this eight-week interval would remove about .5 years (one half of .99) of arithmetic achievement. (Naturally it would be desirable to replace the 0-1 CAI variable with actual amount of time on system; the regression coefficient would then have a good deal more practical value.) Second, the mathematics pretest has a negative coefficient; when CAI and control regressions were run separately, this coefficient is negative for CAI and positive for control. This implies that CAI in sufficient quantity would have an equalizing effect, a point to be further discussed in the next subsection. In a later paper we plan to analyze in much more detail the interaction of CAI and student background characteristics as determinants of scholastic achievement.

B. Reduction in Inequality

Our second criterion of performance concerns the extent to which CAI is inequality reducing. Clearly any compensatory program that has positive achievement gains, if applied only to those sectors of the population who perform least well, will have a tendency to reduce inequality. Often, however, entire schools receive the compensatory education and it is less obvious that the program will be inequality

Table III.7 - Determinants of Achievement Gain[a,b]

Independent variable	Mean	Standard deviation	Regression coefficient	Standard error
Constant term			[illegible].40	
CAI 0 CAI group; 1 control group	.35	.48	-.99	.96
Sum of pretest scores on computation and application	15.3	4.22	-.26	.14
Raw Score on verbal test	27.6	9.9	.17	.06
Age in years	15.9	2.5	-.23	.20
Race 0 Caucasian; 1 Other	.23	.42	-1.44	1.18
Number of people living in child's home	5.63	1.86	.13	.29
Total years of schooling of both parents	15.5	10.52	-.02	.05
Educational aspiration of student, in years of schooling	15.4	4.45	.07	.11
Previous Math GPA of student	2.40	1.30	-.11	.39

[a] Dependent variable is the sum of students' gain scores on arithmetic and computation sections of SAT.

[b] $r^2 = .26$

reducing. Our purpose in this subsection is to use techniques developed for analyzing inequality in the distribution of income to provide concrete measures of the extent to which CAI is inequality reducing. These measures are as applicable in cases where an entire student population receives the "compensatory" treatment as when only some subset of the population does.

We first use a traditional measure of inequality--the Gini coefficient based on the Lorenz curve--to examine before and after inequality in CAI and control groups and to examine inequality in achievement gains. Use of the Gini coefficient as a measure of inequality has, however, a number of shortcomings that are reviewed in A. Atkinson (1970). Prominent among these is that it is not purely an empirical measure but contains an underlying value judgment concerning what constitutes more inequality. Newbery (1970) has shown that it is impossible to make this value judgment explicit by means of any additive utility function. Therefore we also use the inequality measure proposed by A. Atkinson that does make explicit any underlying value judgments.

Use of either the Atkinson measure or Gini coefficients implies that achievement test scores should be measured on a ratio scale (i.e., the achievement measure must be unique up to multiplication by a positive constant). If, for example, achievement measures were only unique up to a positive linear transformation, the Gini coefficient could be made arbitrarily small by adding an arbitrarily large amount to each individual's achievement test score. The reader is cautioned that our assumption that achievement is measured on a ratio scale is quite strong; on the other hand, a ratio scale is essentially implicit in the assumption that one test score is better than another if and only if the number of problems correct on the one test is greater than the number correct on the other.

<u>Inequality measured by the Gini coefficient</u>. Consider a group of students who have taken an achievement test; each student will have achieved some score on the test, and there will be a total score obtained by summing all the individual scores. We may ask, for example,

what fraction of the total score was obtained by the 10 percent of students doing most poorly on the test, what fraction was obtained by the 20 percent of students doing most poorly, etc. The Lorenz curve plots fraction of total score earned by the bottom x percent of students as a function of x.

These concepts may be expressed more formally in the notation of Levine and Singer (1970) as follows. Let $N(u)$ be the achievement-score density function. Then $N(u)du$ represents the number of individuals scoring between u and $u + du$. The total number of students, N, and their average score, A, are given by:

$$N = \int_0^{\infty} N(u)du, \qquad \text{and}$$

$$A = \frac{1}{N}\int_0^{\infty} uN(u)du \quad .$$

The fraction of students scoring a or less is given by

$$f(a) = \frac{1}{N}\int_0^{a} N(u)du \quad ,$$

and the fraction of the total score obtained by students scoring a or less is

$$g(a) = \frac{\int_0^{a} uN(u)du}{NA} \quad .$$

The Lorenz curve plots $g(a)$ as a function of $f(a)$, and a typical Lorenz curve for our results is shown in Figure III.1 below. The $f(a)$, $g(a)$ pairs are obtained by computing these functions for all values of a.

Insert Figure III.1 about here

If there were a perfectly equitable distribution of achievement (everyone having identical achievement) the Lorenz curve would be the 45° line depicted in Figure III.1. The more $g(a)$ differs from the 45° line

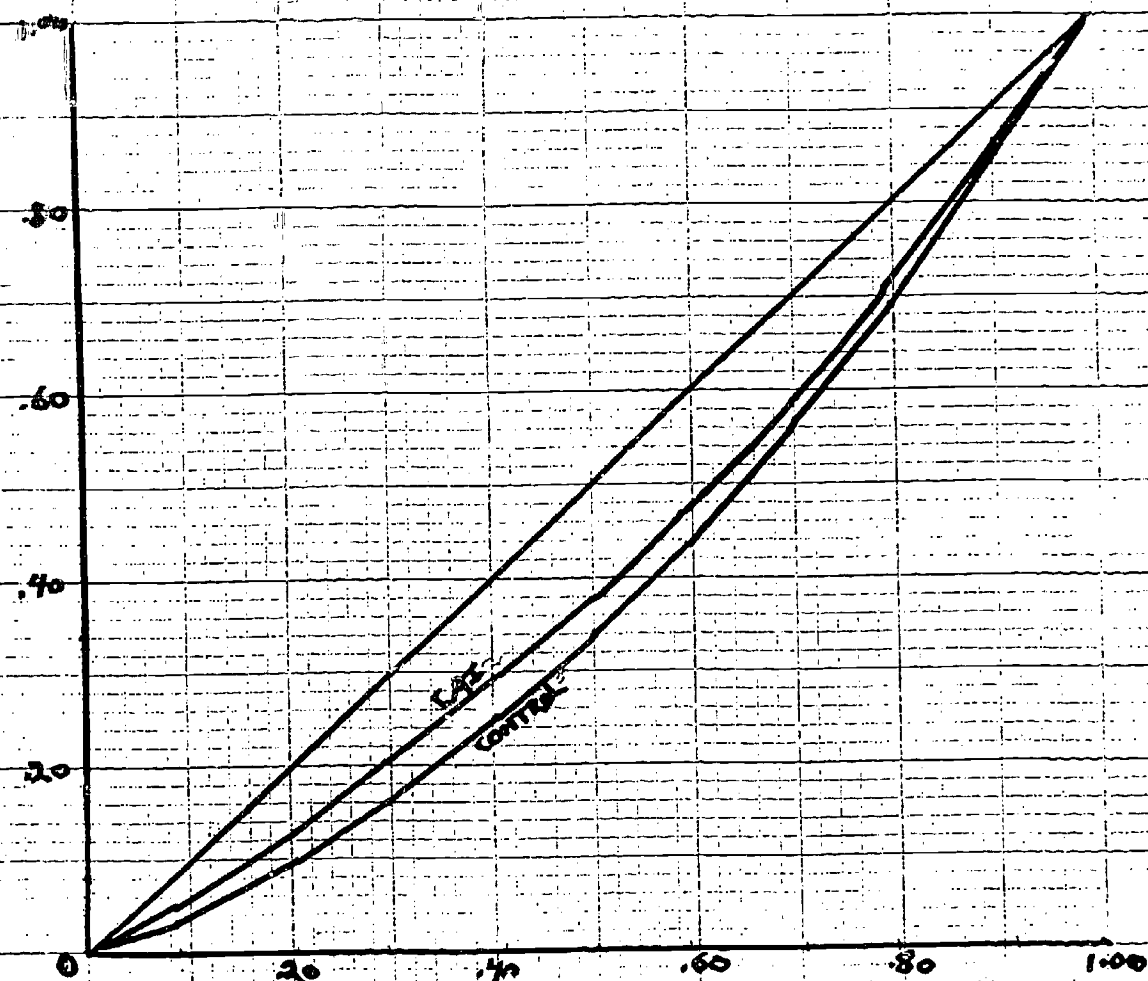

Figure III.1 - Typical Lorenz Curves: Distribution of Reading Posttest Achievement (COOP) the more inequitable is the distribution of achievement. The Gini coefficient is an aggregate measure of inequality that is defined as the ratio of the area between g(a) and the 45° line to the area between the 45° line and the abscissa. If the Gini coefficient is zero the distribution of achievement is completely uniform; the larger the Gini coefficient, the more unequal the distribution.

In order to examine the extent to which the different CAI programs described in Section II of this paper were in fact inequality reducing, we computed Gini coefficients for the distribution of achievement before and after the CAI was made available for both the CAI and the control groups. In Table III.8 these Gini coefficients are presented

---

Insert Table III.8 about here

---

for both the high school level computer programming course and the elementary arithmetic course in Mississippi and California grades 1-6. For each group at each grade level we give the Gini coefficients for the pretest for the group as a whole, the Gini coefficients for the posttest for the group as a whole, and the difference between those two Gini coefficients. Similar information is given for the control group. In the final column of the table the difference between columns 3 and 6 of the table is shown; if this difference is positive, it indicates that there is more of a reduction in inequality in the CAI group than in the control group. For the high school CAI group we computed the Gini coefficients for both raw scores and grade placement scores and the differences between those two computations can be seen in the table. We applied a sign test to the 12 arithmetic cases and the 2 computer programming cases that used grade placement scores to test the significance of the hypothesis that inequality was reduced more in the CAI groups than in the control groups. From column 7 of Table III.8 it can be seen that in only 3 of the 14 cases was the CAI less inequality reducing than no CAI. The sign test then implies an acceptance of the hypothesis that CAI is inequality reducing at the .05 level.

Table III.8 - Gini Coefficients for CAI and Control Groups

Group	CAI PRE	CAI POST	CAI PRE-POST	Control PRE	Control POST	Control PRE-POST	(CAI Pre-Post) - (Control Pre-Post)
Computer Programming							
SAT COMP R.S.[a]	.113	.087	.026	.108	.096	.012	.014
SAT APPL R.S.[b]	.119	.111	.008	.084	.097	-.013	.021
SAT COMP G.P.[c]	.079	.066	.013	.075	.070	.005	.008
SAT APPL G.P.[d]	.080	.079	.001	.059	.069	-.010	.011
Math Drill[e] and Practice							
Miss. 1967-68							
Grade 1	.057	.067	-.010	.037	.062	-.025	.015
2	.064	.039	.025	.055	.050	.005	.020
3	.016	.032	-.016	.035	.038	-.003	-.013
4	.080	.053	.027	.084	.065	.019	.008
5	.095	.070	.025	.078	.079	-.001	.026
6	.068	.077	-.009	.078	.084	-.006	-.003
Calif. 1967-68							
Grade 1	.058	.077	-.019	.054	.075	-.021	.002
2	.075	.056	.019	.073	.062	.011	.008
3	.042	.063	-.021	.050	.060	-.010	-.011
4	.067	.053	.014	.065	.058	-.007	.007
5	.056	.048	.008	.055	.068	-.013	.021
6	.077	.073	.004	.065	.070	-.005	.009

[a]Gini coefficients from Stanford Achievement Test, Computation subscale, raw scores.

[b]Gini coefficients from Stanford Achievement Test, Applications subscale, raw scores.

[c]Gini coefficients from Stanford Achievement Test, Computation subscale, grade placements.

[d]Gini coefficients from Stanford Achievement Test, Application subscale, grade placements.

[e]Gini coefficients for all math drill and practice from Stanford Achievement Test, Computation subscale, grade placements.

In Table III.9 we show the Gini coefficients for CAI and control

----

Insert Table III.9 about here

----

groups for the various sections of the reading achievement posttests. We do not include the pretest scores since different tests were used and the results are thus not directly comparable. In all 7 cases in Table III.9 the Gini coefficient is less for the CAI group than for the control group; the hypothesis that CAI is inequality reducing is substantiated in this case at the .01 level. The widely held subjective impression that no students in the reading CAI groups are "lost" seems, then, to be strongly supported by these data. It is reasonable to expect that the effect of CAI on posttests would correlate positively with the Gini coefficient differences obtained from the CAI and non-CAI subjects. The difference in Gini coefficients should be greatest where the CAI treatment is greatest and this seems to be the case. The effect of CAI is statistically significant on the S/WR, S/PM and S/WS, and for these subtests the Gini coefficient differences is fairly large. There is only a slight positive effect of CAI in the S/VOC, and the Gini coefficient differences for this subtest is correspondingly small.

Value explicit measures of inequality. In this part we will consider a measure of inequality proposed by A. Atkinson (1970) that makes explicit the value judgment entering into the comparison of the inequality of two distributions. Atkinson draws, in his discussion of greater and lesser inequality, on a close parallel between the concept of greater risk (or greater spread) in a probability distribution and the concept of greater inequality in a distribution of income. He is thus able to directly transfer certain results concerning the ordering by riskiness of probability distributions to ordering by degree of inequality of income distributions. He shows that a variety of conventional measures of inequality--including variance, coefficient of variation, relative mean deviation, Gini coefficient, and standard deviation of logarithms--would not necessarily be consistent with the ordering induced by concave utility functions.

Table III.9 - Gini Coefficients for Reading Achievement Posttests[a]

	CAI	Control	Control-CAI
SAT	.134	.174	.040
COOP	.183	.266	.083
DF	.068	.152	.084
S/WR(1)	.140	.209	.069
S/PM(2)	.226	.396	.170
S/WS(3)	.119	.149	.030
S/VOC(4)	.170	.183	.013

[a]Due to careful matching of CAI and control groups by pretest achievement (on the Metropolitan Readiness Test - see Section III.A), pretest Gini coefficients are not shown.

That is, one can in general find a concave utility function that would be inconsistent with the ordering induced by any of the above measures.

Atkinson then proposes that the overall utility, W, of a distribution of achievement scores, $N(u)$, be represented by the following formula:

$$W = \int_0^{\bar{u}} U(u)\, N(u)\, du \quad ,$$

when $\bar{u}$ is the maximum score achieved on the test. It is assumed in the above that $U(u)$ is increasing and concave, i.e., that $U'(u)$ is greater than 0 and that $U''(u)$ is less than 0. The concavity implies, for that particular population, that there is an aversion to inequality. Given this aversion to inequality there will exist a level of achievement, u_e, that is lower than the average level of achievement in the population under consideration such that if everyone in the population had exactly a u_e level of achievement, the overall level of social welfare would remain constant at W. Following Atkinson we will call u_e the "equally distributed equivalent" level of achievement. Clearly, u_e will in general depend on the form of U; however, by direct analogy with the theory of choice under uncertainty, u_e is invariant with respect to positive linear transformations of U.

If μ is the average level of achievement in the society, then a reasonable measure of inequality, I, is given by the following formula:

$$I = 1 - \frac{u_e}{\mu} \quad .$$

The lower I is, the more equal is the distribution of achievement; to put this another way, as u_e gets closer to μ, the "cost" of having inequality gets lower. The measure I ranges between 0 for complete equality and 1 for complete inequality and tells us, in effect, by what percentage total achievement could be reduced to obtain the same level of W if the achievement level were equally distributed.

In order to apply the measure I we need to have an explicit formulation of U. In this paper we consider two classes of functions of U. The first of these is one suggested by Atkinson that has the property of "constant relative inequality aversion." By constant relative inequality aversion it is simply meant that multiplying all achievement levels in the distributions by a positive constant does not alter the measure I of inequality. If there be constant relative inequality aversion it is known from the theory of risk aversion that U(u) must have the following form:

$$U(u) = a + b\,\frac{u^{1-\epsilon}}{1-\epsilon} \quad \text{if } \epsilon \neq 1, \text{ and}$$

$$U(u) = \ln(u) \quad \text{if } \epsilon = 1 \quad .$$

Another possibility that Atkinson considers is that of constant <u>absolute</u> inequality aversion, by which it is meant that adding a constant to each achievement level in the distribution does not affect the measure of inequality. A theorem of Pfanzagl (1959) can be used to show that if there is constant absolute inequality aversion then U(u) must have one of the following two forms:

$$U(u) = au + b \quad , \text{ or}$$

$$U(u) = a\lambda^{u} + b \quad .$$

Strict concavity implies the latter of these two and that $0 < \lambda < 1$.

We thus have two families of utility functions, one indexed by ϵ and the other by λ , which between them would seem to include a large number of qualitatively important alternatives for U . In Figure III.2 U(u) is shown for several values of ϵ and in Figure III.3 U(u) is shown for several values of λ .

Insert Figures III.2 and III.3 about here

Since transforming the functions depicted in Figures III.2 and III.3 by a positive linear transformation does not affect the measure I ,

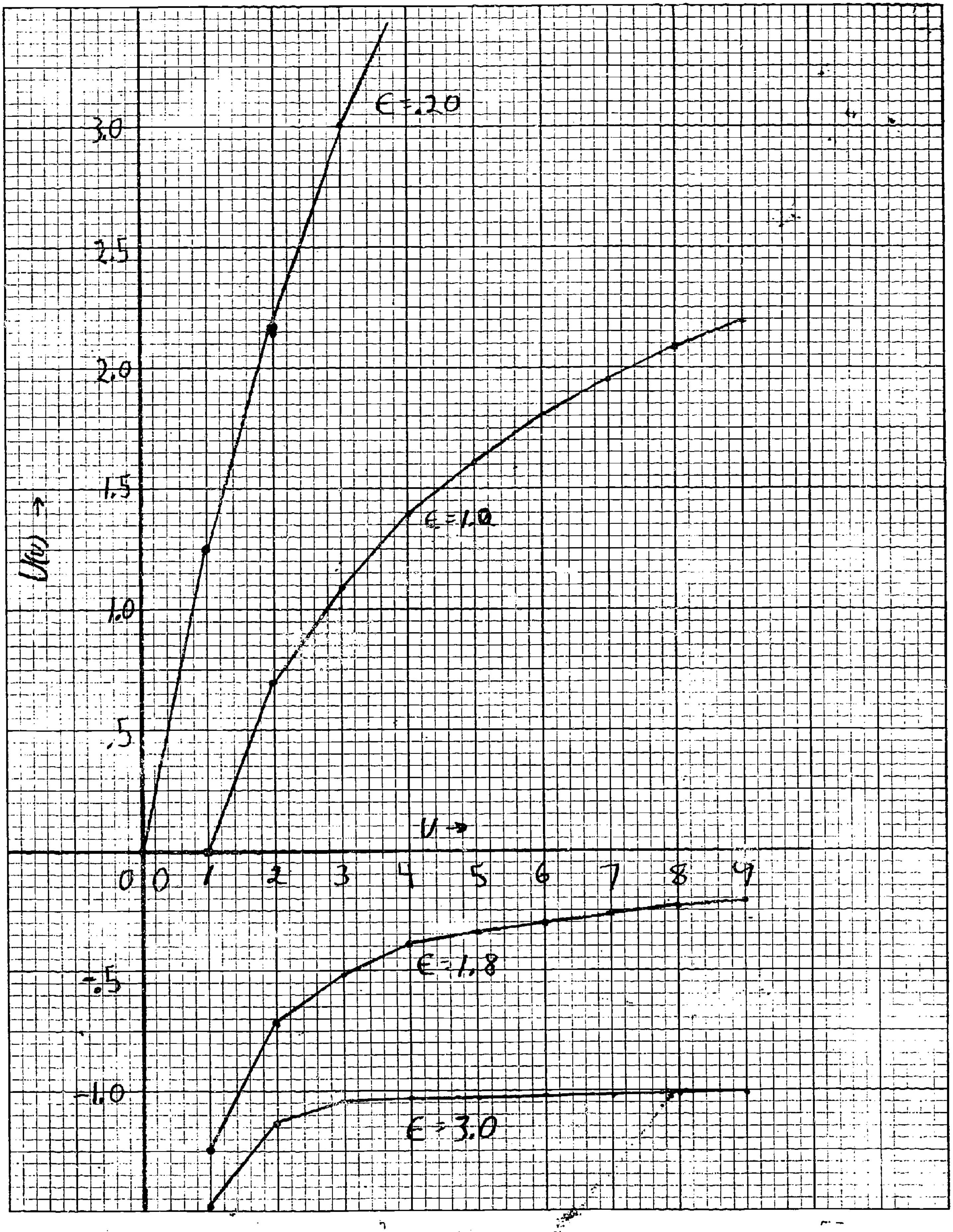

Fig. III.2 - U(u) for several values of ϵ.

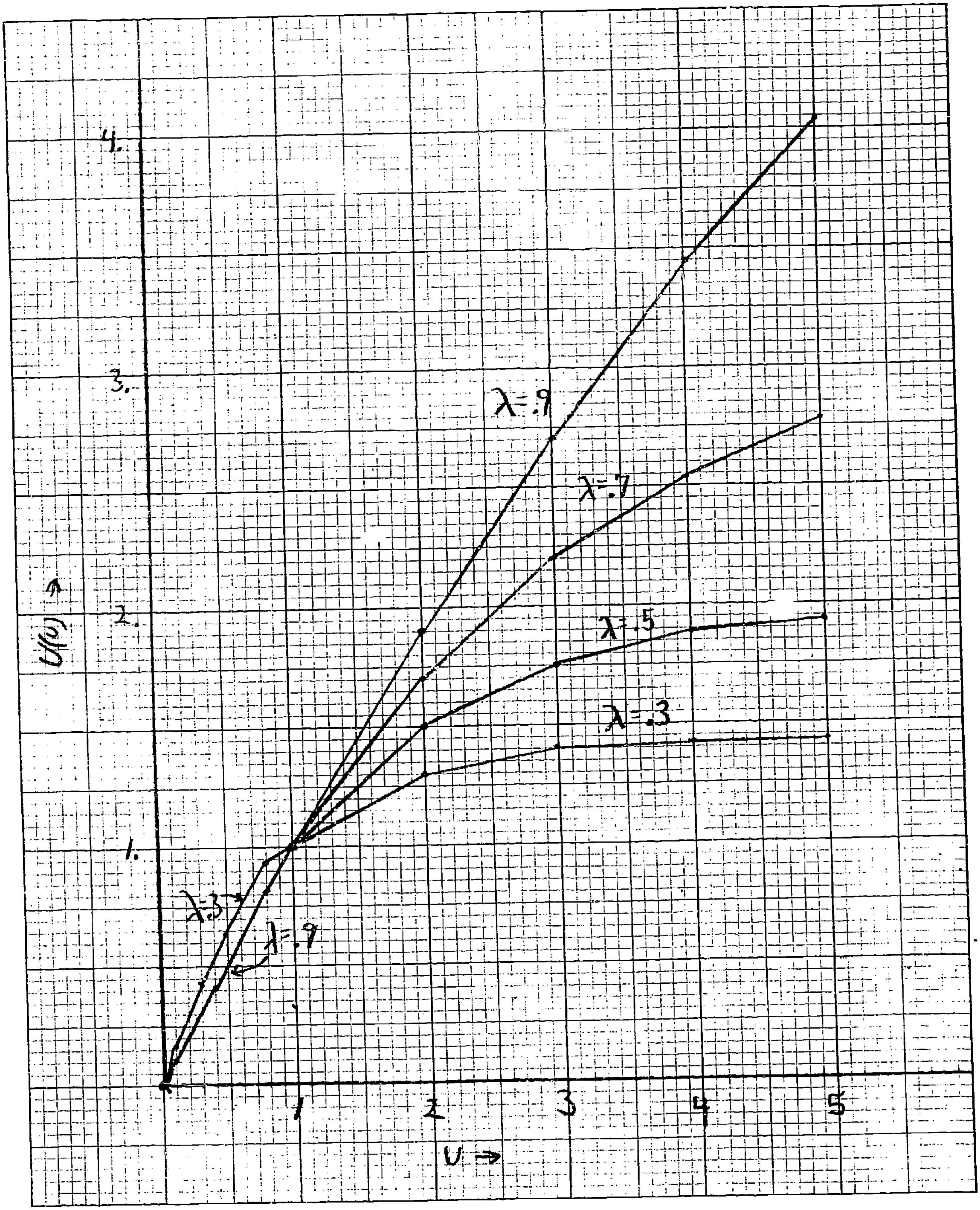

Fig. III.3 - U(u) for several values of λ.

the height and location of the functions in those two figures is arbitrary.

It is clear from the preceding that the measure I of inequality for any fixed distribution of achievement will vary with ϵ or λ. In Figure III.3 we have constrained $U(u)$ to pass through 0 and 1 for all values of λ implying that $U(u) = (1 - \lambda^u)/(1-\lambda)$. For λ very close to 1 inequality is close to 0; as λ gets smaller and smaller then inequality will get larger for any fixed distribution. The way in which I varies with ϵ is just the opposite; low values of ϵ give a low measure of inequality whereas large values of ϵ give large values for I.

In Figures III.4 and III.5 I is plotted as a function of ϵ and as a

---

Insert Figures III.4 and III.5 about here

---

function of λ for one particular CAI group and its control. The distributions $N(u)$ are of posttest scores and they are for a case where there was little difference in inequality on the pretest as measured by the Gini coefficients of the CAI and control groups.

One of the reasons it is of value to have a measure of inequality indexed by some parameter describing degree of inequality aversion (such as λ or ϵ) is that it is possible that the control group may be judged to be more equal for some values of λ and ϵ but less equal for others. In Table III.10 one can look for such reversals as a function of ϵ under the assumption of constant relative inequality aversion. Table III.11 shows the same information as a function of λ. The captions on those tables make them self-explanatory.

Insert Tables III.10 and III.11 about here

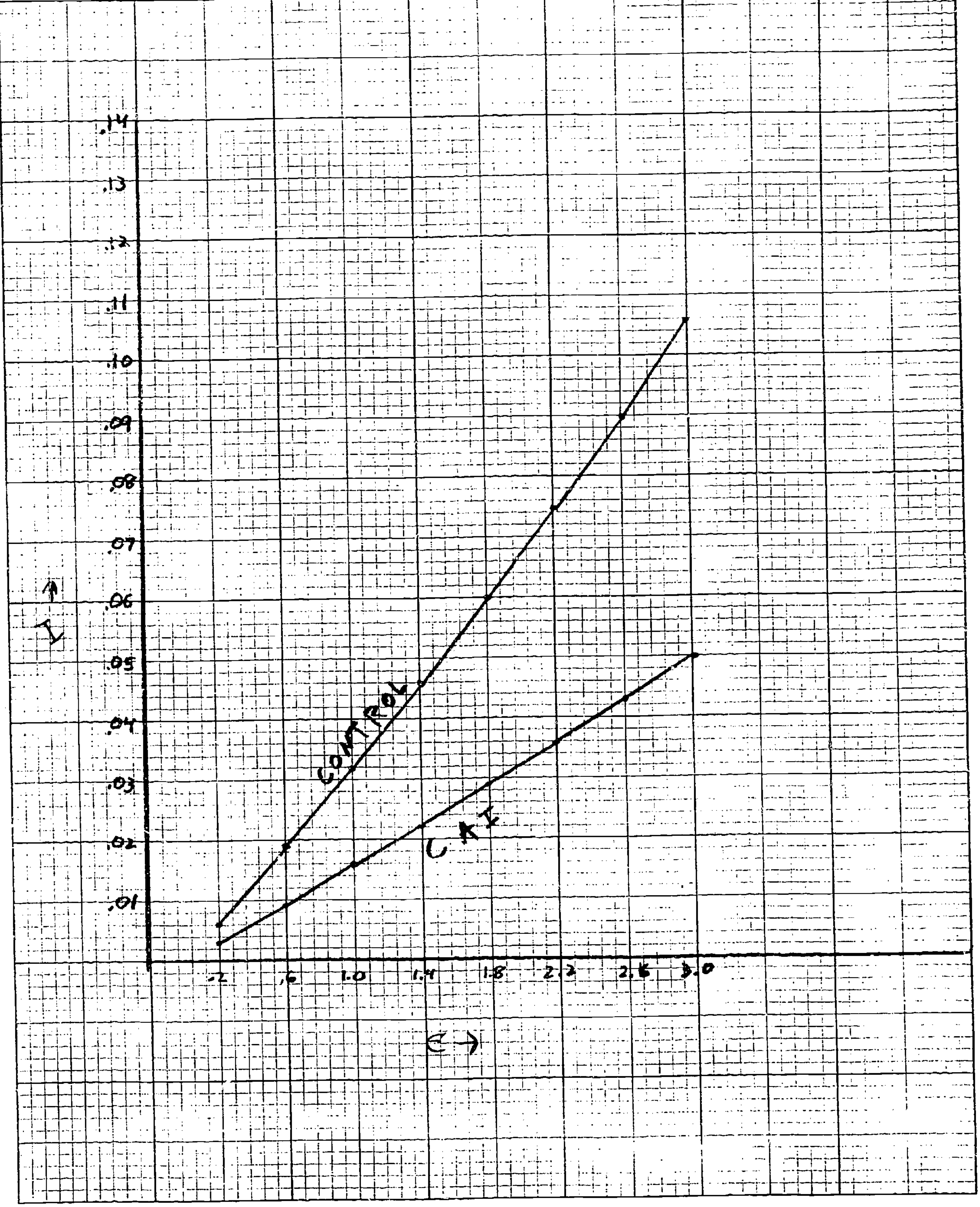

Fig. III.4 - I as a function of ϵ for fifth grade arithmetic, California, 1967-68.

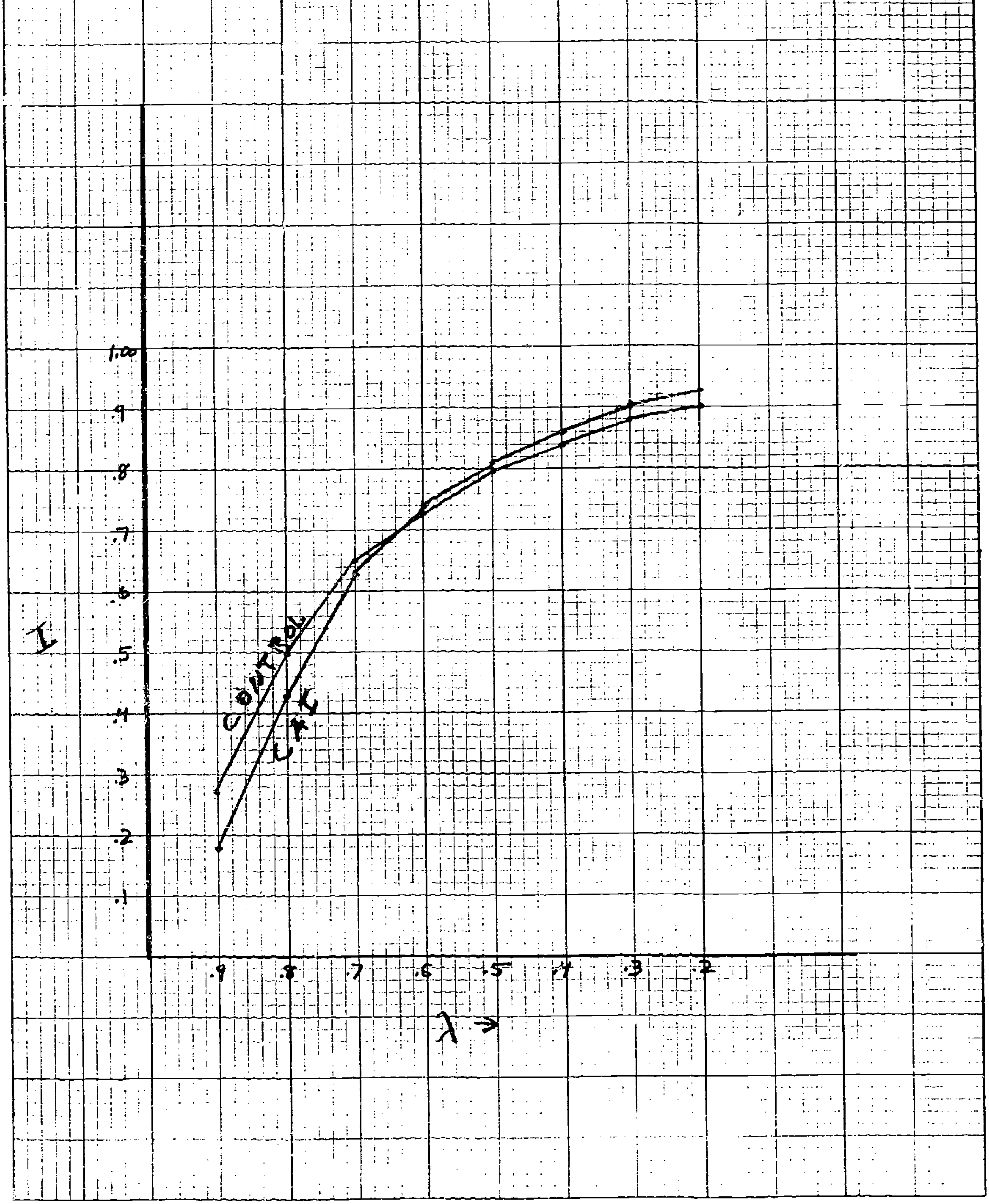

Fig. III.5 - I as a function of λ for fifth grade arithmetic, California, 1967-68.

Table III.10 - CAI Inequality Reduction: Constant Relative Inequality Aversion[a]

Student Group (Math Drill and Practice)		ϵ .20	.60	1.0	1.4	1.8	2.2	2.6	3.0
Miss. 67-68									
Grade	1	.001	.002	.004	.005	.006	.007	.007	.007
	2	.004	.012	.020	.030	.041	.054	.068	.084
	3	-.002	-.005	-.008	-.012	-.015	-.019	-.024	-.029
	4	.002	.005	.009	.014	.020	.028	.038	.050
	5	.005	.012	.019	.023	.026	.027	.025	.022
	6	.000	-.002	-.003	-.004	-.006	-.007	-.009	-.010
Calif. 67-68									
Grade	1	.000	.000	.000	.000	.000	.001	.002	.002
	2	.002	.004	.007	.009	.011	.014	.016	.019
	3	-.002	-.006	-.010	-.015	-.021	-.027	-.035	-.045
	4	.004	.001	.001	.000	-.003	-.007	-.013	-.022
	5	.003	.010	.017	.025	.034	.042	.052	.062
	6	.002	.006	.010	.015	.022	.030	.039	.051

[a]The numbers shown in the table are $I_A - I_B$ as a function of ϵ. I_A is the difference in inequality between CAI and control after treatment (i.e., on the posttest) and I_B is the difference before treatment. If the difference is greater after treatment than before, CAI is inequality-reducing.

Table III.11 - CAI Inequality Reduction: Constant Absolute Inequality Aversion[a]

Student Group (Math Drill and Practice)	λ: .90	.80	.70	.60	.50	.40	.30	.20
Miss. 67-68								
Grade 1	-.001	-.005	-.009	-.011	-.013	-.005	.011	.030
2	.010	.041	.090	.127	.146	.148	.139	.120
3	-.131	-.180	-.237	-.297	-.331	-.331	-.300	-.246
4	-.013	.016	.050	.054	.044	.033	.024	.017
5	.048	.006	-.010	-.007	.000	.004	.009	.016
6	-.083	-.108	-.098	-.078	-.060	-.046	-.037	-.030
Calif. 67-68								
Grade 1	.032	.069	.086	.086	.081	.078	.077	.076
2	-.018	-.038	-.041	-.031	-.020	-.012	-.006	.001
3	-.078	-.116	-.158	-.173	-.160	-.246	-.118	-.096
4	.050	.044	.012	-.010	-.024	-.031	-.033	-.036
5	.092	.071	.021	.002	-.004	-.004	-.005	-.006
6	-.020	.045	.045	.038	.034	.031	.029	.027

[a]The numbers shown in the table are $I_A - I_B$ as a function of λ. I_A is the difference in inequality between CAI and control after treatment (i.e., on the posttest) and I_B is the difference before treatment. If the difference is greater after treatment than before, CAI is inequality-reducing.

We have in this subsection attempted to provide explicit measures of the extent to which the three types of CAI programs that we review are inequality-reducing. We have used the recent work on measurement of inequality that has appeared in the economics literature to show that, ultimately, measurement of inequality rests on either an implicit or explicit value judgment. We have shown measures of inequality for CAI and control groups for several explicit classes of value judgments concerning distribution of achievement. It is perhaps worth stressing that as we were actually designing and implementing our CAI programs we did not have inequality-reduction in mind as an explicit goal; our results, literally, just turned out this way.

The next step to take at this point is, we feel, to try to design patterns of presentation of CAI to students that are optimal by some utility function U maximized subject to a variety of constraints. One sort of constraint would be the distribution of prior achievement in the population we are providing this CAI to; another constraint would be the total number of terminal hours per month available to that population of students; still another possible class of restraints would be possible impositions from the school system administration that no students get less than a certain amount of CAI or more than a certain amount of CAI per day on an average; and a final fundamental constraint would be the production function that relates time on the system and other factors to gains in student achievement. What we plan to examine in the future is how the solution to this optimization problem varies as U varies when the various constraints vary. After so doing we will design patterns of instruction for students that are explicitly tailored to several separate Us and empirically examine the extent to which we are able to obtain the stated objectives. We hope that in this fashion any trade-offs that might exist between total achievement gain and inequality-reduction can be made very explicit both in terms of the underlying technology and the underlying value structure.

IV. COST OF COMPUTER-ASSISTED INSTRUCTION

A. General Considerations

It is useful to place CAI costs into three broad categories. The first category comprises the terminal equipment used by the students. Terminals vary in complexity from a simple teletype slightly modified to a CRT with keyboard, light pen, audio and random-access slide screen, and costs vary accordingly. The second cost category comprises the computer system that decides on and stores instructional presentations and evaluates student responses, and includes the central processing unit, disc and core storage, high-speed line units, and peripheral equipment. The final cost component is the multiplexing and communication system that links the student terminals to the main computer system. This communication system can be reasonably simple when the terminals are located within a few hundred feet of the computer. If the terminals are dispersed, the communication system may include a communication satellite as well as one or more small computers that assemble and disassemble signals.*

Up to this point, we have mentioned only the cost components necessary to provide CAI and have assumed that the curriculum to be used has already been programmed. It is only the cost of provision that we shall consider here. Of course, unless ways are found to share a single curriculum among many users, the per-student cost of curriculum preparation can be prohibitively high. Levien et al. (1970) discuss how to provide incentives and how to recoup costs for CAI curriculum preparation. Since a reasonably large body of tested curriculums already exists, we consider those costs sunk and will not include them here.

There appear to be two trends in design philosophy for the computer component of a CAI system. One trend is toward large, highly flexible

*Terminals now linked to the present Stanford CAI system are scattered over much of the United States; beginning in September, 1971 two clusters of 8 terminals each will be linked to Stanford via NASA's ATS-1 experimental communication satellite.

systems capable of simultaneously providing curricula in many subjects to a large number of simultaneous users. The other trend is toward small, special-purpose computer systems capable of providing only one or two curricula to a few students. A large, general-purpose computer system might have 500 or more student terminals simultaneously in use (the proposed PLATO IV system of the University of Illinois is aiming for 4,000); the small special-purpose system is apt to have 8 to 16 terminals. Naturally the number of terminals per computer has important implications for the communication system. In order to make a large system worthwhile, a reasonably extensive communication system is almost inevitable. On the other hand, even a moderate-sized elementary school could use a 16-terminal system, and only simple communications would be required. The potential scale economies of a large computer system, its broader range of offerings, and its easy updating must be balanced, then, against the lower communication costs of special-purpose systems.

Jamison, Suppes and Butler (1970) examined the cost of providing CAI in urban areas by way of a small special-purpose computer system, the first of which is now in operation in San Diego. Rather than review those costs here, we refer the reader to that paper. Costs per student per year are approximately $50 above the normal cost of educating the child, assuming that the school system in no way attempts to reduce other costs (by, for example, increasing the student-teacher ratio) as a result of introducing CAI.

B. Cost of Providing CAI in Rural Areas

The most distinctive aspects of providing CAI in rural areas are that the students to be reached are highly dispersed and would thus tend to be reasonably distant from a central computer. One could use small computers for rural areas at costs probably somewhat higher than Jamison, et al estimated for urban areas. To obtain the advantages of a large central system, however, the communication system must be rather sophisticated. In this section we examine the cost of providing large-scale CAI in rural areas. To obtain per-student annual-cost figures we examine each of the three cost areas mentioned above and then combine them to give the final figures. Our costs are based on the CAI system at IMSSS, using the curriculum already available; other systems could have different costs.

Terminal costs. The cost of a Model-33 teletype, including modifications, is about $850. To provide the teletype terminal with a computer-controlled audio cassette would increase the cost about $150, but since this is not operational now the additional $150 is not included in our estimates. An alternative would be to lease the teletypes--that cost is about $37 per teletype per month and includes maintenance.

Computer facility costs. Cost estimates are provided for a system capable of running about 1,000 students at a time. The system would be run at "4/5 diversity," i.e., 1,250 terminals would be attached to the system under the assumption that no more than 4/5 of the 1,250 would run at any one time. The assumption of 4/5 diversity is conservative given our past experience.

The system would comprise two PDP-10 computers, each with a 300m byte disc, 512K words of core memory, a swapping drum, and appropriate I/O and interfacing devices. The system would essentially be a doubled 500-terminal system; if, however, appreciably more terminals were desired, other designs would be appropriate.

Table IV.1 shows the initial costs of the system and Table IV.2 shows annual costs. Overhead is *not* included.

Insert Tables IV.1 and IV.2 about here

In order to express all costs as annual costs we multiply the $3,260,000 by .15, assuming about a ten-year equipment lifetime and 10 percent social discount rate. Thus the annual cost of the initial equipment purchase is about $490,000. When added to the direct annual costs, the total is $870,000 per year. With 1,250 terminals, the central facility cost is $690 per terminal per year.

Communication costs. In an unpublished paper, Jamison, Ball and Potter (1971) have examined in some detail the cost of communication between a central computer facility and rural terminals. They con-

Table IV.1 - Initial Costs, Computer Components of CAI System[a]

Component	Cost
Computer system	$2,560
Spare parts and test equipment	200
Planning and installation	350
Building	150
Total	$3,260

[a]Costs in thousands of dollars

Table IV.2 - Annual Costs, Computer Components of CAI System[a]

Component	Annual Cost
System operation	$150
System maintenance	175
Building maintenance	20
Supplies	35
Total	$380

[a]Costs in thousands of dollars

sidered two types of systems--one using commercial phone services and one using a single transponder on a communication satellite. Costs of communicating by way of satellite are independent of distance whereas phone costs are quite distance-dependent. Thus, for longer distances, satellites become increasingly attractive. Figures IV.1 and IV.2 taken from Jamison, Ball and Potter show the annual cost of communication and multiplexing for satellite and terrestrial systems. Both assume that the terminals are clustered in groups of eight. The graphs assume "best estimate" satellite and phone service costs in the 1975 time

---

Insert Figures IV.1 and IV.2 about here

---

frame and 8-year equipment lifetime with 10 percent cost of capital. They also include maintenance and system installation, but do not include overhead.

The present engineering cost estimates for G, the satellite ground-station cost, is $10,000 (but this is the estimate for a feasible, not optimal system--we expect much engineering improvement). Thus Figure IV.1 shows that the annual communication cost for a satellite distribution system would be about $800,000. From Figure IV.2 we see that if D, the average distance between the central computer facility and the terminals, exceeds about 550 miles then communication via satellite is cheaper than via telephone.* Since the average distance to the terminals is quite likely to exceed 550 miles, $800,000 is our estimate of communication and multiplexing cost. This comes to $640 per terminal per year.

Per-student costs. To obtain the annual cost of the terminal we multiply its purchase price ($850) by .15 to obtain $130 and add 10 percent of its purchase price to cover maintenance. The total is $215 per year. Teacher training must also be included and is typically a one-week course given at the school at a cost of about $500, plus transportation per person. Continuing our assumption of eight terminals

---

*A further, and very important, advantage of using satellites is that it eliminates the necessity of working with poorly equipped rural telephone services. IMSSS has experienced many delays and unexpected costs as a result of working with such services in Kentucky and elsewhere.

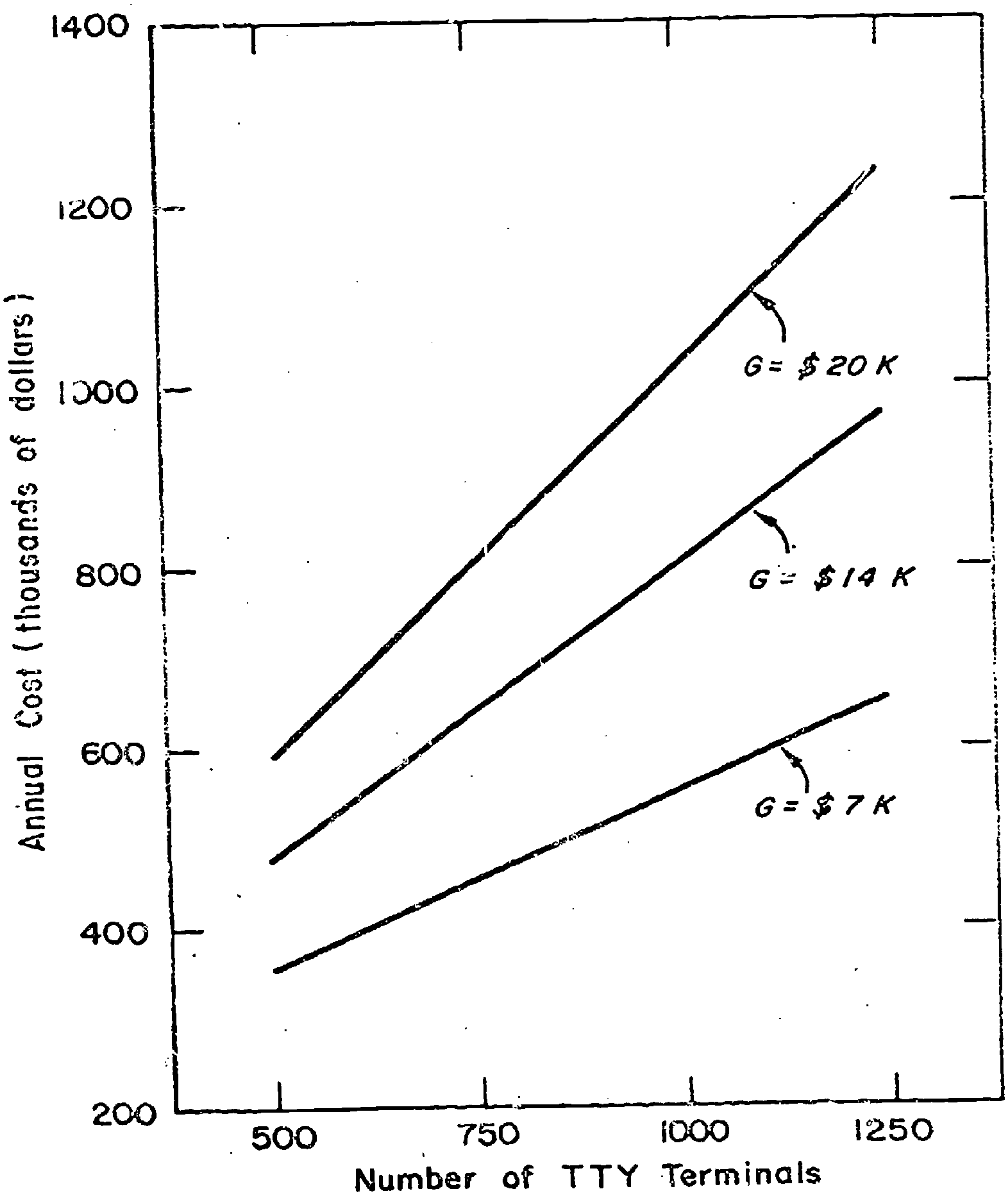

Fig. IV.1 Annual communication and multiplexing cost, satellite system.

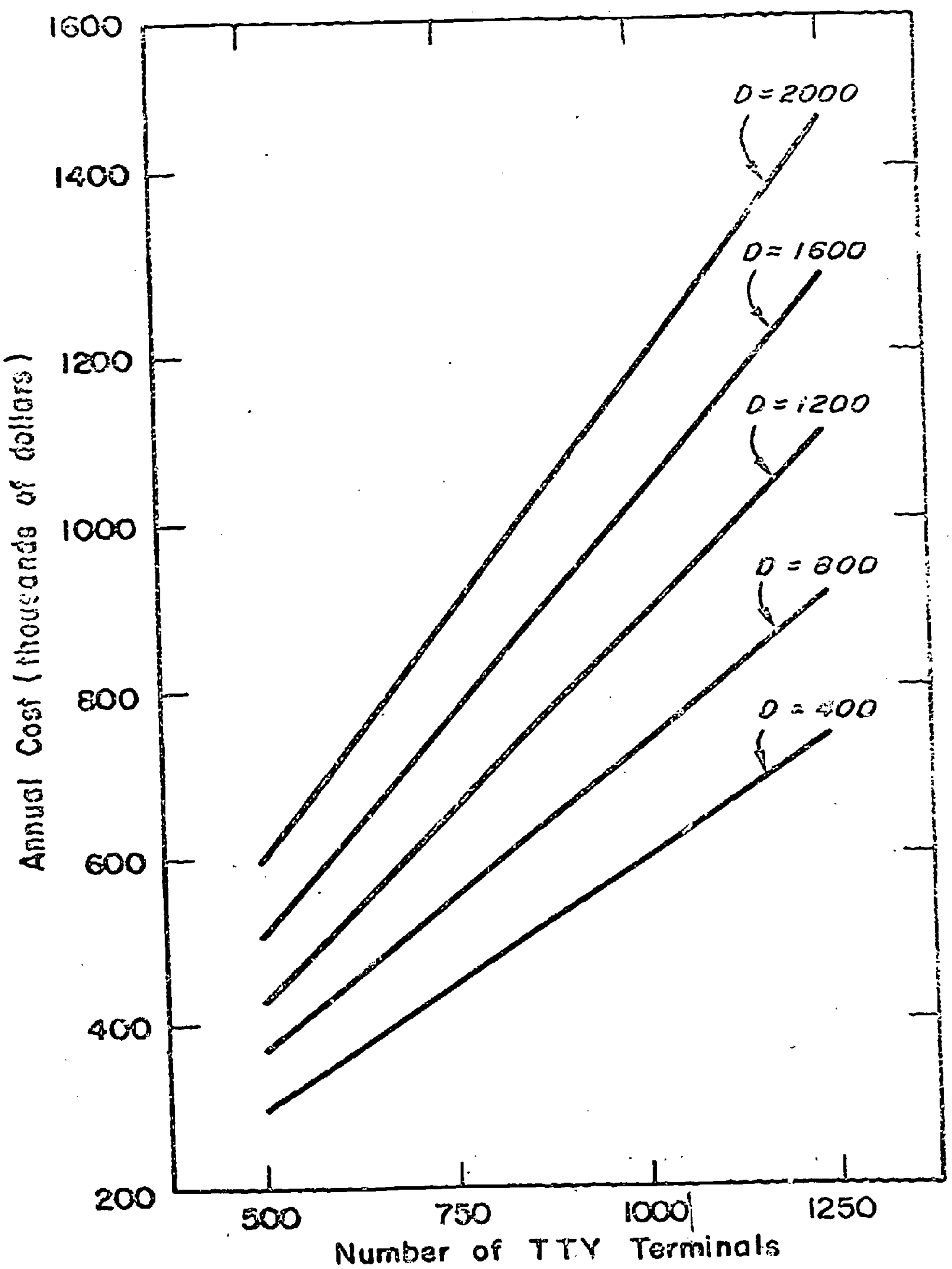

Fig. IV ? Annual communication and multiplexing cost, commercial telephone system.

per school, and assuming that the course will be repeated for at least four years and that transportation costs average $300 per session, the per-terminal annual charge of teacher training is $25. A final cost to be considered is that of the terminal room proctor. Much of this cost can usually be covered by volunteers and inexpensive help and would cost not more than $2,000 per school per year or $250 per terminal per year. We assume space available in the schools due to a declining rural population.

Table IV.3 shows the annual costs per terminal. A utilization rate of 25 students per terminal per day is typical with this sort of system so that the cost per student per year would be on the order of $75.

Insert Table IV.3 about here

Overhead costs might increase this to as much as $125. If the number of terminals per school were increased from eight to ten there would be no increase in communication and multiplexing, teacher training or proctoring costs, so our estimates are conservative in that respect.

Kiesling's (1970) estimates of 1970 costs for conventional compensatory education at about the quality provided by CAI are $200-$300 per student per year in urban and suburban areas. It would presumably be more expensive to provide this quality of compensatory education to rural areas, and salary inflation would also increase his estimates. We thus feel that CAI is a low-cost alternative for providing compensatory education to rural areas.

A possible pattern of development for rural compensatory education is to begin with satellite or long-line communications to a large central system, and then, after a cadre of experienced personnel has been trained, to convert to somewhat less expensive special-purpose systems located in the area.

C. Opportunity Cost of CAI

In the preceding discussion of cost we were estimating *ceteris paribus* costs of adding CAI to the school curriculum. We indicated

Table IV.3 - Annual Cost in 1975 of Rural CAI per Terminal

Item	Cost
Teletype terminal	$ 215
Computer facility cost	690
Communication and multiplexing	640
Teacher training	25
Proctoring	250
Supplies and miscellaneous	25
Total	$1,845

that the add-on costs of CAI were sufficiently less than those of alternative compensatory education programs so that, if additional funds were available for compensatory education, CAI appears a very attractive alternative. If add-on funds are unavailable--and this is apt to be the common case in the present financial environment--then CAI can be introduced only at the cost of providing less of some other school resource to the students. The amount of these other resources foregone represents, then, the opportunity cost of providing CAI to the school. As teacher costs comprise by far the largest component--on the order of 70%--of school costs, our purpose in this section is to examine what must be given up in terms of teacher resources in order to provide CAI for students.

The amount of teacher time required per child per year depends on average class size, average number of days per school year, and average number of class hours per school day. We assume that length of school day and length of school year are rather more fixed than average class size, and will examine only the effect on class size of introducing CAI. The other two variables could, however, be introduced in a straightforward way into the analysis.

Let the "instructional" cost per year for a class be the cost of its teacher's salary plus the cost of whatever CAI the class receives. Let S be the class size before CAI is introduced, T be the teacher's annual salary, and C be the cost per student per year of CAI, including all costs previously indicated in Table IV.3. We wish to compute A, the number of additional students in the class that are required to finance the CAI. With no CAI, the annual instructional cost for the class is T; with CAI, the cost is $T + C(S + A)$. We require that the per student cost with CAI be no greater than the cost without it, that is,

$$\frac{T}{S} = \frac{T + C(S + A)}{S + A} \quad .$$

Solving this equation for A we obtain:

$$A = CS^2 / (T - CS) \quad .$$

The partial derivatives of A with respect to T, C, and S are also of interest, and those are given below:

$$\frac{\partial A}{\partial C} = TS^2 / (T - CS)^2 ,$$

$$\frac{\partial A}{\partial S} = CS(2T - CS) / (T - CS)^2 , \text{ and}$$

$$\frac{\partial A}{\partial T} = -CS^2 / (T - CS)^2 .$$

Table IV.4 below shows A, $\partial A/\partial S$, $\partial A/\partial C$, and $\partial A/\partial T$ for C = \$50 (urban) and \$75 (rural) under the assumptions that T = \$11,000 and S = 26.

Insert Table IV.4 about here

A number of interesting points emerge from the table. First, even if C = \$75, the student to teacher ratio only goes from 26 to 31.6 in order to provide CAI. If the Coleman Report is correct in concluding that student performance is insensitive to student to teacher ratio, this would seem to be a quite attractive reallocation to the extent that it can be made politically feasible. Second, from the values for $\partial A/\partial C$ we see that a \$10 increase in C would require about a .8 increase in A if C is \$75. Third, from the value of $\partial A/\partial S$ we see that an increase of 1 in S causes an increase of .286 in A if C = \$50 but an increase of .477 if C = \$75. Finally, the last row in the table shows that a \$1,000 annual increase in teacher salary would decrease A by about .36 if C is \$50; it decreases A by almost twice that amount if C is \$75. In general the partial derivatives in the table seem quite sensitive to C.

We conclude this section on costs by observing that the cost of CAI seems to have decreased to the point that CAI is now quite attractive compared to alternative compensatory techniques with roughly similar performance. This holds whether one considers CAI as an add-on cost or as a substitute for teacher time.

Table IV.4 - Increment in Class Size Required to Finance CAI

Variable	Expression[a]	Cost of CAI per Student per Year: $50	$75
A	$CS^2/(T - CS)$	3.5	5.6
$\partial A/\partial C$	$TS^2/(T - CS)^2$	.079	.091
$\partial A/\partial S$	$CS(2T - CS)/(T - CS)^2$	.286	.477
$\partial A/\partial T$	$-CS^2/(T - CS)^2$	-.00036	-.00062

[a] S is initial class size and it is assumed to be 26; T is annual teacher salary and it is assumed to be $11,000; C is cost per student per year of CAI and A is the increment in class size required to finance CAI if there are to be no increases in per student annual costs.

REFERENCES

Atkinson, A. B. On the measurement of inequality. Journal of Economic Theory, 1970, 2 221-224.

Atkinson, R. C. Computerized instruction and the learning process. American Psychologist, 1968, 23, 225-239.

Atkinson, R. C., Fletcher, J. D., Chetin, H. C., and Stauffer, C. M. Instruction in initial reading under computer control: The Stanford project. Educational Technology Research Reports Series. Englewood Cliffs, N. J.: Educational Technology Publications, 1971.

Boneau, C. A. The effects of violations of assumptions underlying the t-test. Psychological Bulletin, 1960, 57, 49-64.

Bowles, S., and Levin, H. The determinants of scholastic achievement: An appraisal of some recent evidence. Journal of Human Resources, 1968, 3, 3-24.

Chall, J. Learning to read: The great debate. New York: McGraw-Hill, 1967.

Coleman, J., et al. Equality of Educational Opportunity. Washington, D. C. USGPO, 1965.

Coleman, J. The evaluation of equality of educational opportunity. Santa Monica, Calif.: The Rand Corporation, P-3911, 1968.

Elashoff, R. M. Effects of errors in statistical assumptions. International Encyclopedia of the Social Sciences. New York: Macmillian, Inc., 5 1968, 132-142.

Fletcher, J. D., and Atkinson, R. C. An evaluation of the Stanford CAI program in initial reading (grades K through 3). Technical Report No. 168, March 12, 1971, Stanford University, Institute for Mathematical Studies in the Social Sciences.

Jamison, D., Ball, J., and Potter, J. Communication economics of computer-assisted instruction for rural areas: Satellite vs. commercial phone systems. To be presented at the XXIInd International Astronautical Congress of the International Astronautical Federation, Brussels, September 1971.

Jamison, D., Suppes, P., and Butler, C. Estimated costs of computer-assisted instruction for compensatory education in urban areas. Educational Technology, 1970, 10, 49-57.

Kelley, T. L., Madden, R., Gardner, E. G., and Rudman, H. C. Stanford Achievement Test: Directions for administering. New York: Harcourt, Brace, Jovanovich, 1964.

Jensen, A. How much can we boost I. Q. and scholastic achievement? Harvard Educational Review, 1969, 39, 1-123.

Kiesling, H. A study of successful compensatory education programs in California. March, 1970, internal Rand Corporation Document.

Levien, R., et al. The emerging technology, instructional uses of the computer in higher education. The Rand Corporation, unpublished report, October 1970.

Levine, D. B., and Singer, N. M. The mathematical relation between the income density function and the measurement of income inequality. Econometrica, 1970, 38, 324-330.

Lorton, P., and Slimick, J. Computer-based instruction in computer programming--A symbol manipulation--list processing approach. Proceedings AFIPS 1969 Fall Joint Computer Conference, 1969, 33, 535-543.

Michelson, S. The data: Equal protection and school resources. Inequality in Education, No. 2, 1970, 4-16.

Michelson, S. The political economy of public school inequalities, 1971. To appear in M. Carnoy (Ed.), The Political Economy of Education: New Approaches to Old Problems.

Mood, A. (Ed.) Do teachers make a difference? Washington. D. C.: U. S. Government Printing Office, 1970.

Newbery, D. A theorem on the measurement of inequality. Journal of Economic Theory, 1970, 2, 264-266.

Pfanzagl, J. A general theory of measurement--applications to utility. Naval Res. Logistics Quart., 1959, 6, 283-294. Reprinted in R. Luce, R. Bush, and E. Galanter (Eds.), Readings in Mathematical Psychology, Volume II. New York: John Wiley, 1965.

Piccariello, H. Evaluation of Title I. Unpublished paper, 1969.

Suppes, P., and Morningstar, M. Computer-assisted instruction. Science, 1969, 166, 343-350.

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 Problem of Society. Washington, D.C.: American Psychological Association, Inc., 1970. Pp. 221-236.

Suppes, P., Jerman, M., and Brian, D. _Computer-assisted instruction: The 1965-66 Stanford arithmetic program_. New York: Academic Press, 1968.

Zinn, K. L. List of projects using computers in instruction. Unpublished paper, 1969.