In order to assist decision makers in considering different approaches to improving mathematics and reading performance of elementary school children, a cost-effectiveness study was undertaken of computer-assisted instruction (CAI) and three other interventions. In general, peer tutoring is found to be more cost-effective than CAI, and both are more cost-effective than reducing class size or increasing the length of the school day. A discussion of the cost-effectiveness methodology and its application to educational interventions is stressed.

COST-EFFECTIVENESS OF COMPUTER-ASSISTED INSTRUCTION

HENRY M. LEVIN
Stanford University
GENE V. GLASS
Arizona State University
GAIL R. MEISTER
Research for Better Schools, Inc., Philadelphia

A concern about the overall quality of American education as well as its consequences for U.S. competitiveness in the international economy has led to numerous calls for educational reform (National Commission on Excellence in Education, 1983; Task Force on Education for Economic Growth, 1983). The push for computers in instruction has been central among strategies for educational reform. Computer-assisted instruction (CAI) has grown more visible as an instructional technique for several reasons. Since 1980 the price of microcomputers has fallen by about 50%, and capabilities have risen sharply; concomitantly, there has been a surge in the availability of educational software for teaching such courses as programming, foreign languages, logic, music, design, and mathematics as well as for providing supplementary instruction through drill and practice. These developments and the proliferation of computers in the work place have

EVALUATION REVIEW, Vol. 11 No. 1, February 1987 50-72

contributed to the widely held view that CAI adoption will produce increases in educational productivity of schools.

Before placing such heavy reliance on CAI, however, it is important to have information on its cost-effectiveness compared to other alternatives. Unfortunately, pertinent cost-effectiveness data are not available. The purpose of this article is to fill that gap by providing estimates of the cost-effectiveness of CAI and three other educational interventions for improving mathematics and reading achievement in the elementary school. The remainder of this section will describe the four educational alternatives that are evaluated, and the next section will develop estimates of costs and effects. The final section will evaluate the cost-effectiveness of the alternatives and its implications.

Four educational interventions were chosen for the cost-effectiveness comparison: CAI, cross-age tutoring, reduction in class size, and increases in daily instructional time. Each represents an important policy alternative. For three of the four, a specific study was chosen that was representative of the intervention in general and was typical of its effectiveness. The exception to selection of a single study was for the reduction in class size, where effects were determined by meta-analysis.

The studies on which the cost-effectiveness analyses were based met certain criteria. Each was considered representative of one of the classes of intervention chosen for the study; each appeared to be replicable in other school settings; each showed effect sizes that fell near the typical effectiveness found in many studies of the intervention; each provided clear descriptions of the intervention and its ingredients; and each was accompanied by a careful and reliable evaluation. A description of each of the interventions follows, as well as the specific study that was used for more precise cost and effectiveness analysis.

COMPUTER-ASSISTED INSTRUCTION

Although computer-assisted instruction has been available for at least two decades, the recent drastic decline in costs and sharp increase in capability of microcomputers have engendered a large expansion in the use of computers for instruction. Typical applications of CAI include drill and practice (exercises to reinforce conventional classroom instruction) as well as the teaching of specific subjects, especially programming (Center for Social Organization of Schools, 1983).

Few evaluations of the effects of CAI over a full academic year or longer have been undertaken, however. Evaluations of the effects of CAI on mathematics and reading achievement are generally limited to drill and practice. The costs and effects of the CAI model evaluated in this study conforms to the drill and practice approach as set out by one of the pioneers in the field, the Computer Curriculum Corporation (CCC). The advantages of selecting this particular approach for a cost-effectiveness comparison are that it is one of the most common and historically well-established applications of CAI and that it has been the subject of one of the best instructional evaluations.

The specific CAI approach that we have used to construct cost-effectiveness data was sponsored by the Educational Testing Service and Los Angeles Unified School District (ETS/LAUSD) in 1976-1980 with funding from the National Institute of Education (Ragosta et al., 1982). Elementary students were given 10-minute daily sessions of drill and practice in mathematics, reading, and language arts. Some students had more than one daily session, and the combinations of subjects to which students were assigned differed so that a child studying reading and language arts by computer could serve as a control for assessing the benefits of mathematics instruction by another child studying reading, language arts, and mathematics ("within-group" controls). Since the experiment ran for four years, it was also possible to make comparisons among students with up to four years of CAI and with different combinations of subjects as well as between students who received CAI and those who did not.

The approach evaluated in the ETS/LAUSD study uses a separate classroom with 32 terminals that are connected to a minicomputer. (A similar type of delivery system can be constructed using personal or microcomputers that are arranged in a network with a hard-disk storage device.) The minicomputer holds all computer curricula for all elementary grades and curriculum areas as well as student records on the number of sessions that students have taken and their progress.

Students sign in at their terminals and begin the session where they left off in the previous session. A problem is displayed, typically in a multiple-choice or a fill-in-the-blank format. The student responds, and a message on the display indicates if the answer is correct, followed by a new problem. When a student achieves proficiency on a particular part of the curriculum—as evidenced by some preset proportion of correct answers—the system provides either problems of the same type at a higher level of difficulty or a new type of problem. The curriculum is not designed to introduce new curricular material as much as it is to provide an opportunity to apply concepts and practice tasks that have already been taught in the regular classroom.

CROSS-AGE TUTORING

Cross-age or peer tutoring has a long informal history in American education. In one-room schools, older students routinely helped teach younger students. A compendium of reported benefits of successful peer tutoring efforts includes achievement gains, increases in self-esteem, and enhancement of academic motivation, often for tutors and tutees (Ehly and Larsen, 1980: 12-17, 21-23). The policy importance of tutoring turns on the fact that when a child or paraprofessional, instead of a certificated teacher, fulfills this role, the individualized instruction that results costs less. Perhaps even more important, benefits are expected for both peer tutor and tutee.

The cross-age tutoring intervention used in this study is based on the Cross-Age Structured Tutoring Program for Reading and Mathematics in the Boise (Idaho) Schools (Independent School District of Boise City, 1983a, 1983b, 1983c). For a school of about 300 to 400 students, the full tutoring program relies on four paraprofessionals—an adult tutor manager in reading, an adult tutor manager in math, an adult tutor in reading, and one in mathematics—and 60 upper-grade student tutors who provide tutoring for second and third grade children needing help in reading and mathematics. The adult tutors and tutor managers are trained and supervised by a Tutoring Program Specialist, a central office administrator responsible for 14 schools. Student tutors at each site are trained and supervised by the tutor manager in each subject. Typically, a tutor manager oversees 30 tutoring pairs and tutors 2 additional students directly, and an adult tutor works regularly with 12 or 13 individual upper-grade tutees. Only on occasion, when tutors are absent or tutees need special help, do adults provide tutoring to lower-grade students. Thus one school in the range we are considering hosts 60 student tutors and their 60 tutees, as well as 30 other tutees who work with the adult tutors and tutor managers, for a total of 150 children participating in the tutoring program. The effects of the Boise tutoring program, then, reflect principally those of peer tutoring for students in grades 2 and 3, and adult tutoring for students in grades 4, 5, and 6.

All tutors use a commercially available curriculum, which includes a manual for each adult in each subject (as well as an audiotape in reading). Student tutors are trained with a locally produced manual. As part of their work with tutees, they distribute locally purchased awards and certificates. Tutoring takes place in otherwise unused space around the school, such as an available classroom, hallways, a cafeteria, or a small office. Tutoring sessions conducted by both adult and student tutors last approximately 20 minutes a day.

REDUCING CLASS SIZE

One of the oldest methods thought to improve educational outcomes is reduction of class size. The reduction of class size is not an intervention that is designed to increase achievement directly, however. Rather, it is expected to influence what goes on in the classroom, how teachers interact with students, and what the students themselves do or are allowed to do. The differences in classroom processes resulting from fewer students per teacher, in turn, influence outcomes like student achievement, student attitudes, and teacher morale. In this indirect fashion, then, a class size reduction opens the way for improving classroom processes and, hence, achievement. Glass and Smith (1979) attempted to integrate the extensive literature on the relation between class size and achievement, and their results are used as the basis for calculating the effect sizes in this cost-effectiveness study. Cost-effectiveness comparisons will be made for reducing class size successively from 35 to 30 students, 30 to 25, 25 to 20, and 35 to 20.

INCREASING INSTRUCTIONAL TIME

Although reducing class size has been the most prominent intervention for improving schooling in the past, increasing instructional time has more recently become a favorite recommendation of educational reformers. National reports argue for increases in the amount of time devoted to instruction by lengthening the school day and school year, assigning more homework, and using existing time more effectively (National Commission on Excellence in Education, 1983: 29; Task Force on Education for Economic Growth, 1983: 38).

The evidence behind these policies derives from comparisons of time in instruction between U.S. schools and those of other industrialized nations as well as studies of the effects of time in learning on achevement. The typical U.S. school day lasts 5-6 hours, whereas a 7-hour day is common in other industrialized countries such as Japan. Further, although a 180-day school year is the norm in the United States, 220- to 240-day sessions are found in other nations. Empirical studies suggest that more instructional time as well as greater amounts of "time-on-task" or "engaged learning" will improve educational achievement (Denham and Lieberman, 1980; Karweit, 1983).

The data used here to measure the effectiveness of increased learning time derive from the Beginning Teacher Evaluation Study (BTES), the most important data source on the subject (Denham and Lieberman, 1980; Fisher et al., 1980). The BTES research team carefully observed selected students in a number of second and fifth grade classrooms in 1976-1977 at the same time that teachers in those classrooms kept detailed logs of instructional content in mathematics and reading and time spent on those activities for an 85-day period. Student achievement was assessed by tests geared to the specific content taught.

Scheduled class time is not identical to instructional time since some time will be used for student entry and departure, teacher clerical tasks, student disruptions, field trips, and the like. In adapting the BTES findings to a time intervention for second and fifth grade reading and mathematics, we therefore assumed that only a portion of available time will be used for instruction. We estimate that to lengthen the school year of 180 days by one hour a day will add only 150 hours to instruction, instead of 180 hours.

COSTS AND EFFECTS OF INTERVENTIONS

In this section we will present the estimated effects and costs of CAI and the other three interventions. Details on the analysis of effectiveness for each intervention are reported separately in Glass (1984), so we will report only the basic method and overall results here.

Although most of the interventions used achievement tests that were "normed" for a national sample, the test instruments themselves differed from study to study. The student achievement gains were therefore converted into standard deviation units to provide a comparable measure of effectiveness. For experimental designs, this measure of effect was generally estimated as the average test score difference between treatment and control groups divided by the standard deviation of the control group (Glass et al., 1981).

In the case of quasi-experimental research designs (research in which statistical controls are used to adjust for differences among students rather than random assignment of students to treatments), the effect size was derived by dividing the increase in test scores associated with the regression coefficient for the intervention by the standard deviation of test scores in the sample. Quasi experiments employing covariates implicitly raise the choice of a metric for standardizing mean differences. One may elect to express treatment effects in either the metric of within-group standard deviation on the dependent variable or this same standard deviation corrected for covariate variability. Although a case can be made for each option (Glass et al, 1981: 114ff), we chose here to standardize mean differences by the unadjusted within-group measure of variability (on the grounds that it is more directly observed by practitioners, it is always available from a study report whereas a mixture of studies using and not using covariates would yield an inconsistent jumble of adjusted and unadjusted effects, and the sizes of effects expressed on an adjusted metric depend on arbitrary choices of number and type of covariates used). Thus the effectiveness of an intervention was viewed as the increase in test scores associated with the intervention in standard deviation units. Each standard deviation is approximately equal to gains of an academic year of 10 months, so each tenth of a standard deviation can be viewed as about 1 month of achievement gain.

Our general strategy of effectiveness analysis was to ascertain the range of results of different studies on each intervention and to explore explanations for differences in results, such as testing format, grade level, student population, or variations of the intervention. Once a range of effects was established, a specific study was chosen with effects toward the middle of the range that also met the other criteria set out earlier.

A summary of the effectiveness of each intervention in mathematics and reading achievement is presented in Table 1. All effects are based on the assumption of a full school year of intervention.

EFFECTS OF COMPUTER-ASSISTED INSTRUCTION

Effect sizes of the drill and practice approach of the Computer Curriculum Corporation are based on reanalysis of the results of the four-year experiment carried out by the Educational Testing Service in the Los Angeles Unified School District from 1976-1980 and are asso-

TABLE 1

Effect Sizes per Year of Instruction for Four Educational Interventions

	Mathematics						Reading					
Intervention	*Mean*	*Grade 2*	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Mean*	*Grade 2*	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*
Computer-Assisted Instruction												
(10-minute daily session on mini-computer)												
Overall	.12[c]	.13[a]			.12[b]		.23[c]	.23[a]			.23[b]	
	Computation	.30[a]			.25[b]		Vocabulary	.25[a]			.25[b]	
	Concepts	.00[a]			.00[b]		Comprehension	.20[a]			.20[b]	
	Application	.10[a]			.10[b]							
Cross-Age Tutoring (Boise model)												
Combined peer and adult program	.79[c]	1.02	.91	.79	.68	.55	.42[c]	.50	.46	.42	.39	.35
1 – Peer component	.97[a]	1.02	.91				.48[a]	.50	.46			
1 - Adult component	.67[b]			.79	.68	.55	.38[b]			.42	39	35
Reducing Class Size												
from 35 to 30	.06						.03					
from 30 to 25	.07						.04					
from 25 to 20	.09						.05					
from 35 to 20	.22						.11					
Increasing Instructional Time	.03	.02			.04		.07	.08			.07	
(additional 30 minutes per day for each subject)												

NOTE: a = average for grades 2 and 3; b = average for grades 4, 5, and 6; c = average for grades 2 through 6.

ıted with a 10-minute daily session in each subject. The ETS/LAUSD sign was large and complex. Three different types of control were ıployed (within-group, comparisons using randomization cohort con- ɔls, and comparison schools) and some groups were followed for four ars. The design is discussed at length in Glass (1984).

Table 1 reports results for grades 2 and 5 as well as estimated mean fect sizes. The mean effect size is .12 for mathematics and .23 for ading. The mean score in each area is based on an equal weighting of e three mathematics subtests and two reading subtests. The largest fect size in mathematics is for computation, with a smaller effect for ɔplication and virtually no effect (that is, no superiority over tradi- ɔnal classroom instruction) for concepts. The two subscores (vocabu- ry and comprehension) for the reading effect are in much closer greement.

FFECTS OF CROSS-AGE TUTORING

The cross-age tutoring approach used in Boise, Idaho, consists of ildren in the upper elementary grades tutoring students in grades 2 ıd 3 and adults tutoring students in grades 4, 5, and 6. Other adults ere responsible for training student tutors and for overall coordination f the tutoring program. Comparable achievement gains were found for oth student tutors and tutees. Table 1 breaks down effects into peer and dult components and summarizes them for the combined program.

Overall tutoring effects were substantial, with average effect sizes of ⁾7 and .48 for mathematics and reading, respectively, in the peer ɔmponent, and .67 and .38 for mathematics and reading in the adult ıtoring component. Average effect sizes in the combined peer and adult rogram were .79 for mathematics and .42 for reading. Although the ffect sizes are lower at each successive grade level, it is not possible to scertain if this is something intrinsic to tutoring or the rate of learning asic skills, if the adult tutoring approach used in the upper grades is less ffective than the peer approach used in the lower grades, of if the ifference is due to a measurement artifact.

FFECTS OF REDUCING CLASS SIZE

The effect of reducing class size is based upon a refinement of the esults of a meta-analysis of 77 studies (Glass and Smith, 1979). These authors found a substantial discrepancy between the findings of studies on class size and achievement depending on whether one based estimates of effects on studies using random assignment to classes of different sizes or on studies using preexisting nonrandom groups. Of the 77 studies, 14 randomized studies showed a stronger relationship of class size and achievement and became the basis for estimating the strength and shape of the relationship. After evaluating the 14 randomized studies and exploring unique effects of a variety of mediating factors in them, it was found that the relation between class size differences and learning effectiveness could be estimated by the following relation:

$$\hat{\Delta}_{S-L} = \hat{\beta} \log_e (L/S)$$

where $\hat{\Delta}_{S-L}$ is the estimated effect size for achievement in changing from a large class size of L pupils to a small class size of S pupils, and $\hat{\beta}$ is a constant determined by fitting the model to the data by least squares. The value of $\hat{\beta}$ is about .40 for mathematics and .20 for reading; these estimates are based on subsets of the 14 randomized studies, which evaluated effects either in reading or mathematics. The effect sizes for reducing class size in Table 1 were thus estimated for successive reductions of 5 students from a class of 35 to a class of 20. An estimate was also made for reducing class size directly from 35 to 20. The typical effect sizes associated with a class size reduction of 5 students is about .07 in mathematics and about half that in reading. For a reduction in class size from 35 to 20, the expected increase in effect size is about .22 standard deviation units for mathematics and .11 for reading.

EFFECTS OF INCREASING INSTRUCTIONAL TIME

The estimate of effectiveness for increasing instructional time was based on adding one hour to the elementary school day, divided equally between mathematics and reading. Although this would add 180 hours a year—90 for mathematics and 90 for reading—we also assume that only about 80% of the time would actually be used for instruction (Rosenshine, 1980: 110). We base estimates of engaged instructional time and effects on results from the Beginning Teacher Evaluation Study (BTES), which carried out a detailed analysis of classroom time. We estimated that about 186.5 hours and 232.6 hours were devoted during a school year to reading at grades 2 and 5, respectively, in the BTES classrooms. The corresponding hours of mathematics instruction were 102 hours at grade 2 and 133 hours at grade 5. An additional 75 hours of instruction a year in each subject—the total from adding 80% of 30 minutes per subject per day—would therefore increase the amount of time devoted to reading by about 40% at grade 2 and 32% at grade 5, and would increase learning time in mathematics by about 74% in grade 2 and 56% at grade 5. These represent substantial increases in instruction.

It is important to mention that the fifth grade mathematics result from the BTES data was suspect in that it was greatly inconsistent with the other results and seemed to be due to an anomalously large effect for a single subtest, fractions. Accordingly, the finding was adjusted to provide a result that was more consistent with the other subtests and other studies in the literature (Glass, 1984). The resulting effect sizes were relatively small, with a mean estimated effect of only .03 for mathematics and .07 for reading.

COSTS OF THE INTERVENTION

The goal of the cost portion of the analysis was to ascertain the costs of replicating each intervention so that comparisons across interventions could be made. Replication refers to the ability to undertake the same intervention with similar effects at a different site. Accordingly, the replication costs include only those required to reproduce the intervention in new settings, but not the costs associated with initial development activities that created or evaluations that assessed the intervention.

The procedure for estimating the cost of an intervention is based on a three-stage approach (Levin, 1983). First, the ingredients for replicating a program are specified in detail. Second, an annual cost is placed on each ingredient. The summation of these costs provides an estimated total annual cost for each intervention. Finally, a cost per student is derived by dividing the total annual cost figure by the number of students served.

It is important to emphasize that all of the four interventions represent instructional supplements rather than replacements of basic instructional services. Accordingly, the costing strategy addresses only the additional resources or ingredients required to replicate these supplemental interventions, that is, their marginal costs. For each intervention we identified the ingredients by consulting documents and, where necessary, expert practitioners, to obtain appropriately detailed descriptions of the interventions. The comprehensiveness and detail permitted cost estimates of the necessary resources for the intervention. These were classified according to numbers and types of personnel, facilities and equipment, materials, and other required ingredients.

Assigning a cost to the ingredients themselves entailed a number of steps. First, to obtain a consistent set of costs for a specific year, an attempt was made to set out average "national" costs for 1980. For example, whenever a full-time classroom teacher is used in an intervention, the cost is established as $21,875 per year on the basis of an average salary for 1980 of $17,500 and fringe benefits of $4,375. Similar calculations are made for other personnel, facilities, and all equipment with the exception of computer hardware. By using cost data for the same year, it was then possible to obtain a uniform basis for comparisons. Even though costs have risen since 1980, this is unlikely to affect the *relative* cost patterns with the exception of costs for computer hardware. Finally, to obtain a cost per student, the total cost of each intervention was divided by the number of students.

In the case of computer hardware, the rapid decline in costs since 1980 suggested that we obtain the most recent cost information. Thus the costs of computer hardware are based on prices to schools in the spring of 1984.

Costs of facilities and equipment in all interventions were annualized, that is, converted into a cost per year (Levin, 1983: 67-71) by a procedure that takes into account their replacement cost, life span, and interest rates. The cost of each ingredient and the overall or total cost assigned to each intervention thus represent costs for one year of operation.

For purposes of comparability, we ascertained the full cost of each intervention. Thus the complete costs of personnel and facilities are accounted for, even if some of the personnel were volunteers and facilities were provided "free" or without charge by other units of government. Since the ingredients, costs, and cost sources for each intervention are available in Levin et al. (1984), analysts at any particular site can adjust and update our estimates to make their own cost estimates for their own particular sites. Such adjustments might include substituting local for national figures and current for 1980 prices. To the degree that

TABLE 2
Annual Cost per Student per Subject of
Four Educational Interventions

	Cost per Student per Subject ($)
CAI	
Cross-age tutoring	119
Peer component	212
Adult component	827
Increasing instructional time	61
Reducing class size	
from 35 to 30	45
from 30 to 25	63
from 25 to 20	94
from 35 to 20	201

any potential decision maker can reduce costs through obtaining volunteers or donated facilities, equipment, and supplies, our reported costs can be adjusted accordingly.

In a few cases we identified ingredients for particular interventions for which we did not attribute costs. Generally, these were cases in which the ingredients were truly "costless" in the sense that they were slack resources that had no alternative use other than the intervention at the time that they were employed. For example, the cross-age tutoring model is able to draw upon nooks and crannies in halls, cafeterias, gymnasiums, auditoriums, resource centers, lounges, and vacant classrooms at times when these spaces would not be used for their intended functions. Details of the costing process for each of the interventions follow, and Table 2 presents the annual cost per student.

COST OF COMPUTER-ASSISTED INSTRUCTION

Personnel costs for replicating the CAI intervention include a coordinator, two teaching aides, and a small portion of the time of the principal. The CAI coordinator is responsible for the overall functioning of CAI including scheduling and coordination of instruction, reporting to teachers on student progress, and monitoring of equipment functioning and maintenance. This role is served by a classroom teacher who is trained in an intensive 1½-day program. Teaching aides monitor the performance of students and assist them in understanding and solving the CAI difficulties encountered.

Facilities include a classroom for the CAI laboratory and renovation to instill counters, air conditioning, and security devices. Equipment and materials include the minicomputer, 32 terminals, a printer, curriculum rental, chairs and other furnishings, and supplies. All of the hardware and software costs are based on prices quoted by the provider, Computer Curriculum Corporation, in March 1984. Estimates of use are based on evidence from the Los Angeles evaluation, which suggested that each of the 32 terminals could accommodate about 23 daily sessions resulting in a total of 736 sessions per day.

The total cost per school for a fully equipped computer laboratory, personnel, and other requirements is about $87,000 a year, resulting in an annual cost per student per 10-minute daily session of about $119 (based on 736 sessions per day).[1] In 1978 the cost of a similar system was estimated at $136 per student (Levin and Woo, 1981), so a combination of 1984 hardware and software costs and 1980 costs for other ingredients reduced the overall costs per student by only 12%, despite a large drop in the cost of hardware.

Some analysts assume that declines in hardware costs will substantially reduce the costs of CAI. However, hardware costs represent only about 11% of the cost of the CAI intervention, in comparison with about one-quarter of the costs of the same intervention in 1978. Almost 90% of the present cost for delivering the CAI services is not associated with the hardware, so even drastic future declines in hardware costs would not greatly reduce the overall cost per student. For example, even if the cost of the hardware were to decline by 50%, the cost per student would decline by about 5%-6%—assuming that all other costs remained the same. Since other costs are rising over time, it is conceivable that the overall cost reduction in this scenario would be at least partially offset by higher costs for personnel and other ingredients. The CAI intervention requires considerably more than hardware to provide CAI services.

COSTS OF CROSS-AGE TUTORING

From the various evaluation reports for the tutoring program as well as detailed inquiries and interchanges with the Boise School District, we identified the various ingredients for the entire cross-age tutoring program with its separate peer and adult components. A typical school with 60 tutors and 60 tutees in the student or peer tutoring component and 30 tutees in the adult tutoring component was used as the unit of analysis.

The total costs of the complete tutoring program (peer and adult components combined) were estimated at $41,433 for the 150 students or a cost per student of about $276. Since the peer tutoring approach for grades 2 and 3 and the adult tutoring approach for grades 4, 5, and 6 were separable, estimates were made individually. The peer approach showed a cost of $212 per student participant (which included tutors and tutees), and the adult tutoring approach showed a cost of about $827 per student.

The substantial difference in costs was primarily due to two factors. First, the peer tutoring component produces achievement gains for both tutors and tutees because both are counted as student participants, whereas the adult tutoring component produces achievement gains only for the tutees. Costs of adult tutoring are divided by the smaller number of students affected so the costs are distributed over twice as many students for the peer component. Second, the peer tutoring model assumes no cost for the time of elementary students in terms of market opportunities or lost learning. Tutoring activities do not compete with other mathematics and reading opportunities. In contrast, the time of adult tutors is costly, and each adult can tutor only a limited number of students. Thus the personnel cost for the adult model is higher and is distributed over fewer student participants, resulting in a much higher cost per student for the adult component by itself.

COSTS OF REDUCING CLASS SIZE

A reduction in class size requires the availability of more teachers with additional classrooms and furnishings. A classroom for our purposes includes the physical space, furnishings, energy needs, insurance, maintenance, and a teacher. One classroom in this model costs $28,138 annually, or a cost per student of $804 when class size is 35.

Decreasing class size from 35 to 30 pupils would require an increase of $135, or about 14% in cost per student for that classroom. Similarly, reducing class size from 30 to 25 pupils raises costs by an additional $188 per student, or about 17%. A decrease from 25 to 20 students entails an increase in costs of $281 per student, or 30%. Finally, a single reduction in class size from 35 to 20 implies an increase in per pupil costs of $603, or about 43%.

Reduction in class size is an overall educational intervention that should affect all of the educational activities during the school day, not just the teaching of mathematics and reading. Consequently, only a portion of the additional cost should be viewed as an educational intervention to improve mathematics and reading. We therefore assumed that about one-third of the school day is devoted directly or indirectly to mathematics at the elementary level and one-third to reading, with the remaining one-third devoted to other areas. Although our time-in-learning analysis indicated that formal instruction in mathematics and reading takes up less than two-thirds of the school day, we assumed that the benefits of smaller classes for mathematics and reading should also be conferred from other activities such as social studies, writing, and science. Accordingly, the total additional cost per student for a given reduction in class size was divided by three to obtain an estimated cost per subject comparable to those calculated for the other interventions.

COSTS OF INCREASING INSTRUCTIONAL TIME

Estimation of the cost of increasing the length of the school day is straightforward. We assumed that the only additional cost arises from higher salaries and fringe benefits associated with additional teacher time. This additional cost was calculated by increasing teacher salaries and fringe benefits by one-sixth to accommodate an additional hour of instruction beyond a normal six-hour requirement. Given an average class size of 30, the annual cost of this intervention, then, is estimated at $61 per student per subject. We assumed that such an intervention would not entail additional costs for administration, library, maintenance, or curriculum materials and supplies. We further assumed that no additional facilities will be required (and that no activities will be displaced).

COST EFFECTIVENESS RESULTS

From the data on effectiveness and costs, it is possible to calculate cost-effectiveness ratios and to rank the alternative interventions. Table 3 provides estimates of the cost per student per subject for each of the four interventions as well as effect sizes for each $100 of cost per student. The effect size for each $100 of cost per pupil is our cost-effectiveness ratio. Consider the results for reading and mathematics separately.

TABLE 3
Cost per Student per Subject and Cost-Effectiveness Ratios of Four Interventions (Effect Size for Each $100 Cost per Student)

		Cost-Effectiveness Ratios											
		Mathematics						Reading					
Intervention	*Cost per Student (Per Subject) ($)*	*Mean*	*Grade 2*	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Mean*	*Grade 2*	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*
Computer-Assisted Instruction													
(10-minute daily session on minicomputer)													
Overall	119	.10[c]	.11[a]			.10[b]		.19[c]	.19[a]			.19[b]	
		Computation	.25[a]			.21[b]		Vocabulary	.21[a]			.21[b]	
		Concepts	.00[a]			.00[b]		Comprehension	.17[a]			.17[b]	
		Application	.08[a]			.08[b]							
Cross-Age Tutoring (Boise Model)													
Combined peer and adult program	276	.29[c]	.37	.33	.28	.25	.20	.15[c]	.18	.16	.15	.14	.13
Peer component	212	.46[a]	.48	.43				.22[a]	.23	.21			
Adult component	827	.08[b]			.09	.08	.07	.05[b]			.05	.05	.04
Reducing Class Size													
from 35 to 30	45[d]	.14						.07					
from 30 to 25	63[d]	.12						.06					
from 25 to 20	94[d]	.10						.05					
from 35 to 20	201[d]	.11						.06					
Increasing Instructional Time	61	.05	.04			.06		.12	.12			.11	
(additional 30 minutes per day for each subject)													

NOTE: a = average for grades 2 and 3; b = average for grades 4, 5, and 6; c = average for grades 2 through 6; d = cost per student per subject is one-third of reducing class size across all subjects.

COST-EFFECTIVENESS OF INTERVENTIONS FOR RAISING MATHEMATICS ACHIEVEMENT

Among the alternatives for increasing mathematics achievement, CAI and reducing class size show about equal cost-effectiveness ratios, although the initial reduction (from 35 to 30 pupils) shows somewhat higher cost-effectiveness than successive reductions. The two tutoring interventions—the combined cross-age approach and the peer component—show the largest effects per $100 of cost per pupil, however, with .29 for the combined program and .46 for the peer component. This means that the combined Boise tutoring program provides almost one-third of a standard deviation in test score gain per $100 cost per pupil. Whereas the peer component alone provides almost half a standard deviation gain per $100, the adult component with its higher costs and smaller effects provides a much smaller effect relative to cost. CAI and reducing class size, in contrast, show effect sizes relative to cost only about one-fourth of that for peer tutoring and less than half of that for the combined tutoring approach. Finally, increasing instructional time by a half-hour a day in mathematics has the smallest effect per unit of cost: about half that of CAI and reduced class size, one-sixth that of the combined tutoring approach, and only one-ninth that of the peer tutoring component.

Thus, the preferred alternative among the four interventions for increasing mathematics achievement cost-effectively is the peer tutoring model, followed by the combined tutoring model, CAI, reducing class size, and increasing instructional time. These rankings change somewhat when results for reading are considered.

COST-EFFECTIVENESS OF INTERVENTIONS FOR RAISING READING ACHIEVEMENT

For reading achievement, peer tutoring and CAI show almost equal cost-effectiveness ratios. The peer tutoring model at .22 appears to be slightly more cost-effective than CAI at .19, although the combined tutoring program at .15 is estimated to be slightly less effective. The relatively more expensive adult tutoring model is one of the least cost-effective of the alternatives in reading, along with reducing class size. Increasing instructional time for reading is about twice as cost-effective as reducing class size, a reversal of the results for mathematics, though both ratios are small.

In summary, the results for reading suggest that the most cost-effective approach is peer tutoring, followed closely by CAI. Increasing instructional time and the reduction of class size are less cost-effective alternatives for raising reading scores.

COST-EFFECTIVENESS FOR BOTH SUBJECTS

Because the cost-effectiveness rankings for the four interventions differ by subject, the decision maker may be confronted with a dilemma. In some cases, the solution would be to use different alternatives for different subjects. An example might be to use peer tutoring for mathematics and CAI for reading. However, in other cases such as the reduction of class size, it may be more difficult to separate interventions by subject because adoption of the intervention affects both subjects. Thus, a decision maker might consider the implications of each intervention for both subjects. It is useful for this reason to average the cost-effectiveness ratios for the two subjects to determine if an unambiguous ranking emerges when the two subjects are weighted equally in the calculations.

Table 4 shows the cost-effectiveness ratios for each intervention averaged across mathematics and reading. The peer tutoring component and the combined tutoring approach show the best result, followed by CAI. Reducing class size, increasing instructional time, and the adult tutoring component show poorer cost-effectiveness ratios.

The differences in cost-effectiveness are substantial. For example, the same cost outlay would produce almost four times as large an effect on reading and mathematics achievement through peer tutoring as through reducing class size or increasing instructional time. Although the adult tutoring approach in itself has the poorest cost-effectiveness result among all of the interventions, the high cost-effectiveness of peer tutoring contributes to a combined cost-effectiveness of the peer and adult approach that still exceeds considerably the second best alternative, CAI.

POLICY ASPECTS

The purpose of this report was to compare the cost-effectiveness of a major application of computer-assisted instruction with three other educational strategies for raising mathematics and reading achievement at the elementary school level. The findings run counter to some con-

TABLE 4
Average Cost-Effectiveness Ratios of
Four Interventions for Two Subjects
(Average of Mathematics and Reading Effect Sizes for
Each $100 Cost Per Student Per Subject)

Intervention		*Cost-Effectiveness Ratio*
Cross-age tutoring	Combined peer and adult program	.22
	Peer component	.34
	Adult component	.07
Computer-assisted instruction		.15
Reducing class size	from 35 to 30	.11
	from 30 to 25	.09
	from 25 to 20	.08
	from 35 to 20	.09
Increasing instructional time		.09

ventional expectations. Although the CAI alternative does relatively well according to the cost-effectiveness criterion, it does not do as well as peer tutoring. It is somewhat surprising that a traditional and labor intensive, approach, peer tutoring, appears to be far more cost-effective than an electronic intervention, a widely used CAI approach. Moreover the low ranking of increased instructional time, the centerpiece of many of the calls for educational reform, makes it a relatively poor choice for both reading and mathematics from a cost-effectiveness perspective.

Equally interesting and important is the contrast between the analysis of effects alone and the cost-effectiveness results. Table 1 shows that the adult tutoring model is associated with one of the largest effect sizes, .67 for mathematics and .38 for reading. Yet the costs of the adult tutoring approach are so large that it yields one of the lowest cost-effectiveness ratios in mathematics, and the lowest one in reading and the poorest average cost-effectiveness across both subjects. Accordingly, an evaluation of effectiveness alone might provide highly misleading information for the policymaker concerned with how to allocate additional resources for improving mathematics and reading achievement in the most efficient way. To extend the strong cost-effectiveness advantages of peer tutoring to the upper-elementary grades, it might be desirable to consider the use of seventh and eighth grade students from local middle schools instead of adult tutors. New cost estimates would need to be made.

In using the results of these computations, a number of cautions should be noted. First, each of the results is drawn from a particular version and application of a general class of intervention, so the results should not be used to draw a general conclusion for all possible versions of the intervention. Although we attempted to select specific forms of interventions that were tested, replicable, based upon substantial experience, and that had effects that were representative of that class of intervention, there may be other examples that are potentially more cost-effective.

Moreover, future declines in the cost of CAI and increases in its effectiveness may be reasonable possibilities. It should be noted, however, that the large proportion of non-hardware costs in CAI suggests that a decrease in hardware costs by itself may not substantially reduce the total cost of CAI services. Second, our results pertain to mathematics and reading achievement, so they should not be applied to other outcomes. Third, both costs and effects of interventions may vary from one school to the next, depending on variations in conditions that were not studied here. For example, at some schools and for some interventions, it may be possible to obtain volunteers and donations of facilities and equipment. In those cases, the cost to the sponsor may be reduced and local cost-effectiveness ratios altered in favor of those interventions. In other cases, a long tradition of working with a particular intervention may render it especially cost-effective.

The most appropriate use of these results is to provide guidelines for consideration of alternative interventions for increasing mathematics and reading achievement in elementary schools. On the basis of these findings, educators should question unqualified assertions that CAI is a more cost-effective intervention than other alternatives.

NOTE

1. A separate analysis comparing the costs of hardware for a microcomputer network and a minicomputer with roughly similar capabilities for delivering CAI showed that the minicomputer was slightly less costly. After examining the reliability and personnel needs of the two systems, it was concluded that the minicomputer was likely to have a better performance relative to cost than existing microcomputer networks. See Levin et al. (1984).

REFERENCES

Center for Social Organization of Schools, Johns Hopkins University (1983) School Uses of Microcomputers (June).

DENHAM, C. and A. LIEBERMAN (1980) Time to Learn. Washington, DC: National Institute of Education.

EHLY, S. W. and S. C. LARSEN (1980) Peer Tutoring for Individualized Instruction. Boston: Allyn & Bacon.

FISHER, C. W. et al. (1980) "Teaching behaviors, academic learning time, and student achievement: an overview," pp. 7-32 in C. Denham and A. Lieberman (eds.) Time to Learn. Washington, DC: National Institute of Education.

GLASS, G. V (1984) The Effectiveness of Four Educational Interventions. IFG Project Paper 84-A19, Institute for Research on Educational Finance and Governance, Stanford University, Palo Alto, CA.

GLASS, G. V, B. McGAW, and M. L. SMITH (1981) Meta-Analysis in Social Research. Beverly Hills, CA: Sage.

GLASS, G. V and M. L. SMITH (1979) "Meta-analysis of research on class size and achievement." Educ. Evaluation and Policy Analysis 1: 2-26.

Independent School District of Boise City (1983a) 1982-83 Evaluation Report, PROJECT INSTRUCT, Chapter 1 Program. Boise, ID: Author.

Independent School District of Boise City (1983b) "Submission to joint dissemination-review panel (JDRP), cross-age structured tutoring program in math." (mimeo)

Independent School District of Boise City (1983c) "Submission to joint dissemination-review panel (JDRP), cross-age structured tutoring program in reading." (mimeo)

KARWEIT, N. (1983) Time-On-Task: A Research Review. Report 332, Center for Social Organization of Schools, Johns Hopkins University, Baltimore, MD.

LEVIN, H. M. (1983) Cost-Effectiveness: A Primer. Beverly Hills, CA: Sage.

LEVIN, H. M., G. V. GLASS, and G. R. MEISTER (1984) Cost-Effectiveness of Four Educational Interventions. IFG Project Report 84-A11, Institute for Research on Educational Finance and Governance, Stanford University, Palo Alto, CA.

LEVIN, H. M. and L. WOO (1981) "An evaluation of the costs of computer-assisted instruction." Economics of Education Rev. 1, 1: 1-26.

National Commission on Excellence in Education (1983) A Nation at Risk: The Imperative for Educational Reform. Washington, DC: U.S. Department of Education.

RAGOSTA, M. et al. (1982) Computer-Assisted Instruction and Compensatory Education: The ETS/LAUSD Study Final Report, Project Report 19. Princeton, NJ: Educational Testing Service.

ROSENSHINE, B. (1980) "How time is spent in elementary classrooms," chap. 5 in C. Denham and A. Lieberman (eds.) Time to Learn. Washington, DC: National Institute of Education.

Task Force on Education for Economic Growth (1983) Action for Excellence. Denver, CO: Education Commission of the States.

Henry M. Levin is Professor of Education and Affiliated Professor of Economics at Stanford University. He is a specialist in the economics of human resources and the

economics of evaluation. He was the President of the Evaluation Research Society in 1982 and was recipient of its Myrdal Award for Human Service Delivery in 1985.

Gene V. Glass is Professor of Education at Arizona State University with appointments in the Divisions of Psychology and Leadership and Policy Studies. He is a specialist in statistical methods and measurement and their relation to policy. He is past-president of the American Educational Research Association and was a recipient of the Evaluation Research Society's Lazarsfeld Award in 1984.

Gail R. Meister is Senor Research and Evaluation Associate at Research for Better Schools, Inc., in Philadelphia. She is also a Ph.D. candidate in Administration and Policy Analysis at Stanford University. Her research interests focus on knowledge use and implementation.