

Effectiveness of Computer-Based Education in Elementary Schools

**James A. Kulik, Chen-Lin C. Kulik, and
Robert L. Bangert-Drowns**

UNIVERSITY OF MICHIGAN

Abstract—A meta-analysis of 32 comparative studies showed that computer-based education has generally had positive effects on the achievement of elementary school pupils. These effects have been different, however, for programs of off-line computer-managed instruction (CMI) and for interactive computer-assisted instruction (CAI). The average effect in 28 studies of CAI programs was an increase in pupil achievement scores of 0.47 standard deviations, or from the 50th to the 68th percentile. The average effect in four studies of CMI programs, however, was an increase in scores of only 0.07 standard deviations. Study features were not significantly related to study outcomes.

Computers are fast becoming an important factor in elementary school teaching. The number of computers in American elementary schools has increased by a factor of at least 10 during this decade, and the majority of schools now own them (Becker, 1983). The use of computers in teaching is nonetheless a difficult subject to bring into focus. Researchers and developers disagree on some of the basic issues.

Even the terminology in the area is open to dispute. The acronym CAI is often used, but it is variously interpreted as standing for *computer-assisted instruction*, *computer-aided instruction*, *computer-augmented instruction*, or *computer-administered instruction*. Other terms used in the area are *computer-managed instruction*, *computer-based learning*, and *computer-based instruction*. *Computer-based education*, or CBE, is becoming increasingly popular as a generic term for the area because it encompasses a broad spectrum of computer applications (Hall, 1982).

Researchers also differ in their opinions on the best way to subdivide the area. Early taxonomies of CBE usually distinguished between four uses of the computer in teaching (Atkinson, 1969; Watson, 1972): (a) In *drill-and-practice* applications, the teacher presents lessons to pupils by conventional means, and the

The material in this report is based upon work supported by a grant from the Exxon Education Foundation. Any opinions, findings, and conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the Exxon Education Foundation. The authors thank David Shwalb and Jolie Spring for their invaluable assistance in locating studies for the report.

Requests for reprints should be sent to James A. Kulik, Center for Research on Learning and Teaching, University of Michigan, 109 E. Madison St., Ann Arbor, MI 48109.

computer provides practice exercises as a follow-up; (b) in the *tutorial* mode, the computer both presents the concepts and provides practice exercises on them; (c) in *dialogue* mode, the computer presents lessons and practice, and the student is free to construct natural language responses, ask questions in unrestricted mode, and almost completely control the sequence of learning events; and (d) in *computer-managed instruction*, the computer evaluates students either on-line or off-line, guides students to appropriate instructional resources, and keeps records.

Recent taxonomies show that today's researchers have a broader conception of the role that computers can play in education. Taylor (1980), for example, has described three uses of the computer in schools. First, as a *tutor*, the computer presents material, evaluates student responses, determines what to present next, and keeps records of student progress. Most computer uses described in earlier taxonomies involved the tutoring function of computers. Second, the computer serves as a *tool* when students use it for statistical analysis, calculation, or word processing, such as when students use it as a calculator in mathematics classes, as a map-maker in geography, as a performer in music, or as a text editor and copyist in English. Third, the computer serves as a *tutee* when students give it directions in a programming language that it understands, such as Basic or Logo.

Experts also differ in their assessment of the effects that CBE is likely to have on children. In *Run, Computer, Run*, for example, Oettinger (1969) has argued that claims of positive benefits from CBE are grossly exaggerated. In Oettinger's view, schools are too conservative to make good use of teaching innovations, and computer-based teaching technology is too poorly developed to be of much use to schools. In *Mindstorms*, on the other hand, Papert (1980) has written glowingly of the effects that computation has on children. In Papert's view, computers can open up *microworlds* for children. In the simulated reality of these worlds, children learn to manipulate objects, to apply themselves to long-range projects, and to construct new symbol systems. They thereby gain a greater mastery of the world, a sense of the power of applied knowledge, and a realistic image of themselves as intellectual agents.

Most educators and researchers can live with the differences in opinion about computer terminology and taxonomy. Although troublesome, such differences do not seriously affect their capacity to use computers. Differences in opinion about the effects of CBE on children are an entirely different matter. Such differences can affect planning for the future. Without knowing what CBE has already achieved, teachers and researchers cannot be confident in designing new instructional programs.

To clear up uncertainty, researchers have therefore carried out numerous studies evaluating computer uses in schools. Some of their evaluation reports are impressionistic and of little scientific value. Other evaluation reports, however, describe controlled studies that merit serious attention. In a typical study of this type, a researcher divides a group of students into an experimental and a control group. Members of the experimental group receive part of their instruction with computer assistance, whereas members of the control group receive their instruction by conventional teaching methods. At the end of the experiment, the researcher compares responses of the two groups on a common measure. Such evaluation studies have been carried out often enough to give some indication of the overall worth of CBE.

Reviews designed to integrate the findings from the evaluation studies are of two basic types: box-score tabulations and meta-analyses. Box-score reviews usually report the proportion of studies favorable and unfavorable to CBE, and often provide narrative comments about the studies as well. Reviewers using meta-analysis take a more quantitative approach to their task (Glass, McGaw, & Smith, 1981). They use (a) objective procedures to locate studies, (b) quantitative or quasi-quantitative techniques to describe study features and outcomes, and (c) statistical methods to summarize overall findings and explore relationships between study features and outcomes.

The box-score reviews have produced a generally favorable picture of CBE effects. Vinsonhaler and Bass' (1972) review, for example, examined results from eight separate reports on computer-based drill-and-practice in elementary schools. The eight reports presented results from ten studies, three in language arts and seven in arithmetic. For language arts, the performance of CAI groups was between 0.1 and 0.4 school years beyond the performance of traditionally instructed groups, and all findings were positive. For arithmetic, most comparisons showed a statistically significant advantage in favor of the CAI group.

Edwards, Norton, Taylor, Weiss, and Dusseldorp's (1975) box-score review covered a wider range of applications of CBE, including studies of computer-based drill-and-practice, tutorials, and simulations carried out in elementary schools, high schools, and colleges. Ten of the 27 reports that these investigators located described findings from elementary schools. These reports covered 12 separate sets of results. The reviewers considered 9 of these 12 sets of findings (or 75%) to be positive for CBE.

Meta-analytic research by Hartley (1978) and Burns (1981) focussed on computer-based drill-and-practice and tutorials in elementary and secondary arithmetic. Hartley's analysis covered 22 studies and 89 separate results. A total of 81 of these 89 results were from grades 1 through 8. The average effect of CBE in these grades was to raise arithmetic achievement scores by 0.42 standard deviations. Burns (1981) located 32 separate documents for her meta-analysis. She coded approximately 400 results from these studies, and nearly 90% of these were from the elementary grades. The average effect of CBE in these grades was to raise arithmetic achievement by 0.37 standard deviations.

The value of these early reviews is limited by several factors. First, the conclusions in the reviews are based mainly on studies published before 1975 of computer-based teaching of arithmetic. The reviews give too little attention to recent studies and to computer-based teaching in subjects other than arithmetic. None of the studies reviewed by Vinsonhaler and Bass (1972), for example, was published after 1971; none of those reviewed by Edwards et al. (1975) was published after 1973, and only a handful of the studies of elementary school CBE reviewed by Hartley (1978) and by Burns (1981) were conducted during the past 10 years. All of the studies reviewed by Hartley and Burns, all but three of the studies reviewed by Vinsonhaler and Bass, and five of those reviewed by Edwards et al. focussed on the teaching of arithmetic.

A second factor that limits the value of earlier reviews is their methodology. Early reviews of CBE effectiveness were written at a time when rapid progress was being made in the development of a methodology for research synthesis, and most of the reviewers were not able to incorporate the advances into their analyses. The reviews by Vinsonhaler and Bass (1972) and by Edwards et al. (1975),

for example, relied on box-score tabulations of major findings. The statistical weakness of box-score reviews has recently come in for heavy criticism from research synthesists (Hedges & Olkin, 1980). Although the meta-analytic approach introduced by Glass (1976) has won greater acceptance in recent years, meta-analytic methodology itself has matured with the years. Early users of the method were often unselective in choosing studies for analysis; they often inflated numbers by using nonindependent findings in a single statistical analysis; and they often reported their results in a sketchy fashion. Users of the methodology today are more sensitive to such methodological pitfalls and usually try to avoid them.

This review is meant to supplement earlier reviews on the effectiveness of CBE in elementary schools. It updates them and uses state-of-the-art methods in integrating the research findings. The article asks: How effective is computer-based teaching at the elementary school level? Is it especially effective for certain types of outcomes or certain types of students? Under which conditions does it appear to be most effective?

METHOD

The meta-analytic approach used in this review was similar to that described by Glass et al. (1981). Their approach requires a reviewer (a) to locate studies of an issue through objective and replicable searches; (b) to code the studies for salient features; (c) to describe study outcomes on a common scale; and (d) to use statistical methods to relate study features to outcomes.

Data Sources

The studies considered for use in this meta-analysis came from three major sources. One large group of studies came from the references in major reviews of CBE effectiveness (Burns, 1981; Edwards et al., 1975; Hartley, 1978; Vinsonhaler & Bass, 1972). A second group of studies was located by computer-searching two library data bases using Lockheed's DIALOG on-line information services. The data bases searched in this way were (a) *Comprehensive Dissertation Abstracts*, and (b) *ERIC*, a data base on educational materials from the Educational Resources Information Center, consisting of two files, *Research in Education* and *Current Index to Journals in Education*. A third group of studies was retrieved by branching from bibliographies in the documents located through reviews and computer searches.

These search procedures yielded 32 studies that met four basic criteria for inclusion in our data set. First, the studies had to take place in actual classrooms in grades 1 through 6. Second, the studies had to provide quantitative results from both a computer-taught and a conventionally instructed class. Uncontrolled "experiments" and anecdotal reports were not acceptable. Third, the studies had to be free from such crippling methodological flaws as substantial differences in aptitude of treatment and control groups, unfair "teaching" of the criterion test to one of the comparison groups, and differential rates of subject attrition from the groups being compared. And fourth, the studies had to be

retrievable from university or college libraries by interlibrary loan or from the Educational Resources Information Center (ERIC), the National Technical Information Service (NTIS), or University Microfilms International.

Standards such as these kept us from using all the reports cited by previous reviewers. A total of 14 studies cited by four previous reviewers (Burns, 1981; Edwards et al., 1975; Hartley, 1978; and Vinsonhaler & Bass, 1972) were found suitable for inclusion in our analysis. Of the 14 studies, only one was cited by all four reviewers; 8 of the studies were cited by two reviewers; and 5 studies were cited by only one reviewer. The major reasons for eliminating reports cited in other reviews were (a) the studies did not examine effects in grades 1 through 6; (b) the reports could not be obtained through interlibrary loans or from standard clearinghouses; and (c) the studies used pre-experimental designs that did not call for conventionally instructed control groups.

Study Features

A total of 21 variables was used to describe treatments, methodologies, settings, and publication histories of the studies (Table 1). These 21 variables were chosen on the basis of an examination of variables used to describe study features in previous reviews, and a preliminary examination of dimensions of variation in the studies located for this analysis. Two coders independently coded each of the studies on each of the variables. The coders then jointly reviewed their coding forms and discussed any disagreements. They resolved these disagreements by jointly re-examining the studies whose coding was in dispute.

Outcome Measures

The instructional outcome measured most often in the 32 studies was student learning, as indicated on achievement examinations given at the end of the program of instruction. Other outcomes measured in the studies included (a) performance on a follow-up or retention examination given some time after the completion of the program of instruction; (b) change in student attitude toward school subjects; and (c) change in attitude toward computers.

For statistical analysis, outcomes from a variety of different studies with a variety of different instruments had to be expressed on a common scale of measurement. The transformation used for this purpose was the one recommended by Glass et al. (1981). To reduce measurements to a common scale, Glass and his colleagues code each outcome as an *Effect Size (ES)*, defined as the difference between the mean scores of two groups divided by the standard deviation of the control group. For studies that report means and standard deviations for both experimental and control groups, *ES* can be calculated directly from the measurements provided. For less fully reported studies, *ES* can be calculated from statistics such as *t* and *F*.

The application of the formula given by Glass and his colleagues is straightforward in most cases. In some studies, however, more than one value is available for use in the numerator of the formula for *ES* and more than one value is available for the denominator. For example, some investigators report raw-score differences between groups as well as covariance-adjusted differences, and some report differences on a postmeasure as well as differences in pre-post gains. In

Table 1. Categories Used to Describe Study Features*Use of Computer*

Drill-and-practice—The computer provided practice exercises but not the original lesson on a topic.

Tutorial—The computer presented the lesson as well as practice exercises on the material.

Management—The computer evaluated student performance, guided students to appropriate instructional resources, and kept records of student progress.

Author of Program

Local—Computer materials were developed locally for a specific setting.

Other—Computer materials were developed for use in a wide variety of settings.

*CCC Materials**

Yes—Materials used in the study were developed at Stanford University and were obtained from the Computer Curriculum Corporation of Palo Alto.

No—Other materials.

Type of Computer Interaction

Off-line

Terminal with mainframe

Microcomputer

*Number of CBE Sessions per Week**Total Number of Weeks of CBE**Number of Minutes per CBE Session**Total Amount of Time on CBE (in minutes)**Subject Assignment*

Random—Subjects were randomly assigned to the experimental and control groups.

Nonrandom—A quasi-experimental design was used.

Control for Instructor Effects

Same instructor—The same teacher or teachers taught both the experimental and control groups.

Different instructors—Different teachers taught the two groups.

Control for Historical Effect

Same semester—Subjects in experimental and control groups were taught concurrently.

Different semester—Two groups were not taught concurrently.

Control for Time-on-task

Experimental > Control—Experimental subjects received regular instruction *plus* supplemental computer assistance.

Experimental = Control—Total amount of instructional time was equal for experimental and control groups.

Control for Test-author Bias

Commercial—A standardized test was used as the criterion measure for student achievement.

Local—A locally developed test was used as the criterion measure.

Control for Bias in Test Scoring

Objective—Objective, machine-readable examinations were used to measure student achievement, e.g., multiple-choice tests.

Nonobjective—Subjective decisions had to be made in scoring tests, e.g., essay tests.

Field-tested Computer Materials

Yes

No

Control for Evaluator Involvement

Involved—The evaluator was involved in developing the CBE material and/or in conducting the CBE program.

Not involved

Class Level

Primary—Subjects included in the study came from grades 1 through 3.

Primary and middle—Subjects came from both primary and middle grades.

Middle—Subjects came from grades 4 through 6.

Table 1 (continued)

Course Content

Mathematics

Science

Language Arts and/or Reading

Combined—More than one of the above were studied.

Subject Ability Level

Low

Average or mixed

High

Source of Study

Unpublished—ERIC document, paper presented at a convention, etc.

Dissertation

Published—Journal article, book, etc.

Year of the Report

*CCC was the only supplier whose CAI materials were used in more than one or two studies, and so this source was singled out for special coding.

such cases, we used as the numerator of *ES* the difference that gave the most accurate estimate of the true treatment effect. That meant using covariance-adjusted differences rather than raw-score differences, and differences in gains rather than differences on posttests. In addition, some reports contained several measures of variation that might be considered for use as the denominator of *ES*. We used the measure that provided the best estimate of the unrestricted population variation in the criterion variable.

Statistical Analysis

Some studies reported more than one finding for a given outcome area. Some of these multiple findings resulted from the use of more than one experimental or control group in a single study. Others resulted from the use of several subscales and subgroups in measuring a single outcome. Using several different *ES*s to represent results from one outcome area in one study seems inappropriate to us. The multiple *ES*s are not independent; they often come from a single group of subjects or overlapping subject groups, and in any case they represent the effects of a single program implemented in a single setting. To represent a single outcome by several *ES*s violates the assumption of independence necessary for many statistical tests and also gives undue weight to studies with multiple groups and multiple scales.

The procedure we adopted, therefore, was to calculate only one *ES* for each outcome area of each study. A single rule helped us to decide which *ES* best represented the study's findings. The rule was to use the *ES* from what would ordinarily be considered the most methodologically sound comparison when comparisons differed in methodological adequacy. Specifically, when a study included both a conventionally taught control group and a no-treatment control group, results from the comparison with the conventionally taught group were coded for analysis. This procedure controlled for the possibly confounding effects of differential time-on-task. When results from both a true experimental comparison and a quasi-experiment were available in the same study, results of the true

Table 2. Major Features and Achievement Effect Sizes in 32 Studies of Computer-Based Education

Study	Place & Time	Grade	Course Content	Use	Weeks of Instruction	Achievement ES
Akkerhuis (1974)	Germany, 1972-73	6	M	Management	18	-0.08
Atkinson (1969)	California, 1966-67	1	L	Tutorial	25	0.44 ^a
Chiang, Stauffer & Cannara (1978)	California, 1977-78		C	D & P	35	0.19
Coffman & Olsen (1980)	Louisiana, 1976-78	3,4	C	Management	72	-0.01
Cranford (1976)	Mississippi, 1975-76	5,6	M	D & P	12	0.64
Davies (1972)	California, 1971	3-6	M	D & P	16	0.34
Delon (1970)	Mississippi, 1968-69	1	M	D & P	36	1.08
Dunn (1974)						
Study I	Maryland, 1972-73	4	M	D & P	11	0.45
Study II	Maryland, 1972-73	6	M	D & P	30	0.44
Durward (1973)	Canada, 1973	6,7	M	D & P	6	0.19
Easterling (1982)	Texas, 1982	5	C	D & P	16	0.02
Fejfar (1969)	Indiana, 1967-68	4	M	D & P	4	0.80
Fletcher & Atkinson (1972)	California, 1969-70	1	L	D & P	22	0.81
Grocke (1982)	Australia, 1981	1-6	L	D & P	4	0.82
Haberman (1977)	Pennsylvania, 1974	4-6	M	D & P	8	0.57
Israel (1968)	New York, 1967-68	1,2	L	Tutorial	36	0.16
Litman (1977)	Illinois, 1972-73	4-6	L	D & P	36	0.23
MacLean (1974)	Pennsylvania, 1970-71	3-5	M	Tutorial	3	0.62
Mendelsohn (1972)	New York, 1968-69	2-6	M	D & P	20	0.49
Metric Associates (1981)	Massachusetts, 1980-81	1-6	C	D & P	36	0.32
Miller (1984)	Oregon, 1981-83	5-8	M	D & P	36	0.38
Morgan, Sangston & Pokras (1977)	Maryland, 1975-76	3-6	M	D & P	60	0.23
Nabors (1974)	Missouri, 1973	5-6	S	Management	36	0.36
Palmer (1973)	California, 1972-73	4-6	M	D & P	16	0.36
Prince (1969)	Mississippi, 1967-68	1-6	M	D & P	36	0.64
Regosta (1983)	California, 1977-81	1-6	C	D & P	108	0.30
Roberts (1982)	Utah, 1977-81	3-6	L	Management	108	0.02
Suppes & Morningstar (1969)	California, 1966-68	1-6	M	D & P	22	0.28
Suppes & Morningstar (1970)	New York, 1968-69	1-6	M	D & P	18	0.44 ^a
Thompson (1973)	Texas, 1970-71	4-6	L	D & P	36	0.11
Warner (1979)	Ohio, 1978-79	6	M	Tutorial	36	1.31
Wilson & Fitzgibbon (1970)	Michigan, 1969-70	4,5	L	D & P	16	0.40

Note. M = mathematics; L = language arts; S = science; C = combined; D & P = drill-and-practice.

^aThese studies yielded statistically significant effects, but the reports did not include enough detail for direct calculation of ES. The ES reported here and used in the analysis is an estimated value; it is the median ES in all studies of CBE that reported a statistically significant effect.

experiment were coded. When results from a long and short CBE implementation were available, results from the longer implementation were used. When transfer effects of CBE were measured in a study in addition to effects in the area of instruction, the direct effects were coded for the analysis. In all other cases, our procedure was to use total score and total group results rather than subscore and subgroup results in calculating *ES*.

Statistical analyses were carried out separately for each outcome area. One study feature—control for historical effects—was dropped from the statistical analyses because of the lack of variation among studies in this feature. In all but one study, both the experimental and the control treatments were administered concurrently. Another study feature, type of computer interaction, was eliminated from the analyses because of its unacceptably high correlation with the variable indicating type of computer-based instruction. All studies of computer-managed instruction involved off-line use of the computer, whereas no study of drill-and-practice and tutorial instruction involved off-line instruction.

RESULTS

Of the 32 studies used in this analysis, 28 reported results from computer-assisted instruction (CAI) programs, involving drill-and-practice or tutorial instruction. Only four studies reported results from computer-managed instruction (CMI). The two sets of studies differed strikingly in their study features. In the 28 CAI studies, for example, students used the computer interactively; in the 4 CMI studies, the computer processed student records off-line. But more important, preliminary examination of results showed that the CMI and CAI studies produced strikingly different results (Table 2). For this reason, results from CMI and CAI studies were analyzed separately.

Computer-Managed Instruction

The achievement of the control students exceeded slightly the achievement of students taught with computer management in two studies (Akkerhuis, 1974; Coffman & Olsen, 1980), but the difference between groups in these studies was trivial and nonsignificant. The achievement of CMI students was trivially higher than that of control students in a study by Roberts (1982), but again the difference between groups was nonsignificant. In a study of CMI by Nabors (1974), however, the effect of CBE was positive and moderately high. The average *ES* in the four implementations, however, was 0.07. The standard deviation of *ES* was 0.196, and the standard error was 0.10.

The four studies provided little evidence for other positive effects of CMI. Only the study by Akkerhuis (1974) examined noncognitive outcomes of instruction. In Akkerhuis's study the average *ES* on attitude toward subject was -0.20 , and the average *ES* on attitude toward computers was -0.07 . Both values are small or trivial in size, and neither can be considered statistically significant.

Computer-Assisted Instruction

The effects of CAI were clearly more positive than those of CMI. The clearest results were available on end-of-course achievement measures, but in other areas also results were basically positive.

Achievement examinations. In each of 28 studies with results from achievement examinations, students from the CAI class received the better examination scores; in no study did students from the conventional class get better scores on a final examination on course content. A total of 23 of these 28 studies reported, in addition, that the difference between CAI and conventional classes was statistically significant. Overall, these box-score results strongly favored CBE.

The index of effect size *ES* provides a more exact picture of the degree of benefit from CAI in the typical study. The average *ES* in the 28 studies was 0.47; the standard deviation of *ES* was 0.29; and its standard error was 0.055. The average *ES* for these CAI studies was significantly different from the average *ES* for CMI studies, $t(30) = 2.57, p < .05$.

The average *ES* of 0.47 for the CAI studies means that in the typical study, performance of CAI students was raised by 0.47 standard deviations. To interpret this effect more fully, it is useful to refer to areas of the standard normal curve. Approximately 68% of the area of this curve falls below a *z* score of 0.47. We can conclude, therefore, that students from CAI classes performed at the 68th percentile on their examinations, whereas the students who received only conventional instruction performed at the 50th percentile on the same examinations. Or put in another way, 68% of the students from CAI classes outperformed the average student from the control classes.

Study features and achievement effects. Although the effect of CAI was moderate in the typical study, the size of effect varied from study to study. Effects of CAI ranged in size from high of 1.3 standard deviations (Warner, 1979), to a low of 0.02 standard deviations (Easterling, 1982). It seemed possible that this variation in study outcome might be systematic, and we therefore carried out further analyses to determine whether different types of studies were producing different results. These analyses, however, did not disclose any significant relationships between study features and final achievement scores (Tables 3 and 4).

Subgroup and subscore achievement effects. Several studies that were coded for overall achievement effects provided, in addition, information on scores of specific subgroups of students and on subtest scores as well as total test scores. Table 5 presents separate effect sizes for high- and low-ability students, primary and middle grade students, and scores on language and mathematics subtests in studies that reported more differentiated results.

Five studies examined effects of CBE on both high- and low-aptitude students. The average effect of CBE in these studies was to raise achievement test scores of low-aptitude students by 0.41 standard deviations and to raise scores of high-aptitude students by 0.08 standard deviations. Although the average effect on low-ability students appeared to be larger than the average effect on high-ability students, the difference between the two effects was not statistically significant, $t(4) = 1.32, p > .10$. Eight studies examined effects on both primary and middle grade students. Although the average effect in the primary grades ($M = 0.39$) appeared to be larger than the average effect in the middle grades ($M = 0.25$), the difference was not statistically significant, $t(7) = 2.07, p > .05$. Finally, the five studies that examined language and arithmetic results did not produce evidence for differential effects. The effect on language tests ($M = 0.17$) in these studies was very similar to the effect on mathematics tests ($M = 0.16$), $t(4) = 0.08, p > .10$.

Table 3. Means and Standard Errors of Effect Sizes for 28 CAI Studies Classified by Study Features

Categories	N	Effect Size	
		M	SE
<i>Use of Computer</i>			
Drill-and-practice	24	0.44	0.05
Tutorial	4	0.63	0.24
<i>Author of Program</i>			
Local	10	0.45	0.08
Other	18	0.48	0.08
<i>CCC Material</i>			
Yes	15	0.46	0.06
No	13	0.48	0.10
<i>Duration of Instruction</i>			
One semester or less	12	0.47	0.07
One semester-one year	14	0.49	0.10
More than one year	2	0.26	0.03
<i>Subject Assignment</i>			
Random	5	0.57	0.21
Nonrandom	23	0.44	0.05
<i>Instructors</i>			
Same	7	0.44	0.09
Different	21	0.48	0.07
<i>Time-on-task</i>			
Experimental > Control	16	0.44	0.08
Experimental = Control	12	0.50	0.08
<i>Test Author Bias</i>			
Commercial test	22	0.45	0.06
Local test	6	0.52	0.11
<i>Evaluator Involvement</i>			
Involved	19	0.51	0.06
Not involved	9	0.37	0.10
<i>Field-tested Material</i>			
Yes	20	0.47	0.07
No	8	0.45	0.09
<i>Class Level</i>			
Primary (1-3)	4	0.62	0.20
Primary & middle (1-6)	11	0.42	0.06
Middle (4-6)	13	0.45	0.09
<i>Course Content</i>			
Mathematics	17	0.54	0.07
Language/reading	7	0.42	0.11
Combined	4	0.21	0.08
<i>Ability of Subjects*</i>			
Low	15	0.44	0.07
Average/mixed	13	0.51	0.09
<i>Nature of publication</i>			
Unpublished	10	0.46	0.11
Dissertation	8	0.34	0.08
Published	10	0.57	0.08
<i>Year of publication</i>			
Before 1969	5	0.46	0.11
1970-1974	12	0.48	0.08
1975-1979	6	0.53	0.17
1980-1984	5	0.37	0.13

* Populations examined in the 28 studies were either disadvantaged or representative of the school population in ability. None of the studies was restricted to a gifted, or high-ability population.

Table 4. Relationships of Time-on-Task Variables to Achievement Effects in 28 CAI Studies

Variable	<i>M</i>	Range	<i>r</i> with <i>ES</i>
Sessions per week	4.18	1–5	0.08
Number of weeks	26.43	3–108	–0.18
Minutes per session	14.96	5–30	0.08
Total time-on-task in hours	25.96	1.33–180	–0.10

Table 5. Effects Sizes for Subgroups and Subtests within Studies

Study	Ability		Content		Grades	
	Low	High	Language	Mathematics	Primary	Middle
Chiang, Stauffer, & Cannara (1978)			0.24	0.14		
Coffman & Olsen (1980)	–0.17	0.15	–0.01	–0.01	0.14	–0.17
Davies (1972)					0.18	0.39
Dunn, Morgan, & Richardson (1974)— Study I	0.55	0.35				
Durward (1973)	0.54	0.08				
Easterling (1982)			0.01	0.02		
Mendelsohn (1972)					0.62	0.40
Metric Associates (1981)			0.50	0.15	0.42	0.28
Morgan, Sangston, & Pokras (1977)	0.23	0.13			0.17	0.24
Prince (1969)	0.89	–0.32			0.87	0.44
Ragosta (1983)			0.10	0.49	0.35	0.25
Suppes & Morningstar (1969)					0.40	0.16

Follow-up examinations. Five of the CAI studies reported results from follow-up achievement examinations administered after the completion of computer-based and conventional teaching. In each of the studies, the follow-up scores were higher in the CAI class than in the conventional class. Delon (1970) reported a follow-up *ES* of 0.30; Dunn, Morgan, and Richardson (1974), an *ES* of 0.47; Litman (1977), an *ES* of 0.08; Prince (1969), an *ES* of 0.38; and Miller (1984), an *ES* of 0.14. The average *ES* in the five studies was 0.27; the standard deviation of *ES* was 0.16, and the standard error was 0.07.

Attitudes toward subject. Only one of the CAI studies presented student attitude results in a fashion that yielded an *ES* (Cranford, 1976). That study showed a small and statistically nonsignificant ($ES = 0.10$) positive effect of CBE on student attitudes towards mathematics.

DISCUSSION

The major finding in this study was the positive effect that CAI had on achievement of elementary school children. In the typical application, students received approximately 26 hours of CAI: 15 minutes per day, for 4 days a week, and for a total of 26 weeks. The effect of this instruction was to raise student achieve-

ment scores (relative to conventional instruction) by 0.47 standard deviations, or from the 50th to the 68th percentile. This effect is similar to effects reported in earlier reviews. Burns (1981) and Hartley (1978), for example, each reported gains of approximately 0.4 standard deviations from CAI in elementary school mathematics. The closeness in results from these different meta-analyses is impressive in view of the minimal overlap in the studies they covered and the different methodologies used in the analyses.

The results of this meta-analysis lend some support to the notion that CBE effects differ at different instructional levels (Kulik, 1981). CAI, for example, appears to have its strongest effects in elementary schools; its effects appear to be weaker in secondary schools and weakest in colleges. The average effect of CAI in elementary school programs examined in this analysis was to raise achievement scores by 0.47 standard deviations. Bangert-Drowns, Kulik, and Kulik (1985) found an average effect of 0.36 standard deviations for CAI in secondary schools, and Kulik and Kulik (1985) found an average effect of 0.26 standard deviations in colleges and universities. High school and college students apparently have less need for the highly structured, highly reactive instruction provided in computer drills and tutorials. They may be able to acquire basic textbook information without the cues and feedback that CAI systems provide.

Results from only four studies were available to assess the impact of CMI at the elementary school level. Although the reader must be cautious given the limitations of this information, it appears that CMI is less effective with younger learners and better suited to older students. The average effect of CMI in elementary schools was to raise student test scores by 0.07 standard deviations, whereas its average effect in secondary schools was to raise scores by 0.40 standard deviations (Bangert-Drowns et al., 1985) and its average effect in colleges was to raise scores by 0.35 standard deviations (Kulik & Kulik, 1985). This pattern may be attributable to the features of the computer exploited in most CMI systems. In CMI, the computer simply acts as the teacher's clerk; it scores tests, keeps records, and arranges schedules. Like other forms of individualized instruction, CMI requires learners to pace themselves properly, work independently, and make their own choices. These requirements may exceed the abilities and motivation of very young learners; they may be more appropriate requirements for older students. It is not surprising therefore that the greatest successes of noncomputerized individual instructional packages have come at the higher grade levels (Bangert, Kulik, and Kulik, 1983; Kulik, Kulik, & Cohen, 1979).

Like other meta-analyses carried out in recent years, this one did not find strong relations between study features and outcomes. Studies with different features produced similar outcomes. *ESs* were very similar, for example, in true experiments and quasi-experiments. In other meta-analyses that our research team has carried out on effects of instructional technology, results from true and quasi-experiments have been nearly identical (Bangert-Drowns, Kulik, & Kulik, 1984), and so we were not surprised by our failure to find differences on this study feature. Results with other study features, however, were less predictable. Kulik, Bangert, and Williams (1983) have suggested, for example, that CBE programs are growing more effective with time. In the present meta-analysis, *ESs* from different time periods were very similar. Other researchers have speculated that time-on-task might explain some of the variation in outcomes of

computer-based teaching (e.g., Suppes & Morningstar, 1969). In this meta-analysis, *ESs* were very similar for implementations with and without controls for time-on-task.

It is important to note that the record of CBE effectiveness that we have described rests on specific computers used in specific ways for specific purposes. The record may not apply to machines, approaches, and objectives not examined in the studies in our meta-analysis. Current interest in instructional computing, for example, has been stimulated greatly by the development of microcomputers in the last 15 years. Microcomputer-based systems have their own characteristics: their own software, their own management systems, and their own scale of operations. Only 1 of the 32 studies located for this meta-analysis examined the effects of a microcomputer-based system. Evaluators will have to give much more attention in the future to effects of such microcomputer-based systems.

As dramatic as changes in hardware have been in recent years, they are no more important than the changes that have occurred in conceptions of computer-based teaching. Early applications of the computer in teaching emphasized computer tutoring and drills. Computers presented lessons, provided drills on course material, and kept records. Recent developments have broadened this conception of the computer's role in education. In addition to serving a tutorial function, today's computers also assist in instruction when they perform routine calculating and clerical tasks for students and even when they serve as "tutees." Our computer searches produced no adequate evaluation studies of these exciting new developments in computer-based elementary-school teaching. Evaluation work on these areas is badly needed.

Evaluators also need to investigate a wider range of educational outcomes than they have in the past. They have repeatedly examined computer effects on achievement scores, but they have given inadequate attention to the computer's effects on attitudes toward school, attitudes toward computers, and instructional efficiency. In addition, evaluators of CBE have given almost no attention to the computer's effect on acquisition of higher order skills, transfer of gains to other areas, and interpersonal outcomes of computer uses in the classroom. If educational evaluators turn their attention to such matters, the years ahead will produce evaluation studies that better reflect the breadth of computer uses in education and the breadth and depth of goals of elementary education.

REFERENCES

- Akkerhuis, G. (1974). A comparison of pupil achievement and pupil attitudes with and without the assistance of batch computer-supported instruction. *Dissertation Abstracts International*, 4, 6345A. (University Microfilms No. 74-09,052)
- Atkinson, R.C. (1969). Computerized instruction and the learning process. In R.C. Atkinson & H.A. Wilson (Eds.), *Computer-assisted instruction: A book of readings* (pp. 143-165). New York: Academic Press.
- Bangert, R.L., Kulik, J.A., & Kulik, C.-L.C. (1983). Individualized systems of instruction in secondary schools. *Review of Educational Research*, 53, 143-158.
- Bangert-Drowns, R.L., Kulik, J.A., & Kulik, C.-L.C. (1984). *The influence of study features on outcomes of educational research*. Paper presented at the annual meeting of the American Psychological Association, Toronto.
- Bangert-Drowns, R.L., Kulik, J.A., & Kulik, C.-L.C. (1985). *Effectiveness of computer-based education in secondary schools*. Ann Arbor, MI: The University of Michigan, Center for Research on Learning.

- Becker, H.J. (1983). *School uses of microcomputers: Reports from a national survey* (Issue No. 1). Baltimore, MD: The Johns Hopkins University, Center for Social Organization of Schools.
- Burns, P.K. (1981). A quantitative synthesis of research findings relative to the pedagogical effectiveness of computer-assisted instruction in elementary and secondary schools. *Dissertation Abstracts International*, **42**, 2946A. (University Microfilms No. 81-28,378)
- Chiang, A., Stauffer, C., & Cannara, A. (1978). *Demonstration of the use of computer-assisted instruction with handicapped children, final report*. (Report No. 446-AH-60076A). Arlington, VA: RMC Research Corporation. (ERIC Document Reproduction Service No. ED 166 913)
- Coffman, W.E., & Olsen, S.A. (1980). *The first two years of PLAN: An evaluation of program impact*. Iowa City, IA: Iowa Testing Programs. (ERIC Document Reproduction Service No. ED 190 674)
- Cranford, H.R. (1976). A study of the effects of computer assisted instruction in mathematics on the achievement and attitude of pupils in grades five and six in a rural setting. *Dissertation Abstracts International*, **37**, 5660A. (University Microfilms No. 77-5932)
- Davies, T.P. (1972). An evaluation of computer assisted instruction using a drill and practice program. *Dissertation Abstracts International*, **32**, 6970B. (University Microfilms No. 72-18,627)
- Delon, F.G. (1970). A field test of computer assisted instruction in first grade mathematics. *Educational Leadership*, **28**, 170-180.
- Dunn, A., Morgan, C.E., & Richardson, W.M. (1974). *Computer-assisted instruction program. A three year report covering July 1, 1971 through June 30, 1974*. Rockville, MD: Montgomery County Public Schools. (ERIC Document Reproduction Service No. ED 100 361)
- Durward, M. (1973). *Computer assisted instruction in arithmetic at South Hill Elementary School*. British Columbia: Vancouver Board of School Trustees. (ERIC Document Reproduction Service No. ED 088 915)
- Easterling, B.A. (1982). The effects of computer assisted instruction as a supplement to classroom instruction in reading comprehension and arithmetic. *Dissertation Abstract International*, **43**, 2231 A. (University Microfilms No. 82-23, 032)
- Edwards, J., Norton, S., Taylor, S., Weiss, M., & Dusseldorp, R. (1975). How effective is CAI? A review of the research. *Educational Leadership*, **33**, 147-153.
- Fejfar, F.L. (1969). ISU lab school fourth graders learn through CAI. *Contemporary Education*, **40**, 296-297.
- Fletcher, J.D., & Atkinson, R.C. (1972). Evaluation of the Stanford CAI program in initial reading. *Journal of Educational Psychology*, **63**, 597-602.
- Glass, G.V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, **5**, 3-8.
- Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage Publications.
- Groccke, M. (1982). *Interactive Development of Reading Skills in an Educational Clinic*. Australia, paper presented at the Annual National Conference of the Australian Group for the Scientific Study of Mental Deficiency. (ERIC Document Reproduction Service No. ED 223 993)
- Haberman, E.L. (1977). Effectiveness of computer assisted instruction with socially/emotionally disturbed children. *Dissertation Abstracts International*, **38**, 1998A. (University Microfilms No. 77-21, 221)
- Hall, K.A. (1982). Computer based education. In H.E. Mitzel (Ed.), *Encyclopedia of educational research* (pp. 353-363). New York: The Free Press.
- Hartley, S.S. (1978). Meta-analysis of the effects of individually paced instruction in mathematics. *Dissertation Abstracts International*, **38**(7-A), 4003. (University Microfilms No. 77-29,926)
- Hedges, L.V., & Olkin, I. (1980). Vote counting methods in research synthesis. *Psychological Bulletin*, **88**, 359-369.
- Israel, B.L. (1968). *Responsive environment program: Brooklyn, N.Y.: report of the first full year of operation. The talking typewriter*. Brooklyn, NY: New York City Board of Education, Responsive Environment Program Center. (ERIC Document Reproduction Service No. ED 027 742)
- Kulik, C.-L.C., & Kulik, J.A. (1985). *The effectiveness of computer-based education in colleges*. Ann Arbor, MI: The University of Michigan, Center for Research on Learning and Teaching.
- Kulik, J.A. (1981). *Integrating findings from different levels of instruction*. Paper presented at the American Educational Research Association, Los Angeles. (ERIC Document Reproduction Service No. ED 208 040)
- Kulik, J.A., Bangert, R.L., & Williams, G.W. (1983). Effects of computer-based teaching on secondary school students. *Journal of Educational Psychology*, **75**, 19-26.

- Kulik, J.A., Kulik, C.-L. C., & Cohen, P.A. (1979). A meta-analysis of outcome studies of Keller's Personalized System of Instruction. *American Psychologist*, **34**, 307-318.
- Litman, G. (1977). Relation between computer-assisted instruction and reading achievement among fourth, fifth, and sixth grade students. *Dissertation Abstracts International*, **38**, 2003A. (University Microfilms No. 77-20,883)
- MacLean, R.F. (1974). A comparison of three methods of presenting instruction in introductory multiplication to elementary school children (total computer, partial computer, and non-computer). *Dissertation Abstracts International*, **35**, 1430. (University Microfilms No. 74-19,759)
- Mendelsohn, M. (1972). CAI in New York City: The slow learner in mathematics. *National Council of Teachers of Mathematics Yearbook*, 355-364.
- Metric Associates. (1981). *Evaluation of the computer-assisted instruction Title I Project, 1980-81*. Chelmsford, MA: Metric Associates. (ERIC Document Reproduction Service No. ED 233 122)
- Miller, S.W. (1984). A comparison of computer-assisted instruction with prescription learning and the traditional "pull-out" program used in supplementing of mathematics basic skills to Chapter I (Title I) students. *Dissertation Abstracts International*, **44**, 2397A. (University Microfilms No. 83-26,453)
- Morgan, C.E., Sangston, B.J., & Pokras, R. (1977). *Evaluation of Computer-Assisted Instruction, 1975-76*. Rockville, MD: Montgomery County Public Schools. (ERIC Document Reproduction Service No. ED 139 655)
- Nabors, D.G. (1974). A comparative study of academic achievement and problem-solving abilities of black pupils at the intermediate level in computer-supported instruction and self-contained instructional programs. *Dissertation Abstracts International*, **36**, 3241A. (University Microfilms No. 75-26,294)
- Oettinger, A.B. (1969). *Run, computer, run: The mythology of educational innovation*. New York: Collier Books.
- Palmer, H. (1973). *Three evaluation reports of computer-assisted instruction in drill and practice mathematics*. Los Angeles, CA: Los Angeles County Superintendent of Schools. (ERIC Document Reproduction Service No. ED 087 422)
- Papert, S. (1980). *Mindstorms*. New York: Basic Books.
- Prince, J.D. (1969). *A practitioner's report results of two years of computer-assisted instruction in drill and practice mathematics*. McComb, MS, McComb Schools. (ERIC Document Reproduction Service No. ED 032 769)
- Ragosta, M. (1983). Computer-assisted instruction and compensatory education: A longitudinal analysis. *Machine-Mediated Learning*, **1**, 97-127.
- Roberts, A.S. (1982). The effects of split-day scheduling and computer-managed instruction on the reading achievement of intermediate students. *Dissertation Abstracts International*, **43**, 1482A. (University Microfilms No. 82-23,584)
- Suppes, P., & Morningstar, M. (1969). Computer-assisted instruction. *Science*, **166**, 343-350.
- Suppes, P., & Morningstar, M. (1970). Technological innovations: Computer-assisted instruction and compensatory education. In F.F. Korten, S.W. Cook, & J.I. Lacey (Eds.), *Psychology and the problems of society* (pp. 221-236). Washington, DC: American Psychological Association.
- Taylor, R.P. (Ed.) (1980). *The computer in the school: Tutor, tool, tutee*. New York: The Teachers College Press.
- Thompson, B.B. (1973). Effect of computer-assisted instruction on the language arts achievement of elementary school children. *Dissertation Abstracts International*, **33**, 4077-4078A. (University Microfilms No. 73-3574)
- Vinsonhaler, J.F., & Bass, R.K. (1972). A summary of ten major studies on CAI drill and practice. *Educational Technology*, **12**, 29-32.
- Warner, T. (1979). *The Jackson elementary school computer based education mathematics project*. Akron, OH: The University of Akron, Computer Based Education Center.
- Watson, P.G. (1972). *Using the computer in education*. Englewood, NJ: Educational Technology Publications.
- Wilson, H.A., & Fitzgibbon, N.H. (1970). Practice and perfection: A preliminary analysis of achievement data from the CAI elementary English program. *Elementary English*, **47**, 576-579.