

Review of Recent Research Literature on Computer-Based Instruction

JAMES A. KULIK AND CHEN-LIN C. KULIK

Center for Research on Learning & Teaching, The University of Michigan

Separate meta-analyses have been completed recently on the effectiveness of computer-based education (CBE) in elementary schools, high schools, colleges, and nontraditional postsecondary institutions. Several general conclusions can be drawn from these meta-analyses. First, CBE programs usually have positive effects, as measured by several different criteria of instructional effectiveness. Effects are not uniformly high, however, for all types of CBE programs at all instructional levels. Second, effects vary as a function of the evaluation design used to measure CBE effectiveness. And third, study outcomes also vary as a function of the type of publication in which the results are found. Important questions about ways to improve instruction now await increased research attention. © 1987 Academic Press, Inc.

The history of education is largely a story of gradual evolution, but education has also had its revolutions. The first use of writing as a tool in teaching transformed education many centuries ago when it freed teachers from the constraints of oral tradition. The invention of printing in the 15th century made books widely available and had a similarly dramatic effect on the history of education. Now, in the 20th century, the invention of the computer may have an equally profound impact on education.

Educational developers long ago demonstrated that they could program computers to work in schools as drill masters, tutors, testers, and schedulers of instruction. But until recently the cost of computer-based teaching systems was too high for most schools. With the development of small, quick, inexpensive microcomputers during the last decade, computing costs have dropped dramatically, and a computer revolution in education now seems a real possibility.

Will such a revolution have a beneficial effect on education—as the revolutions based on writing and printing did—or will it have negative effects? Most researchers believe that there is one straightforward way to answer this question. And that is to compare the performance of students who have learned with and without computer assistance. Outcome

Requests for reprints should be sent to Dr. James A. Kulik, Center for Research on Teaching and Learning, 109 East Madison Street, Ann Arbor, MI 48109. The material in this report is based upon work supported by a grant from the Exxon Education Foundation and by National Science Foundation Grant MDR-8470258.

studies are commonly understood to be the basic way of determining whether instructional innovations help or hurt students.

The earlier revolutions in education ran their courses without the benefit of such studies. No one tried to measure the effects of writing or of books when they were first introduced as tools in education. No one used statistical methods to predict or influence results. Earlier educational revolutions just happened—without measurement, prediction, or control. The computer revolution is different. It is occurring at a time when educators have tools for evaluating programs and tools for drawing general conclusions from a collection of evaluations.

These tools have already been used in the evaluation of computer-based instruction. Researchers have carried out numerous outcome studies over the years, and reviewers have synthesized the findings on a number of occasions. The purpose of this article is simply to review what we have learned from this work. The article relies heavily on research syntheses that we have carried out with our colleagues at the University of Michigan.

QUANTITATIVE SYNTHESIS

The methodology that we have used in our research syntheses was originally used by Glass (1976), who gave it the name *meta-analysis*. By meta-analysis, Glass meant the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings. Researchers who carry out a meta-analysis first locate studies of an issue by clearly specified procedures. They then characterize the outcomes of all studies on a common scale of effect size. The effect size for a given study describes in standard deviation units the difference in performance of the experimental and control groups. Meta-analysts next describe study features in categorical or quasi-quantitative terms. Finally, they use multivariate techniques to describe findings and relate characteristics of studies to study outcomes.

We applied Glass's methodology in four separate meta-analyses (Bangert-Drowns, Kulik, & Kulik, 1985; Kulik & Kulik, 1986; Kulik, Kulik, & Shwalb, 1986; Kulik, Kulik, & Bangert-Drowns, 1985). These analyses covered a total of 199 comparative studies: 32 in elementary schools; 42 in high schools; 101 in universities and colleges; and 24 in adult education settings. The meta-analyses covered use of the computer in (a) computer-assisted instruction, or CAI, including drill-and-practice and tutorial instruction; (b) computer-managed instruction, or CMI; and (c) computer-enriched instruction, or CEI, including the use of the computer as a calculating device, programming tool, and simulator. Each of the 199 studies included in the meta-analyses was a controlled, quantita-

tive study that met our predefined standards for methodological adequacy.

OVERALL RESULTS

Most of the studies reported that computer-based instruction has positive effects on students (Table 1).

1. Students generally learned more in classes when they received help from computers. The average effect of computers in all 199 studies used in our meta-analyses was to raise examination scores by 0.31 standard deviations, or from the 50th to the 61st percentile.

2. Students also learned their lessons with less instructional time. The average reduction in instructional time in 28 investigations of this point was 32%.

3. Students also liked their classes more when they received computer help. The average effect of computer-based instruction in 17 studies was to raise attitude-toward-instruction scores by 0.28 standard deviations.

4. Students developed more positive attitudes toward computers when they received help from them in school. The average effect size in 17 studies on attitude toward computers was 0.33.

5. Computers did not, however, have positive effects in every area in which they were studied. The average effect of computer-based instruction in 29 studies of attitude toward subject matter was near zero.

STUDY FEATURES AND OUTCOMES

A few study features were consistently related to achievement outcomes of studies (see Table 2).

TABLE 1
AVERAGE EFFECT OF CBE ON STUDENTS IN 199 STUDIES

Outcome measure	Number of studies	Average effect ^a
Final examination	199	0.31
Attitude toward instruction	17	0.28
Attitude toward computers	17	0.33
Attitude toward subject	29	0.05
Instructional time	28	68%

^a The average effect is measured by the statistic ES for the following variables: final examination, attitudes toward computers, instruction, and subject matter. Savings in instructional time are measured in percentage of time saved (x/c).

TABLE 2
AVERAGE ES^a FOR FINAL EXAMINATION FOR SELECTED STUDY FEATURES

Study Feature	Number of studies	Average ES
Publication source*		
Published	65	0.46
Unpublished	131	0.23
Control for instructor effects**		
Control present	100	0.24
No control	86	0.40
Publication year**		
Before 1975	92	0.24
1975 and after	104	0.36
Duration of instruction***		
Fewer than 9 weeks	79	0.36
9 weeks or more	114	0.27

^a ES = effect size.

* $p < .01$.

** $p < .05$.

*** $p > .10$.

1. Study results were consistently stronger in published studies and weaker in unpublished ones ($p < .01$). The average effect of computer-based instruction in published studies was to raise student examination scores by 0.46 standard deviations, whereas its average effect in unpublished studies was to raise scores by only 0.23 standard deviations.

2. Effects were larger when different teachers taught the experimental and control groups ($p < .05$). Effects were smaller when the same teacher was responsible for both groups. With the same teacher in charge of experimental and control groups, average size of effect on examination scores was 0.24 standard deviations. With different teachers in charge of the groups, the average effect was 0.40 standard deviations.

3. Effects tended to be larger in more recent studies and smaller in older studies ($p < .05$). The average effect of computer-based instruction in studies published before 1975 was to raise examination scores by 0.24 standard deviations; the average effect in studies published in later years was a score increase of 0.36 standard deviations.

4. Effects were also somewhat larger in short duration studies and weaker in longer ones ($.10 < p < .20$). The average effect of computer-based instruction in short studies was to raise examination scores by 0.36 standard deviations, whereas its average effect in longer studies was to raise scores by 0.27 standard deviations.

Because study features were moderately intercorrelated, multiple regression analyses were carried out on study feature data. In the multiple

regression equation developed from the full data set, three of the four study factors had significant weights ($p < .05$): publication source, control for instructor effects, and study duration. The regression weight for the fourth study feature—study year—reached only a borderline level of significance ($p < .10$).

A few relationships between study features and effects held only at a specific instructional level. The most important of the relationships of this type involved the type of computer use. The effects of computer-enriched instruction, or CEI, were strikingly different at precollege and postsecondary levels of instruction. At the precollege level of instruction, CEI effects were near zero. In colleges and adult settings, CEI produced effects that were moderate in size. Young adults and older adults seemed to profit from CEI; precollege students seemed not to benefit from this type of computer use.

DISCUSSION

The message from evaluation studies on computer-based instruction seems clear enough. Computer-based instruction has raised student achievement in numerous studies. It has given students a new appreciation for technology and has had positive effects on student's attitudes toward schools and teaching. And computers have helped teachers save instructional time. These are the major findings from our meta-analyses, and they have been confirmed in research syntheses by other investigators.

Why have studies of computer-based instruction produced such positive results? Different factors may have contributed to the favorable picture in the literature. Among the factors that must be considered are the following:

1. *Editorial gatekeeping.* Journal editors and reviewers may prefer to publish strong and significant results rather than weak and insignificant ones.
2. *Experimental design flaws.* Design flaws in evaluation studies may allow researcher biases and expectations to color study results.
3. *Instructional quality.* The positive results from meta-analytic studies may reflect real differences in the quality of conventional and computer-based instruction.

Editorial Gatekeeping

It is often said that reviewers, researchers, and editors base their publication decisions on the size and statistical significance of study effects rather than on the quality of studies. Such editorial gatekeeping would prevent many nonsignificant and negative findings from being published,

and it would also give readers of the scientific literature a distorted picture of a field. Instead of reflecting reality, the scientific literature would reflect the preconceptions and biases of those who hold positions of authority in an area.

Some reviewers have suggested that the difference in published and unpublished results is merely the tip of the iceberg. What are results like, they have asked, in reports that end up in file drawers or in office wastebaskets? What are the results in studies that are abandoned in mid-stream? Some reviewers have suggested that such "studies" would give us an accurate picture of computer effects, and they have also advised researchers to discount the results in published studies.

Before throwing away the articles of some of our most renowned researchers, we should consider another possibility. The difference in results in published and unpublished reports may have another source. The authors of journal articles and dissertation studies, for example, are different individuals working under different circumstances. They differ in their research experience, in their resources, in their relationship to instructional developers, and in many other respects. Such differences can explain—just as well as editorial gatekeeping—the differences in results found in dissertations and journals.

It seems to us that we know far too little to choose between the competing interpretations at this time. The only thing that is certain is that the issue of editorial gatekeeping is an important one. Differences among results in published and unpublished papers are too common in the meta-analytic literature to be ignored. It is time for us to find out what factors cause these differences.

Experimental Design

Can flaws in experimental design explain—or explain away—the positive findings from studies of computer-based instruction? Some reviewers think so. They believe that with imperfectly controlled experiments, results are not interpretable. Among the factors that might systematically distort results in an imperfectly controlled evaluation are differences in time-on-task, self-selection differences in assignment to comparison groups, and uncontrolled teacher effects.

The evidence from our meta-analyses is that not all such factors are important. Time-on-task differences cannot explain the positive results of evaluations of computer-based instruction, for example. Studies that control for time-on-task have produced nearly the same results as studies without strict controls on instructional time. Actual records of instructional time have been collected in several studies, and these records suggest that students in computer groups often receive instruction for shorter periods than conventional students do.

Nor do our meta-analyses support the idea that subject assignment is an important methodological flaw in studies of computer-based instruction. Probably no other methodological point has received as much attention in evaluation research as the distinction between random experiments and quasi-experiments. Random experiments are generally thought to produce clear and consistent results; quasi-experiments are often thought to produce inconsistent and biased results. In our meta-analyses, random experiments and quasi-experiments produced the same results.

Results from studies with and without controls for instructor effects are somewhat different, however. In the typical study with the same instructor teaching experimental and control classes, the effect of computer-based teaching seemed modest. In the typical study with different instructors in experimental and control classes, the effect of computer-based teaching seemed more substantial.

Why should one-instructor and two-instructor experiments produce somewhat different results? It is not at all obvious to us. It may be, for example, that in two-instructor experiments, the poorer instructor is usually assigned to the control condition and the better instructor to the experimental condition, and the difference between conditions is magnified because of these teacher assignments. If this is the case, then one-instructor studies more accurately assess the effects of computer-based instruction. It may also be, however, that in one-instructor studies there is diffusion of the innovative treatment to the control condition. Involvement of a teacher in an innovative approach to instruction may have a general effect on the quality of the instructor's teaching. Outlining objectives, constructing lessons, preparing evaluation materials, and working with computer materials—requirements in computer-based instruction—may help a teacher to do a better job in a conventional teaching assignment. If this is the case, then two-instructor studies provide the better basis for estimating the size of an experimental effect.

Instructional Design

Computer-based instruction is often well-designed instruction. The hard work of an instructional design team often ensures the quality of computer materials. Objectives are usually clear and explicit. Instruction is carefully sequenced. The materials engage the learner's attention and encourage learner activity. The program provides frequent feedback to the learner. Instructional design teams often spend 100 h developing just 1 h of computer lessons.

Certain features of the computer make it an especially attractive medium for instructional designers. Computers can generate attractive and complex graphics quickly. Computers can simulate motion. They can

give undivided attention to a single learner. They can provide complex evaluations of a learner's performance. They can wait patiently. They can be programmed to model a learner's cognitive processes.

In certain respects, computer lessons seem to have an advantage over lessons presented by classroom teachers. Few classroom teachers can put 100 h of preparation time into each 1-h lesson. Classroom teachers cannot give each individual in a large classroom their continuous, undivided attention. Classroom teachers can be notoriously slow at grading student work and preparing reports. And their patience is often tried by their charges.

Can such differences account for the superior record of computer-based instruction in evaluation studies? They might. It is possible that the computer has fared so well in evaluation studies because programs of computer-based instruction have generally been well designed, and computers have delivered instruction in an attractive and engaging way. It is possible, in other words, that we should take the findings of studies of computer-based instruction at face value.

Evaluation studies of other kinds will be necessary to test this notion more fully. Studies must be carried out to address the question, Why has the computer helped students to the extent it has? How does it influence learning? Which features of the computer make it so effective? Is it the absolute consistency in the computer's response or the computer's complete impartiality? Does computer effectiveness stem from the novelty that the computer brings to instruction? Or is the immediacy of the computer's response the central thing? Are computer lessons equally effective when presented without electronic technology? How important is the social setting in which the computer is used? Process rather than outcome studies are needed to answer such questions, and process studies are needed in sufficient numbers for quantitative synthesis of results. It will take an enormous effort to produce the studies and to synthesize their findings, but, judging by what has already been achieved, the effort may prove to be worthwhile.

CONCLUSIONS

Among the conclusions that can be drawn from our meta-analyses, the following seem especially important:

1. Most programs of computer-based instruction have had positive effects on student learning. Future programs of implementation and development of computer-based instruction should therefore be encouraged. If such programs are as carefully designed as present ones are, they will most likely produce positive results.

2. Both journal articles and dissertations present a basically positive picture of computer-based instruction, but the findings reported in

journal articles are clearly more favorable. Research is badly needed to determine the factors that produce differences in journal and dissertation results. Does editorial gatekeeping lead professional journals to present a distorted picture of social science findings? Or do dissertation authors simply measure experimental effects less well than do more seasoned researchers?

3. Although a variety of different research designs can be used to show the effectiveness of computer-based instruction, certain research designs seem to produce more positive results. Studies where the same instructor teaches both experimental and control classes, for example, report somewhat weaker effects than do studies with different experimental and control teachers. Studies of long duration often report weaker effects than do short studies. Reasons for the difference in results from studies using different experimental designs are imperfectly understood, however. Research on such factors should be encouraged.

4. Computer-based instruction has not been uniformly successful in all its guises and at all instructional levels. Elementary and secondary schools have had a good deal of success in increasing student achievement through CAI programs; they have had less demonstrable success with CEI. College and adult courses have benefited moderately from CAI, CMI, and CEI. Future evaluations of computer-based instruction should try to explain such patterns of effectiveness.

REFERENCES

- BANGERT-DROWNS, R. L., KULIK, J. A., & KULIK, C.-L. C. (1985). Effectiveness of computer-based education in secondary schools. *Journal of Computer-Based Instruction*, 12, 59-68.
- GLASS, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8.
- KULIK, C.-L. C., & KULIK, J. A. (1986). Effectiveness of computer-based education in colleges. *AEDS Journal*, 19, 81-108.
- KULIK, J. A., KULIK, C.-L. C., & BANGERT-DROWNS, R. L. (1985). Effectiveness of computer-based education in elementary schools. *Computers in Human Behavior*, 1, 59-74.
- KULIK, C.-L. C., KULIK, J. A., & SHWALB, B. J. (1986). Effectiveness of computer-based adult education. *Journal of Educational Computing Research*, 2, 235-252.