Educational Evaluation and Policy Analysis
Spring 1988, Vol. 10, No. 1, pp. 51–69

Cost-Effectiveness and Educational Policy

Henry M. Levin
Stanford University

The present time of educational reform and budgetary stringency is a propitious one for improving resource allocation in education through a greater reliance on cost-effectiveness analysis. This article provides a summary of the technique and its applications to educational policy. It concludes that there is great potential for the use of cost-effectiveness applications in education, but there is little capability for doing so among most policymakers. Examples are provided of productive cost-effectiveness applications, and recommendations are made with regard to increasing the capacity of educational evaluators, policy analysts, and decisionmakers to use the tools appropriately for more efficient resource allocation.

The time has probably never been more propitious for incorporating cost-effectiveness analysis into educational policy. Cost-effectiveness analysis refers to the consideration of decision alternatives in which both their costs and consequences are taken into account in a systematic way. A combination of educational reform movements, declining enrollments in some jurisdictions and rising enrollments in others, exploration of ways to improve the education of a burgeoning population of disadvantaged students, and increasing pressure on educational budgets has created inducements to consider systematically both the costs and effects of alternative courses of action.

For example, the report of the National Commission on Excellence in Education (1983) has provided many recommendations for improving education. Even modest versions of the reforms were estimated to require $20 billion a year in additional budgetary resources for 1984, the year following the report, according to estimates made by Frances Kemmerer and Alan Wagner of the State University of New York in Albany. A similar concern is found in Odden (1984). But federal budgetary deficits have reduced the prospects for additional federal funding. Although state and local governments have responded generously (Gold, 1986), both levels of government are hard-pressed to expand educational funding in light of rising needs for other constituencies (e.g., health care and public assistance for the elderly and poor).

The result is that hard choices must be made in selecting from among many alternatives for educational reform and in responding to other educational contingencies such as rises and declines in enrollments and an increasing population of educationally disadvantaged students. That costs as well as educational results should be taken into account in making choices seems obvious.

A recent study evaluated the costs and effectiveness of four different interventions (cross-age tutoring, computer-assisted instruction, lengthening the school day, and reducing class size) for improving mathe-

This paper was prepared for the National Center for Policy Research in Education of Rutgers University, University of Wisconsin, Michigan State University, and Stanford University. Helpful comments were provided by Andy Porter and an anonymous referee.

matics and reading achievement at the elementary level (Levin, Glass, & Meister, 1984, 1987). The most cost-effective program, cross-age tutoring, was found to yield four times the increase in achievement for each dollar of cost as the least cost-effective program, lengthening the school day. Stated in another way, the least cost-effective approach among the four alternatives would have required about $400 per student to provide the same achievement gains attained by a $100 investment per student when allocated to the most cost-effective approach.

At the present time, cost-effectiveness analyses are rarely used in educational decisionmaking. Educational researchers and evaluators are trained in fields that do not incorporate cost analysis, so it is rare that costs are included in educational evaluations. Indeed, educational policymakers and practitioners are rarely familiar with the concepts, even though they often state that they seek cost-effective policies. This apparent contradiction is resolved when one realizes that they think of cost-effectiveness as a generic criterion associated with "good" programs rather than as a specific technique of evaluation that provides particular information on costs and effects.

What Is Cost-Effectiveness Analysis?

Before proceeding, it is important to differentiate the term *cost-effectiveness* from cost-benefit and cost-utility (Levin, 1983). Each of these describes an approach that can compare the costs and outcomes of educational alternatives, but cost-effectiveness assesses outcomes in educational terms (e.g., student achievement), cost-benefit assesses outcomes in terms of their monetary value, and cost-utility evaluates outcomes in terms of their subjective value to the decisionmaker. They share a common methodology in terms of cost measurement, but they differ in their measurement of educational outcomes.

Economists have traditionally had a greater interest in cost-benefit analysis, because it serves as an overall guide for making public investments (Gramlich, 1981). In contrast, cost-effectiveness is more useful to the evaluator or policymaker in a particular area of policy, such as education, who must determine the least cost approach to meeting such educational objectives as reducing dropouts or raising reading achievement. Cost-utility analysis is most useful to the administrator who must choose among alternatives with very incomplete information on effects or benefits (Keeney & Raiffa, 1976).

Early cost-benefit studies in education evaluated the relations among different types and levels of education on the earnings of workers to ascertain the value of reducing dropouts (Weisbrod, 1965), of vocational education (Hu, Lee, & Stromsdorfer, 1971), or of undertaking additional education (Becker, 1964). These studies are more appropriately termed cost-benefit studies because they compare the costs and benefits of different investment alternatives in education where both costs and benefits are expressed in monetary terms.

But, most decision problems facing educational policymakers and administrators tend to focus on program and resource decisions where the outcomes can be assessed only in educational terms or in terms of the likelihood of reaching specific educational objectives. For example, typical program goals include increasing student achievement, reducing dropouts, raising participation in particular courses of study, and so on. Outcome measures include tests of student achievement, dropout rates, and levels of student participation, but it is not always possible to readily convert such results into dollar benefits. Accordingly, the use of cost analysis for educational decisionmaking has tended to focus primarily on cost-effectiveness where the costs and educational effects of different alternatives can be compared. Only recently have the implications of cost utility analysis for education been developed (Hawley, Fletcher, & Piele, 1986; Levin, 1980, 1983).

Cost-effectiveness studies have been carried out on teacher selection (Levin, 1970), educational television and radio (Jamison, Klees, & Wells, 1978), choice of a mathematics curriculum (Quinn, VanMondfrans, & Worthen, 1984), computer-assisted instruction (Hawley, Fletcher, & Piele, 1986;

Levin, Glass, & Meister, 1984, 1987; Levin, Leitner, & Meister, 1986), and increasing the school day, reducing class size, and cross-age tutoring (Levin, Glass, & Meister, 1984, 1987).

The basic technique has been to derive results for educational effectiveness of each alternative by using standard evaluation procedures or studies (Rossi & Freeman, 1985) and to combine such information with cost data that are derived from the ingredients approach. The ingredients approach was developed to provide a systematic way for evaluators to estimate the costs of social interventions (Levin, 1975, 1983). It has been applied not only to cost-effectiveness problems, but also to determining the costs of different educational programs for state and local planning (Chambers, 1981; Hartman, 1981).

Measuring Cost-Effectiveness

The main strength of the ingredients method is that it is based upon a straightforward approach to estimating costs that is comprehensible to evaluators and policymakers while meeting rigorous standards of economics methodology. The approach can be mastered by persons without substantial technical background in economics and cost methods, so that only minimal technical advice or counsel is needed. Although the approach is outlined and applied in some detail in Levin (1983), it is useful to set out its basic stages.

Before doing the cost-analysis, we need to know what the decision problem is, how to measure effectiveness, which alternatives are being considered, and what their effects are. We assume that a problem has risen on the policy agenda that requires a response. A careful understanding of the problem is crucial to addressing its solution (Levin, 1983, pp. 34–35). Once the problem has been formulated, it will be necessary to consider how to assess the effectiveness of a solution. For this purpose, clear dimensions and measures of effectiveness will be needed. Table 1 provides some hypothetical measures of effectiveness for typical educational program objectives.

Given the problem and criteria for assessing the effectiveness of proposed solutions, it is necessary to formulate alternative programs or interventions. The search for such interventions should be as wide-ranging and creative as possible. This procedure sets the stage for the evaluation of effectiveness of the alternatives, a process akin to the standard use of evaluation methods (e.g., Rossi & Freeman, 1985). Estimates of effectiveness can be derived from previous evaluations or from tailored evaluations for the present purpose. Obviously, the latter is far more costly in both time and resources, but it may be desirable if previous studies are not avail-

TABLE 1

Examples of effectiveness measures

Program objective	Measure of effectiveness
Program completions	Number of students completing program
Reducing dropouts	Number of potential dropouts who graduate
Employment of graduates	Number of graduates placed in appropriate jobs
Student learning	Test scores in appropriate domains utilizing appropriate test instruments
Student satisfaction	Student assessment of program on appropriate instrument to measure satisfaction
Physical performance	Evaluation of student physical condition and physical skills
College placement	Number of students placed in colleges of particular types
Advance college placement	Number of courses and units received by students in advance placement, by subject

able and if the resource implications of alternatives are substantial. Clearly, the value of any evaluation must depend on what is at stake and whether the cost of doing an evaluation is justified (Alkin & Solmon, 1983; Levin, 1983, chap. 2).

It is important to emphasize that the evaluation of effectiveness is separable from the evaluation of costs. Most standard evaluation designs for assessing the effectiveness of an intervention are also suitable for incorporation into cost-effectiveness studies. These can be found in the standard evaluation literature (see, e.g., Cook & Campbell, 1979; Rossi & Freeman, 1985). The cost analysis is not typically found in the general evaluation literature and has been developed independently as a subspecialization (Levin, 1983).

Cost Estimation

The costs of an intervention are defined as the value of the resources that are given up by society to effect the intervention. These are referred to as the ingredients of the intervention, and it is the social value of those ingredients that constitute its overall cost. At a later stage one can assess the distribution of these costs among the decisionmaking agency and other entities. Accordingly, the method sets out systematically to identify and ascertain the value of the ingredients that are required for each alternative under consideration.

The ingredients approach to cost estimation entails three distinct phases: (1) identification of ingredients, (2) determination of the value or cost of the ingredients and the overall costs of an intervention, and (3) an analysis of the costs in an appropriate decision-oriented framework.

1. *Identification of ingredients.* The first step is to ascertain which ingredients are required for an intervention (Levin, 1983, chap. 3). Most educational interventions are labor-intensive, so an initial concern is to account for the number and characteristics of personnel. Clearly, we are concerned with whether personnel are part-time or full-time and the types of skills or qualifications that they need. Beyond this we need to identify the facilities, equipment, materials, and other ingredients or resources which are required for the intervention.

Identification of ingredients requires a level of detail that is adequate to ensure that all resources are included and are described adequately to place cost values on them. For this reason, the search for ingredients must be systematic rather than casual. The primary sources for such data are written reports, observations, and interviews. Written reports usually contain at least a brief history and description of the intervention. Some also provide a detailed list of personnel and other resources. These can be used to construct the basic list of required ingredients. Our experience is that the information in such reports is often incomplete and lacks sufficient detail on qualitative aspects of the ingredients. For example, the report may list a number of teachers, coordinators, and other personnel, without stipulating such important details as their levels of experience, training, education, or other special personnel requirements.

Accordingly, other sources of information must be used to corroborate and supplement data on ingredients from evaluations and descriptive reports. If the intervention is present at a nearby site, it may be possible to visit and gather additional data on ingredients through observation. A third valuable source is interviews, where present or former personnel are asked to identify resources from among a number of different classifications. The three principal types of information, reports, observations, and interviews can be used to assure the accuracy of the data by comparing the findings from each source and reconciling differences: the process of triangulation.

2. *Determination of costs.* Once the ingredients have been identified and stipulated, it is necessary to ascertain their costs (Levin, 1983, chap. 4). In doing this, all ingredients are assumed to have a cost, including donated or volunteer resources. That is, they have a cost to someone, even if the sponsoring agency did not pay for them in a particular situation. At a later stage we will distribute the costs among the constituencies who paid them, but at this stage we wish to ascertain the total costs of the intervention.

Ingredients can be divided into those that are purchased in reasonably competitive markets, and those that are obtained through other types of transactions. In general, the value of an ingredient for costing purposes is its market value. In the case of personnel, market value may be ascertained by determining what the costs would be for hiring a particular type of person. Such costs must include not only salary, but also fringe benefits and other employment costs that are paid by the employer. Many of the other inputs can also be costed by using their market prices. These include the costs of equipment, materials, utilities, and so on. Clearly the cost of leased facilities can also be ascertained in this way.

Although the market prices of some ingredients such as personnel can often be obtained from accounting data for educational enterprises, such data are not reliable sources for ascertaining overall program costs. The accounting systems that are used by schools were designed for ensuring consistent reporting to state agencies rather than for providing accurate and consistent cost data on educational interventions. For example, they omit completely or understate the cost of volunteers and other donated resources. Capital improvements are charged to such budgets and accounts during the year of their purchase, even when the improvements have a life of 20–30 years. Normal cost accounting practices would ascertain the *annual* costs of such improvements by spreading them over their useful lives through an appropriate method (Levin, 1983, pp. 67–71). Thus, data from accounting and budgetary reports must be used selectively and appropriately and cannot be relied upon for all ingredients.

There exist a variety of techniques for ascertaining the value of ingredients that are not purchased in competitive markets. For example, the method for ascertaining the value of volunteers and other contributed ingredients is to determine the market value of such resources if they had to be purchased. The value of facilities can be determined by estimating their lease value. The annual value of facilities and equipment can be estimated through a relatively simple approach that takes account of depreciation and interest foregone by the remaining capital investment. Details for these techniques are found in Levin (1983, chap. 4).

3. *Analysis of Costs.* Once each of the ingredients is costed, the costs can be added to obtain a total cost for the intervention. The next stage entails the use of these costs in an analytic framework (Levin, 1983, chap. 5). The two most important concerns for cost summary and analysis are the appropriate unit for expressing costs and who pays the costs.

Clearly, the question of the appropriate unit for expressing costs depends upon how effectiveness is measured and the nature of the decision. Usually, educational effectiveness is measured in terms of achievement gains per student or some other per-student measure. In that case, it is necessary to convert total costs to a per-student cost figure for comparing cost-effectiveness of alternative interventions. Cost-effectiveness ratios are usually based on the average effects and costs per student. However, it is possible to do an analysis on total project or program costs and effects. In other cases it may be the additional or marginal costs versus additional or marginal effectiveness that is the subject of scrutiny. For example, we may want to ascertain the number of additional students who will graduate from high school relative to the additional costs of alternative approaches for reducing dropouts.

Who pays the costs is a very different issue. The overall cost-effectiveness ratio may be irrelevant to a decisionmaker who pays only part of the costs for one intervention, but all of the costs for an alternative. For example, assume two alternatives of equal effect, where the cost of the first is twice as high as the cost of the second. In this case, we would say that the first is only half as cost-effective as the second. But, what if the first alternative attracts contributed facilities and volunteers so that most of its cost is "paid" by constituencies external to the decisionmaker, while all of the costs of the second alternative will be borne by the decisionmaker. In this case, the decisionmaker is likely to choose the first alternative, with its high subsidies, even though it is less cost-effective in the use of society's total resources.

For this reason, it is important to ascertain total costs of an intervention and to separate out those that are borne by the decisionmaker in considering different alternatives. However, we should bear in mind that since different decisionmaking units have different opportunities to obtain volunteers and contributed resources, we should not assume any particular cost subsidy to the decisionmaker. The basic estimate of costs that is used for all subsequent cost analyses is the overall cost of the intervention. Subsequent analyses can distribute the costs among those who will bear them to ascertain the implications of that distribution for decisions.

Use of Cost-Effectiveness Results

The combination of information on both the cost and effectiveness of each alternative provides useful information for decisions that must be made within a context of limited resources. In general, this information will be displayed in the form of cost-effectiveness ratios among the alternatives. Those with the lowest costs relative to effects should have the highest priority for decisions.

Table 2 shows hypothetical results for three alternatives. The usual evaluation of effectiveness—which does not consider costs—would rank alternative B as the best one because it shows much higher effectiveness than the other two alternatives. If costs alone were considered, alternative A would be ranked highest, since it has the lowest cost. When costs and effects are combined in cost-effectiveness ratios, alternative C is the most attractive alternative. It has a lower cost-effectiveness ratio than any of the other alternatives. Alternative A might be a close second choice.

However, these results should never be translated mechanically into decisions. There are four issues that must be raised before making a final decision. First, are the results representative for the level of scale of the intervention that is appropriate? Some interventions have very different average costs per student when used for 30 students than when used for 1,000. For example, a computer-assisted instruction (CAI) system that used an expensive mini-computer might be very costly for 30 students relative to stand-alone micro-computers. However, when providing instruction to several hundred students, the costs might be considerably lower than for micro-computers. *Thus, it is appropriate to make sure that the scale of use on which the cost-effectiveness ratios is based is pertinent to the decisionmaking context.*

TABLE 2

Hypothetical results for three alternative interventions

	A	B	C
Effects per student	15	25	20
Cost per student	$300	$625	$360
Cost/effect ratio	$ 20	$ 25	$ 18

A second concern is whether the results can be extrapolated for effects that are beyond those found in the interventions. For example, if $18 per student for alternative C generates an additional unit of effectiveness, would $180 generate 10 additional units of effectiveness? The answer to that question is that we do not know without further analysis of effectiveness under different intensities of resource use. In general, the results should be expected to hold only for small changes within the range of consideration rather than for global changes over a much larger range of service delivery. *Any generalization beyond the range covered by the study must be based on additional evidence rather than being extrapolated.*

A third concern in interpreting cost-effectiveness ratios is the magnitude of differences among alternatives. When the differences are small, for example, within a 10%–20% range, it is probably best to use other criteria to make the decision. The technique is not so precise that differences of this magnitude should be taken seriously. The precision of the estimates will depend upon the quality of data and the appropriateness of the evaluation methodology and implementation. When estimated cost-effectiveness results are not markedly different, given the attainable precision in results, other considerations such as the ease of implementation, acceptability by staff, and experience with each class of intervention should be used to make the decision (Levin, 1983, chap. 7). However, in other cases, differences in cost-effectiveness among alternatives may be as large as 4, 5, or 10 to 1. In these cases, there are likely to be important gains in using resources for the more cost-effective of the alternatives. *Decisions should not be based upon results that suggest only small differences in cost-effectiveness.*

A final concern is the appropriateness of the evaluation of effectiveness. A cost-effectiveness evaluation is only as good as the underlying evaluations of effectiveness and costs. Although the analyst can have a great deal of control over the estimation of costs, it is often necessary to depend upon an external study of effectiveness. But evaluation results are often confounded by differences among student populations that are not properly controlled for. In those cases, apparent differences in treatments or interventions are really caused by the otherwise unmeasured differences in student populations. Other types of confounding can also take place. It is important to make certain that the evaluation of effectiveness is studied closely to make sure that differences in effectiveness are likely to be associated with differences in interventions rather than confounding influences that were not properly controlled.

The other aspect of effectiveness that must be scrutinized is whether all appropriate outcomes have been taken into account. For example, different approaches to school organization such as reducing class size, cooperative learning, or peer tutoring may have effects other than those on student achievement. They may improve attendance of both teachers and students and create more positive learning attitudes. Effectiveness studies are often unidimensional and do not take these other potential outcomes into account. We will note below that a major area for improving cost-effectiveness studies is that of taking multi-dimensional outcomes into account, but it is first necessary for these to be reflected in evaluation studies of effectiveness. *Cost-effectiveness analyses should consider the quality and appropriateness of the effectiveness evaluations on which they are based.*

Some Applications in Education

The purpose of this section is to illustrate the wide diversity of actual and potential applications of cost-effectiveness in education. The examples are presented to illustrate the issues rather than to provide firm conclusions on cost-effectiveness for each topic. Results in any particular setting may be subject to differences in both costs and effectiveness. Four different areas will be selected: curriculum, teacher selection, general educational reforms, and educational technology. Each of these will be reviewed briefly with respect to the decision problem, the alternatives, the analysis, and findings.

Curriculum

Curriculum is an area that is very appropriate for cost-effectiveness analysis. In the quest for educational improvement, schools face numerous ways in which they can organize the pace, content, and method of instruction. Some are likely to be more effective than others, and there may also be substantial differences in resource requirements with respect to such ingredients as teacher time, materials, equipment, and so on. As with many other areas of education, the potential for using cost-effectiveness analysis has barely been tapped in the curriculum field.

An excellent cost-effectiveness study of fifth grade mathematics curriculum was carried out by Quinn, VanMondfrans, and Worthen (1984). Their evaluation compared two approaches to teaching mathematics. The GEMS math curriculum was developed by the school district that served as the site for their study. It was designed to provide highly individualized instruction with a special instructional method for teaching concepts. The alternative curriculum was one based on standard textbooks. It was found that students in the GEMS math classes had higher mathematics achievement gains than those in the traditional, textbook classes. Moreover, differences among students from different socioeconomic backgrounds were considerably reduced for GEMS, but not for the textbook classes.

Using the ingredients method, the authors found that the GEMS Math program required more resources than the Text Math. Total cost per pupil on an annual basis was $194 for Text Math and $288 for GEMS Math. Depending upon how student achievement was measured, GEMS Math was found to be from 60% to 300% more cost-effective—cost per point of achievement score—than Text Math. The analysis was also done separately for students from different socioeconomic status (SES) backgrounds. For students from low socioeconomic origins, the advantages in cost-effectiveness for GEMS were even more dramatic, but for high SES students that Text math had a cost-effectiveness advantage.

The study concluded that for low and medium SES students, the GEMS curriculum was considerably more cost-effective, while for high SES students the text math was slightly more cost-effective. The test results were comparable between the two curricula for the high SES students, but the GEMS math was more costly. If a single curriculum were to be used for all students, the GEMS math would be more cost-effective than text math. This study provides a sophisticated and well-documented illustration of a productive application of cost-effectiveness analysis to curriculum evaluation.

Teacher Selection

A very important policy challenge at the present time is teacher selection and retention (Bridges, 1986). During the 1980s and 1990s, substantial numbers of teachers who were hired in the 1950s and 1960s will be retiring. In addition, school enrollment growth in this period will require new teachers. Thus, a major decision issue is what types of teachers to select. Teachers with certain educational attributes, abilities, and personality traits might be more effective than teachers who lack these characteristics.

However, to the degree that the characteristics that make for effective teaching are also ones that make persons more productive in other occupations, it is likely that more effective teachers will require higher salaries to attract them into teaching. Accordingly, one must be concerned not only with ascertaining the types of teacher traits that are likely to be most productive in increasing student achievement. One must also know what it will cost to select teachers with these traits and which set of characteristics will maximize teacher effectiveness relative to the salary costs for obtaining those teachers.

Using the results of a study of the determinants of educational achievement of sixth graders by Eric Hanushek (1968) based upon data from the Equal Opportunity Survey (Coleman et al. 1966), Levin (1970) estimated the effects on student achievement of hiring teachers with more experience or greater verbal ability. Differences in teacher educational attainments were not used because they did not show a relation to student achievement. Separate estimates were made for black and for white students.

Estimates of costs for teachers with different characteristics were made from the Equal Opportunity Survey data for a major metropolitan area (Levin, 1968). Earnings functions were estimated for teachers on the basis of their qualifications and working conditions. The earnings coefficient for each variable represents the cost for an additional unit of the appropriate characteristic, for example, teacher verbal score and teacher experience.

Table 3 shows the cost-effectiveness results for improving student performances

TABLE 3

Relative costs of increasing student verbal achievement

	Approximate cost for increasing a student's verbal score by one point	
Strategy	Negro	White
Teacher's verbal score	$ 26	$ 26
Teacher experience	$128	$253

Note. From "A Cost-Effectiveness Analysis of Teacher Selection" by H. M. Levin, 1970, *Journal of Human Resources*, *5*(1), 32.

under alternative recruitment strategies. In terms of relative costs, a given test score gain for blacks could be attained at about one-fifth the cost by selecting teachers with higher verbal scores rather than by getting teachers with more experience. Selection on the basis of teacher verbal score was found to be 10 times as efficient from a cost-effectiveness perspective as selection on the basis of teacher experience for white students. The overall implications were that schools should be using their budgets to hire more verbally able teachers than ones with more experience.

Another observation from Table 3 was that teacher experience appeared to be twice as effective per dollar of expenditure for black students as for white ones. This suggests that for any given budget, the school could get higher student achievement by assigning more experienced teachers to the black students. These results should not necessarily be viewed as being pertinent today. They were obtained on the basis of 1965 data. However, it should be noted that other studies have found teacher test scores to be related to student achievement (Strauss & Sawyer, 1986), and teacher testing has become much more prominent in recent years as a basis for credentialing and selection. Thus, it would be useful to examine the cost-effectiveness of teacher test results and other selection criteria as a basis for teacher hiring.

General Educational Reforms

The 1980s have been a period of major educational reform with a dozen or more reports calling for improvement of the schools. Among the strategies that have been suggested for educational improvement are new technologies such as computers, longer school days and school years, higher standards, smaller classes, and better teacher selection. Unfortunately, the costs of implementing these and other instructional strategies that have been proposed would far exceed available resources for education, even with substantial increases in expenditures. Thus, cost-effectiveness analysis can be a major ally in choosing among the potential strategies for improving education.

Levin, Glass, and Meister (1984, 1987) carried out a cost-effectiveness analysis among four interventions that might be utilized in the reforms. These interventions included: a longer school day, computer-assisted instruction (CAI), cross-age tutoring, and reduced class size. The purpose of the analysis was to compare the cost-effectiveness of the four reforms for raising student achievement at the elementary level in reading and mathematics.

Longer School Day. The extension of the daily school session is a recommendation of almost all of the national reports on educational reform. The specific intervention considered was to lengthen the school day by one hour, half devoted to reading and half to mathematics. Data on effectiveness were taken from the Beginning Teacher Evaluation Study (BTES) as re-evaluated by Glass (1984).

Computer-Assisted Instruction. The specific approach to CAI that was chosen was the drill and practice program of Computer Curriculum Corporation as described in Ragosta, Holland, and Jamison (1982). The CCC method is one of the most widely used in the United States and provides students with 10 minutes of practice daily in each subject through the use of student terminals connected to a minicomputer. Effectiveness results were taken from the 4-year experiment of the CCC system that was undertaken by Educational Testing Service (Ragosta, Holland, & Jamison, 1982) and was re-evaluated by Glass (1984).

Cross-Age Tutoring. Cross-age tutoring utilizes older students to tutor younger ones, under the supervision of adults. Evidence suggests that this approach provides achievement gains for both the tutor and the tutored student (Ehly & Larsen, 1980, 12–17, 21–23). The specific tutoring program that was used for measuring effectiveness was one that had received national recognition from the Joint Dissemination Review Panel of the U.S. Department of Education (Independent School District of Boise City, 1983). Fifth and sixth grade students tutored students in the lower elementary grades. For students in fifth and sixth grades, adult tutors were provided. Thus, the study provided estimates of effectiveness for both peer and

TABLE 4

Estimated cost and effectiveness of four educational interventions in months of additional student gain per year of instruction

Intervention		Annual cost per student per subject (in dollars)	Effectiveness in estimated months of achievement gain	
			Mathematics	Reading
Longer school day		61	0.3	0.7
CAI		119	1.2	2.3
Cross-Age tutoring				
Peer component		212	9.7	4.8
Adult component		827	6.7	3.8
Reducing class size				
From	To			
35	30	45	0.6	0.3
30	25	63	0.7	0.4
25	20	94	0.9	0.5
35	20	201	2.2	1.1

TABLE 5

Estimated annual cost to obtain additional month of student achievement per year of instruction

Intervention		Mathematics	Reading
Longer school day		$203	$ 87
CAI		100	52
Cross-Age tutoring			
Peer component		22	44
Adult component		123	218
Reducing class size			
From	To		
35	30	75	150
30	25	90	158
25	20	104	188
35	20	91	183

adult tutors as reanalyzed in Glass (1984, 21–34).

Reduced class size. A smaller class size has been viewed traditionally as a principal way to improve instruction. Evidence of the effects of class size on mathematics and reading achievement is derived from a statistical analysis of 14 evaluations in which students were randomly assigned to classes of different sizes (Glass, 1984, 3–20; Glass & Smith, 1979). Specific estimates were made of effects on mathematics and reading achievement for reductions in class size from 35 to 30 students; 30 to 25 students; 25 to 20 students; and 35 to 20 students.

Costs were estimated by using the ingredients approach. Table 4 shows the annual cost per student for each subject as well as the estimated additional months of gains in achievement during the academic year as a result of the interventions. The least costly interventions were reductions in class size of five students and increasing the length of the school day. The most costly were the adult and peer tutoring. CAI was relatively inexpensive in comparison with cross-age tutoring or large decreases in class size.

Peer tutoring showed the largest effects: nearly a year of student gain in mathematics and about one-half year in reading. The longer school day and reducing class size by five student decrements showed the smallest effects. CAI was associated with gains in the middle of the range of results.

Table 5 indicates the cost-effectiveness of the different interventions by showing estimated annual costs to obtain an additional month of student achievement per year of instruction. To obtain an additional month of mathematics achievement, it would cost about $200 a year with a longer school day, but only about $22 a year with peer tutoring. In fact, a longer school day is less than half as cost-effective in raising mathematics achievement as is CAI or reducing class size, despite the prominence of calls for longer school days in the national reports on educational reform (Levin, 1984). The most cost-effective approach, peer tutoring, requires only one-ninth of the resources to obtain the same effect on mathematics achievement.

To obtain an additional month of reading achievement would cost about twice as much for increasing the school day as using peer tutoring, with CAI almost as efficient as the latter. Reducing class size appears to be particularly inefficient with respect to raising reading achievement. For both mathematics and reading achievement, adult tutoring is relatively costly. Although its effectiveness is very high, its cost is also high. These results provide useful information in comparing instructional strategies to raise achievement.

Computer-Assisted Instruction

The previous study compared a popular version of CAI with a number of other interventions. On the basis of a cost-effectiveness criterion, it was found that CAI was ranked below peer tutoring, but considerably above adult tutoring, longer school days, and smaller class sizes.

However, the CAI approach that was considered is not the only one. There are many different approaches to CAI and different implementations as well. Alternatives include differences in curriculum, software, hardware configurations, and personnel use. A recent study explored the cost-effectiveness of different CAI approaches to increasing elementary school reading and mathematics achievement (Levin, Leitner, & Meister, 1986).

The comparisons among different approaches to CAI were somewhat disappointing in that it was difficult to find replications of approaches. Other than those of CCC—where there are many replications—other approaches were single efforts that were implemented and evaluated by their progenitors. The results of such efforts should be replicated beyond the initial version before consideration by decisionmakers for adoption.

However, among a comparison of four CCC sites and four other approaches to CAI for improving reading and mathematics achievement at the elementary level, a rather unexpected finding emerged that has important policy consequences. Only one of the eight sites was fully utilizing its CAI capacity. Full utilization was judged by the standard that was found to be feasible in the 4-year experiment that was evaluated by the Educational Testing Service (Levin & Woo, 1981). On the average, the per student cost at the eight sites for CAI was almost $300 a year. However, if the CAI had been fully utilized, the cost would have fallen to a little more than $180 per student per year.

This means that even with the same interventions, cost-effectiveness could have been increased by 80%, on average, by more fully utilizing CAI capacity. Such a finding will hardly come as a surprise to those who have observed the large number of unused computers and associated resources that are commonly found in schools. But, it does suggest that there is a potentially powerful strategy for improving the cost-effectiveness of existing programs as opposed to seeking new ones.

Some Issues in Cost-Effectiveness Analysis

Several examples were provided to show both the policy relevance of cost-effectiveness analysis in education as well as some illustrative findings. But at least five issues should be raised about the potential of cost-effectiveness analysis for educational policy: its cost; its value; its conservatism; its treatment of multiple outcomes; and its integration with a major new branch of evaluation, meta-analysis.

Cost of Cost-Effectiveness Analysis

Obviously, cost-effectiveness analysis requires resources to carry it out. One concern is the magnitude of the additional costs. Does the addition of cost analysis to the standard effectiveness design add significantly to the cost of evaluation? Our experience suggests that it does not. At least 80%–90% of the cost of an evaluation will be associated with the study of effectiveness rather than cost. This is especially true if cost considerations are built in to the evaluation design.

The 4-year CAI experiment cost between $1.5 and $2 million (Ragosta, Holland, & Jamison, 1982). But the cost analysis component of the intervention was carried out for less than $10,000. Obviously, the relative costs of the effectiveness evaluation and the cost evaluation will depend on the specific circumstances of the overall evaluation. However, the cost component will generally be a very small portion of the total. Accordingly, the main issue is whether the additional information provided by a cost-effectiveness comparison justifies the relatively small marginal cost of adding the cost dimension to an evaluation of effectiveness.

Value of Cost-Effectiveness Analysis

Does it pay to do a cost-effectiveness analysis? Clearly, this depends upon what is at stake. If the decision framework is relatively inconsequential with respect to resource allocation, it is probably not advisable to do such an analysis. In contrast, if the resource consequences of a decision are substantial, a cost-effectiveness analysis is probably justified. For example, in a typical school district with an enrollment of 10,000 students and a per student expenditure of $3,500 a year, the total budget will be about $35,000,000. If even 1% of that amount could be saved or reallocated to new activities through cost-effectiveness studies, that would amount to an annual saving of $350,000 a year. This annual saving is the equivalent of investing $3.5 million at a 10% rate of interest.

Now consider the same logic at the state or national level. At the national level, for example, an estimated $170 billion was spent on elementary and secondary schools and another $112 billion on higher education in 1986–87 for a total of over $282 billion (U.S. Department of Education 1987, p. 25). A 1% savings would amount to about $2.8 billion dollars, or the equivalent of the return on an investment of $28 billion at a 10% rate of interest. Moreover, the fact that the educational sector has not been scrutinized systematically for efficiency suggests a potential for increased productivity of considerably greater than 1%—perhaps as high as 10%.

Yet, it is doubtful that as much as $1 million a year is presently being spent on cost-effectiveness studies in education by federal, state, and local agencies together. Thus, even an increase to $5 or $10 million would represent a major improvement, although considerably short of the probable optimum investment, given what is at stake.

Conservatism of Cost-Effectiveness Analysis

Any form of evaluation which requires a summation of results from *existing* alternatives will tend to be conservative. That is, it must necessarily limit consideration of alternatives to those interventions that exist or that can be evaluated. It cannot provide evaluative information on those that are not yet available. While this means that all existing alternatives can be considered, future developments can not be easily accommodated in the analysis. Cost-effectiveness analysis and evaluation tools more generally do not lend themselves to areas where change may occur rapidly. Thus, an evaluation of existing CAI may say little about approaches to CAI that are likely to emerge in future years from the harnessing of artificial intelligence to learning or the development of new, computer courseware.

This is less of a concern than critics of evaluation have made it out to be. First, it is never possible to evaluate that which does not exist. This is not as much a criticism of cost-effectiveness analysis as it is of a world in which one cannot be prescient. Fortunately, the history of education suggests that change is not so rapid and discontinuous that evaluation of today's alternatives will miss the breakthroughs of tomorrow. Although technological change may be rapid, its applications to and impact on education have been much more gradual than predictions have suggested. This is evidenced by the claims that have been made for every new instructional technology (motion pictures, radio, television, videocassettes, videodisks, and computers) which heralded them as having revolutionary implications for education. In every case their educational impacts fell short of that promise, but not because they were defeated by evaluation or cost-effectiveness analysis. They simply did not evolve in the ways that were predicted (Cuban, 1986; Levin & Meister, 1985).

Second, as soon as new educational techniques are developed and applied, it is possible to consider them. There need be little time lag between the establishment of new alternatives and their consideration in a cost-effectiveness analysis. Although cost-effectiveness analysis cannot be readily applied to choosing among research and development projects for new instructional approaches, it can certainly be applied to evaluate those interventions that emerge from that process.

Multiple Objectives

Although most evaluations focus on a single outcome, the evaluative context is often characterized by alternatives in which many outcomes should be considered. Educational interventions are often designed to raise student achievement and improve student motivation. In some cases, such as the reduction of class size or extension of the school day, the interventions might be expected to influence achievement in a number of subject areas. In these cases, the assumption of a single objective and measure of outcome is inappropriate. Cost-benefit studies address this problem by converting all outcomes into dollar values. But, as we noted in the introduction, not all of the outcomes of many types of interventions can be converted into monetary values for a cost-benefit analysis.

The challenge is how to convert multiple outcomes into a single index which can be compared among alternative interventions. This challenge is one for the field of evaluation generally, not just cost-effectiveness evaluations. The most common response to the challenge is to convert the various outcomes to ratings on utility scales which share common measurement properties, although other techniques have also been suggested (Nagel, 1983). The separate utility scales can be aggregated into a total utility index through a variety of procedures (Bell, Keeney, & Raiffa, 1977; Keeney & Raiffa, 1976). The utility scores can be combined with costs for each alternative intervention to derive cost-utility ratios. In the final section of this paper, a plan is suggested for integrating this approach into educational decisionmaking.

Cost-Effectiveness and Meta-Analysis

Clearly, one of the major developments in educational evaluation over the last decade has been that of meta-analysis. Meta-analysis represents the attempt to use various analytic techniques to generalize about a specific phenomenon from a wide variety of independent and disparate studies (Glass, McGaw, & Smith, 1981). Typically, a meta-analysis will collect all of the available studies on a particular phenomenon and estimate an effect size. The effect size will be a standard score such as a z-score on the criterion variable. Meta-analyses have been carried out on a wide variety of topics including the effects of reducing class size (Glass, Cohen, Smith, & Filby, 1982), computer-assisted instruction (Bangert-Drowns, Kulik, & Kulik, 1985; Kulik, Kulik, & Bangert-Drowns, 1985), tutoring (Cohen, Kulik, & Kulik, 1982), and a large range of other educational interventions (Walberg, 1984).

A representative meta-analysis study will provide two types of outcomes. First, it will provide an average "effect size" for a particular type of intervention. Usually this will be expressed in terms of an average effect of an intervention among a set of studies in terms of some standard score relative to what would be expected in the absence of the intervention. The second type of outcome represents an attempt to assess the sources of differences in outcomes among a set of studies. For this purpose, studies are coded according to their features to see if such classifications are associated with different effect sizes.

Although meta-analysis has been widely accepted in the evaluation literature, it is also the subject of criticism. Among these concerns are the issues of inclusion of studies that are based upon poor design and procedures, issues of how to "average" results among different studies, coding and classification aspects of the studies, and appropriate interpretation of results (Hedges & Olkin, 1985; Slavin, 1984, 1986). But a very different issue is whether the result of meta-analytic summaries—even ones of high quality—should be used for cost-effectiveness analysis.

The attraction of using meta-analysis results for cost-effectiveness is straightforward. Assume that good meta-analyses exist to summarize the effects of two alternative interventions. Why not obtain the costs of the two interventions and combine them with their respective effect sizes to compare the cost-effectiveness of the two alternatives? The basic problem is that meta-analysis deals with averages rather than actual programs, while decisionmakers are concerned with choosing among concrete alternatives. This can be seen more clearly if we look at a specific meta-analytic result.

Cohen, Kulik, and Kulik (1982) did a meta-analysis of peer tutoring in which they obtained an effect size of .4 in student achievement. Accordingly, one might estimate the costs of the program that produced this effect size and compare the costs and effects with other alternatives such as computer-assisted instruction, which has shown an average effect size of about .47 according to Kulik, Kulik, and Bangert-Drowns (1985). The problem is that there is no "program" to cost out, since the effect size is based upon an average of many programs: those that failed and those that succeeded; those that provide training for tutors and those that do not; those that monitor tutors and provide instructional materials and those that simply tell some students to tutor others.

From a cost-effectiveness perspective for decisionmaking, this type of result is not helpful for several reasons. First, there is no specific program intervention that is associated with the effect size of .4. That effect size is not based on a program at all, but it is an average based on an amalgam of many different programs. Decisionmakers are concerned with what results they will be able to obtain with a specific intervention, not with the average effects from many different versions of that intervention, some of which were poorly designed and executed and others that were well-designed and executed. Thus, the average effect size makes no sense to the decisionmaker who must consider the impacts of concrete interventions that can be clearly specified and considered for implementation. The effect size associated with meta-analysis is an average associated with a large number of heterogeneous interventions that are grouped under a single topic rather than replications of the same basic intervention.

Second, it is not possible to provide a cost for a "program" that provides the average effect size. The costing of programs requires that we have a specific set of processes and resource ingredients. The fact that the effect size is not based upon a single program, but a mixture of many different programs means that there is no conceptual or practical way to identify costs. The effect size simply does not refer to a program alternative with specific resource requirements.

Third, decisionmakers are primarily concerned with a specification of alternative interventions that have the following properties. They can be readily implemented and do not depend upon special circumstances beyond the capabilities of the decision unit. Unfortunately, many of the studies that are reflected in the meta-analysis literature violate this standard because they were initiated and implemented by researchers rather than practitioners. Indeed, many were doctoral dissertations or "laboratory" studies carried out by university researchers in which the scope conditions for the intervention are not easily generalizable to non-experimental settings. For example, one can hardly assume that effect sizes for an experimental intervention of a few weeks are extrapolatable to a year of instruction. Kulik, Bangert, and Williams (1983) found an effect size of .56 for computer-assisted instruction in projects lasting less than 4 weeks, but an average effect size of .20 for projects lasting more than 8 weeks. Such results are not terribly useful to the decisionmaker unless it can be shown that the interventions can be readily implemented into nonexperimental situations and replicated beyond the original site.

It is also important that the results be based upon evaluations of an acceptable design and quality. Again, many of the studies captured by the meta-analyses are extremely poor in quality. One evaluation found that three-quarters of the evaluations that were summarized in a specific meta-analysis had serious flaws (Clark, 1985).

Finally, the types of interventions that should be considered for policy must not only meet the criteria of replicability with information from sound evaluations. They should also be those that show success. That is, decisionmakers are not interested in considering an *average* of good and poor interventions. Rather, they want information on how to replicate those specific alternatives that have shown good results. The conflicts between what meta-analysis has to offer at the present time and what decisionmakers need is reflected in a debate comparing the cost-effectiveness of peer tutoring with computer-assisted instruction between Levin and Meister (1986); Niemiec, Blackwell, and Walberg (1986); and Levin, Glass, and Meister (1986).

Probably the most general conclusion that can be drawn on what we have learned about the potential link between cost-effectiveness analysis and meta-analysis is that it is tenuous at best. The two techniques were developed to address very different purposes. Meta-analysis was developed to summarize the results of a large number of disparate studies on a single general subject. In doing so, it was never designed to identify exemplary practices that have been found to be replicable and that meet the other standards that are appropriate for policy consideration. It throws together the good and the bad in terms of both successful and unsuccessful interventions and low quality and high quality evaluations. Its very value derives from its claims to summarize very diverse types of approaches and evaluations within a general rubric.

Cost-effectiveness analysis was designed to assist decisionmakers in choosing among successful practices on the basis of which ones would maximize the impact of available resources. In this vein the quality of existing evaluations and replication experience are crucial; those practices or interventions that have failed are ruled out. An average effect for all evaluations of a particular class of interventions is not very useful for this purpose, since the average is not embodied in any specific program. And the lack of a specific program means that what is embedded in the average effect size cannot be costed or implemented.

Future Tasks

In this paper I have asserted that cost-effectiveness analysis has an important role to play in informing decisions within the context of limited resources. This assertion is in sharp contrast with the reality that little cost-effectiveness analysis is carried out or used in education. A 1983 survey of departments of education among the 50 state governments found that they had produced a total of only 18 cost-effectiveness studies over the previous 5 years or an average of about .35 studies over the 5-year period for each state (Smith & Smith, 1985). When asked what impediments existed to doing such studies, the three that were mentioned most frequently were: "We are seldom asked to do cost studies"; "It is difficult to relate cost data to educational outcomes"; and "We have few guidebooks, texts, or examples to follow in conducting cost studies."

It is probable that staff are seldom asked to do such studies because the decisionmakers are not familiar with cost-effectiveness analysis and professional staff are not competent to suggest such studies or carry them out. It is clear that a major training effort will be required to overcome these obstacles. This training initiative must extend to both university training programs and in-service endeavors. Graduate programs for educational administrators should require a solid unit in understanding and use of cost-effectiveness analysis for decisionmaking. A major stride toward this goal can be effected by requiring such training for the administrative credential at the elementary and secondary level and possibly testing credential applicants for their competencies in this area. Educational evaluators should obtain background in the design, execution, and interpretation of cost-effectiveness analysis. In many cases it will only be necessary to work closely with an economics department or department of public policy to obtain the instruction and expertise. In addition, professional staff in evaluation and policy positions should receive in-service training through intensive workshops and courses in cost-effectiveness analysis. These should emphasize not only the concepts and illustrations from the literature, but they should require "hands-on" experience through assignments of case studies.

Building Cost Analysis into Evaluation Design

Existing studies of cost-effectiveness analysis in education are based on adding a cost analysis to a completed analysis of effectiveness. That is, the cost analysis is appended to an effectiveness evaluation that was designed and executed independently. The reason for this disjuncture is that studies of educational effectiveness normally are not designed to consider costs. Therefore, a specialist in cost-effectiveness analysis must graft the cost analysis to an earlier effectiveness study. This is a very inefficient process, since the identification of ingredients must be done retrospectively by reviewing documents and doing interviews.

It would be more efficient to build the cost analysis into the overall study design. In this way, ingredients could be identified during the evaluation process rather than in a separate exercise retrospectively. The integration of cost analysis into evaluation designs would reduce the cost of doing such an analysis and increase its accuracy by avoiding data gaps and errors of retrospection. This development will require training of evaluators in cost-effectiveness analysis as described in the previous section as well as the establishment of guidelines for integrating cost analysis into evaluation designs. In this way a cost-effectiveness component could be a standard product of evaluation designs.

Future Research

A major goal for the future would be to develop methods for educational decisionmakers to do cost-effectiveness analysis as part of the decision process. In suggesting this, we must be cognizant of three obstacles. First, decisionmakers do not have proficiencies in the cost-effectiveness area; second, many decision problems are characterized by multiple objectives rather than a single one; third, most decisionmaking is done under conditions of highly imperfect information.

The first two obstacles have been discussed earlier in this paper. The third was not discussed because we assumed the existence of good evaluation studies that would provide information on the effectiveness of alternatives. Not only is such information not always available, but it is often very costly to develop in terms of time and resources. Accordingly, it is important to consider ways in which cost-effectiveness types of analysis can be done with partial or imperfect information.

One solution is to build on the ubiquitous availability of micro-computers by designing approaches that will be accessible to the decisionmaker and that can be addressed by computer. This could be done by preparing interactive software that will assist the educational administrator or decisionmaker to define and address problems, place them in a cost-effectiveness or cost-utility framework, and solve them with the help of powerful data-base management and financial spreadsheet packages. The computer could assist the decisionmaker in structuring a problem by giving guidance and instructions, asking questions, providing prompts and examples, and using the decisionmaker's response to determine the parameters of the problem as well as alternatives for solving it.

Programs can be prepared to assist the user in rating alternatives, evaluating the likelihood of certain events occurring, and convening interactive group process such as the Delphi technique to make evaluations of outcomes. Such interactive applications can serve simultaneously to train the decisionmaker to solve problems of a general type as well as to construct solutions for a specific problem.

Multiple outcomes could be treated through the application of a utility rating approach (Edwards & Newman, 1982; Keeney & Raiffa, 1976), and uncertainty can be handled through the use of Bayesian techniques (Pollard, 1986). All of these features can be implemented through the design and preparation of interactive software routines that provide instructions, exercises, operational algorithms for calculations, and visual displays to assist the decisionmaker in contemplating and considering judgments.

In short, what is needed is the development of a user friendly tool that is computer-based and that will handle the technical side of the calculations while working with the decisionmaker to elicit information and judgments. More details on the steps required to develop this approach can be found in Levin (1985). It is likely that the development of such a computer-based approach would go far in introducing educational policy analysts and decisionmakers to the resource implications of alternative ways of improving the effectiveness of American education.

References

Alkin, M. C., & Solmon, L. C. (Eds.). (1983). *The costs of evaluation.* Beverly Hills, CA: Sage.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1985, Summer). Effectiveness of computer-based education in secondary schools. *Journal of Computer-Based Education, 12*(3), 59–68.

Becker, G. S. (1964). *Human capital.* New York: Columbia University Press.

Bell, D. E., Keeney, R. L., & Raiffa, H. (1977). *Conflicting objectives in decision.* New York: John Wiley.

Bridges, E. (1986). *The incompetent teacher.* Philadelphia: The Falmer Press.

Chambers, J. (1981). *Cost and price level adjustments to state aid for education: A theoretical and empirical review* (IFG Project Report No. 81-A3). Stanford, CA: Stanford University, Institute for Research on Educational Finance and Governance.

Clark, R. E. (1985, Winter). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. *Educational Technology and Communication Journal, 33*(4), 249–262.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982, Summer). Educational outcome of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19,* 237–248.

Coleman, J. S., et al. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Office of Education, U.S. Government Printing Office.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation.* Chicago: Rand McNally.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920.* New York: Teachers College Press.

Edwards, W., & Newman, J. R. (1982). *Multiattribute evaluation* (Quantitative Applications in the Social Sciences No. 26). Beverly Hills, CA: Sage.

Ehly, S. W., & Larsen, S. C. (1980). *Peer tutoring for individualized instruction.* Boston: Allyn and Bacon.

Glass, G. V. (1984). *The effectiveness of four educational interventions* (Project Report No. 84-A19). Stanford, CA: Stanford University, Institute for Research on Educational Finance and Governance.

Glass, G. V, Cohen, L., Smith, M. L., & Filby, N. (1982). *School class size.* Beverly Hills, CA: Sage.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Glass, G. V, & Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis, 1,* 2–26.

Gold, S. D. (1986, November/December). Spending spree. *State Legislatures,* 23–25.

Gramlich, E. (1981). *Benefit-cost analysis of government programs.* Englewood Cliffs, NJ: Prentice Hall.

Hanushek, E. (1968). *The education of Negroes and Whites.* Unpublished doctoral dissertation, Department of Economics, Massachusetts Institute of Technology.

Hartman, W. T. (1981, July/August). Estimating the costs of educating handicapped children: A resource-cost model approach—summary report. *Educational Evaluation and Policy Analysis, 3*(4), 33–48.

Hawley, D. E., Fletcher, J. D., & Piele, P. K. (1986). *Costs, effects, and utility of microcomputer-assisted instruction.* Eugene: University of Oregon, Center for Advanced Technology in Education.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hu, T., Lee, M. L., & Stromsdorfer, E. W. (1971, Winter). Economic returns to vocational and comprehensive high school graduates. *Journal of Human Resources,* 25–50.

Independent School District of Boise City. (1983, July). *1982–83 evaluation report* (PROJECT INSTRUCT, Chapter 1 Program).

Jamison, D. T., Klees, S. J., & Wells, S. J. (1978).

The costs of educational media. Beverly Hills, CA: Sage.

Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs.* New York: John Wiley.

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. C., & Bangert-Drowns, R. L. (1985). Effectiveness of computer-based education in elementary schools. *Computers in Human Behavior, 1,* 59–74.

Levin, H. M. (1968). *Recruiting teachers for large-city schools* (mimeo). Washington, DC: The Brookings Institution.

Levin, H. M. (1970). A cost-effectiveness analysis of teacher selection. *Journal of Human Resources, 5*(1), 24–33.

Levin, H. M. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag & E. Struening (Eds.), *Handbook of evaluation research* (Vol. 2, pp. 89–122). Beverly Hills, CA: Sage.

Levin, H. M. (1980). Teacher certification and the economics of information. *Education Evaluation and Policy Analysis, 2*(4), 5–18.

Levin, H. M. (1983). *Cost-effectiveness: A primer.* Beverly Hills, CA: Sage.

Levin, H. M. (1984, Summer). About time for educational reform. *Educational Evaluation and Policy Analysis, 6*(2), 151–163.

Levin, H. M. (1985, July). *Cost-effectiveness analysis for educational decision makers.* Unpublished paper prepared for the Economic Development Institute, The World Bank, Washington, DC.

Levin, H. M., Glass, G. V, & Meister, G. R. (1984, May). *A cost-effectiveness analysis of four educational interventions* (Project Report No. 84-A11). Stanford, CA: Stanford University, Institute for Research on Educational Finance and Governance.

Levin, H. M., Glass, G. V.,& Meister, G. R. (1986, September). The political arithmetic of cost-effectiveness analysis. *Phi Delta Kappan, 68*(1), 69–72.

Levin, H. M., Glass, G. V, & Meister, G. R. (1987, February). A cost-effectiveness analysis of computer-assisted instruction. *Evaluation Review, 11*(1), 50–72.

Levin, H. M., Leitner, D., & Meister, G. R. (1986). *Cost-effectiveness of alternative approaches to computer-assisted instruction* (87-CERAS-1). Stanford, CA: Stanford University, Center for Educational Research at Stanford.

Levin, H. M., & Meister, G. R. (1985). *Educational technology and computers: Promises, promises, always promises* (Project Report No. 85-A13). Stanford, CA: Stanford University, Institute for Research on Educational Finance and Governance.

Levin, H. M., & Meister, G. R. (1986, June). Is CAI cost-effective? *Phi Delta Kappan, 67*(10), 745-749.

Levin, H. M., & Woo, L. (1981, Winter). An evaluation of the costs of computer-assisted instruction. *Economics of Education Review, 1*(1), 1-26.

Nagel, S. (1983). Nonmonetary variables in benefit-cost analysis. *Evaluation Review, 7,* 37–64.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Department of Education.

Niemiec, R. P., Blackwell, M. C., & Walberg, H. J. (1986, June). CAI can be doubly effective. *Phi Delta Kappan, 67*(10), 750–751.

Odden, A. (1984, January). Financing educational reform. *Phi Delta Kappan, 65,* 311–318.

Pollard, W. E. (1986). *Bayesian statistics for evaluation research.* Beverly Hills, CA: Sage.

Quinn, B., VanMondfrans, A., & Worthen, B. R. (1984, Spring). Cost-effectiveness of two math programs as moderated by pupil SES. *Educational Evaluation and Policy Analysis, 6*(1), 39–52.

Ragosta, M., Holland, P. W., & Jamison, D. T. (1982, April). *Computer-assisted instruction and compensatory education: The ETS/LAUSD study* (Final Report, Project Report No. 19). Princeton, NJ: Educational Testing Service.

Rossi, P. H., & Freeman, H. E. (1985). *Evaluation: A systematic approach* (3rd ed.). Beverly Hills, CA: Sage.

Slavin, R. E. (1984, October). Meta-analysis in education: How has it been used? *Educational Researcher, 13*(8), 6–15, 24–27.

Slavin, R. E. (1986, November). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*(9), 5–11.

Smith, N. L., & Smith, J. K. (1985, June). State-level evaluation uses of cost analysis: A national descriptive survey. *New Directions for Program Evaluation, 26,* 83–97.

Strauss, R. P., & Sawyer, E. A. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review, 5*(1), 41–48.

U.S. Department of Education, Office of Educational Research and Improvement. *Digest of education statistics 1987.* Washington, DC: U.S. Government Printing Office.

Walberg, H. J. (1984, May). Improving the productivity of America's schools. *Educational Leadership, 41*(8), 19–27.
Weisbrod, B. A. (1965). Preventing high school dropouts. In R. Dorfman (Ed.), *Measuring benefits of government investments* (pp. 117–148). Washington, DC: The Brookings Institution.

Author

HENRY M. LEVIN, Professor of Education and Economics, CERAS Bldg., Stanford University, Stanford, CA 94305. *Specialization:* economics of education.