average class size, average number of days per school year, and average number of class hours per school day. We assumed that length of school day and length of school year are less flexible than average class size, and examined only the effect on class size of introducing CAI. The other two variables could, however, be introduced into the analysis in a straightforward way.

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

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

Solving this equation for A we obtain

$$A = CS^2/T - CS)$$

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

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

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

and

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

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

A number of interesting points emerge from the table. First, even if C = \$75, the student-to-teacher ratio only goes from 26 to 31.6 in order to provide CAI. If the Coleman Report is correct in concluding that student performance is insensitive to student-to-teacher ratio, this would seem to be an attractive reallocation to the extent that it can be made politically feasible.[4] Second, from the values for $\partial A/\partial C$ we see that a \$10 increase in C would require about a .8 increase in A if C is \$75. Third, from the value of $\partial A/\partial S$ we see that an increase of 1 in S causes

TABLE IV-4 Increment in Class Size Required to Finance CAI

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

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

an increase of .286 in A if $C = \$50$ but an increase of .477 if $C = \$75$. Finally, the last row in the table shows that a $1,000 annual increase in teacher salary would decrease A by about .36 if C is $50; it decreases A by almost twice that amount if C is $75. In general, the partial derivatives in the table seem quite sensitive to C.

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

NOTES

1. Sen (1972) criticized the restrictiveness of the additive functional form that Atkinson assumed for determining W. Sen provided a definition of inequality similar to Atkinson's based on a more general functional form. However, the practical usability of the additive form remains a strong argument in its favor.
2. Ball and Jamison (1973) presented updated and more detailed cost estimates for all aspects of a CAI system designed to serve rural populations; their cost estimates differ only a little from the more preliminary ones used here.
3. A further, and very important, advantage of using satellites is that it eliminates the necessity of working with poorly equipped rural telephone services. IMSSS has experienced many delays and unexpected costs as a result of working with such services in Kentucky and elsewhere.
4. Jamison, Suppes, and Wells (1974) reviewed additional evidence that indicates student performance to be insensitive to the student-to-teacher ratio; their review summarized the literature on the effectiveness of various educational technologies as well as various forms of conventional instruction.

REFERENCES

Atkinson, A. B. "On the Measurement of Inequality." *Journal of Economic Theory* 2 (1970): 221–224.

Atkinson, R. C. "Computerized Instruction and the Learning Process." *American Psychologist* 23 (1968): 225–239.

Atkinson, R. C., and Fletcher, J. D. "Teaching Children to Read with a Computer." *Reading Teacher* 25 (1972): 319–327.

Ball, J., and Jamison, D. "Computer-Assisted Instruction for Dispersed Populations: System Cost Models." *Instructional Science* 1 (1973): 469–501.

Boneau, C. A. "The Effects of Violations of Assumptions Underlying the *t*-Test." *Psychological Bulletin* 57 (1960): 49–64.

Bowles, S., and Levin, H. "The Determinants of Scholastic Achievement: An Appraisal of Some Recent Evidence." *Journal of Human Resources* 3 (1968): 3–24.

Chall, J. *Learning to Read: The Great Debate.* New York: McGraw-Hill, 1967.

Coleman, J. *The Evaluation of Equality of Educational Opportunity.* P-3911. Santa Monica, Calif.: Rand Corporation, 1968.

Coleman, J. S.; Campbell, E. Q.; Hobson, C. J.; McPartland, J.; Mood, A. M.; Weinfeld, F. D.; and York, R. L. *Equality of Educational Opportunity.* Washington, D.C.: U.S. Office of Education, 1966.

Elashoff, R. M. "Effects of Errors in Statistical Assumptions." *International Encyclopedia of the Social Sciences.* New York: Macmillan, 1968, Vol. 5, pp. 132–142.

Fletcher, J. D., and Atkinson, R. C. "An Evaluation of the Stanford CAI Program in Initial Reading (Grades K through 3)." *Journal of Educational Psychology* 63 (1972): 597–602.

Jamison, D.; Ball, J.; and Potter, J. "Communication Economics of Computer-Assisted Instruction for Rural Areas: Satellite vs. Commercial Phone Systems." Presented at the 22nd International Astronautical Congress of the International Astronautical Federation, Brussels, September, 1971.

Jamison, D.; Suppes, P.; and Butler, C. "Estimated Costs of Computer-Assisted Instruction for Compensatory Education in Urban Areas." *Educational Technology* 10 (1970): 49–57.

Jamison, D.; Suppes, P.; and Wells, S. "The Effectiveness of Alternative Instructional Media: A Survey." *Review of Educational Research* 44 (1974): 1–67.

Jensen, A. "How Much Can We Boost I.Q. and Scholastic Achievement?" *Harvard Educational Review* 39 (1969):1–123.

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

Kiesling, H. *Input and Output in California Compensatory Education Projects.* R-781-CC/RL. Santa Monica, Calif.: Rand Corporation, 1971.

Lekan, H. A., ed. *Index to Computer-Assisted Instruction.* 3rd ed. New York: Harcourt Brace Jovanovich, 1971.

Levien, R. *The Emerging Technology: Instructional Uses of the Computer in Higher Education.* New York: McGraw-Hill, 1972.

Levine, D. B., and Singer, N. M. "The Mathematical Relation between the Income Density Function and the Measurement of Income Inequality." *Econometrica* 38 (1970): 324–330.

Lorton, P., and Slimick, J. "Computer-Based Instruction in Computer Programming—a Symbol Manipulation–List Processing Approach." *Proceedings of the American Federation of Information Processing Societies 1969 Fall Joint Computer Conference* 33 (1969): 535–543.

Michelson, S. "The Data: Equal Protection and School Resources." *Inequality in Education* 2 (1970): 4–16.

———. "The Political Economy of Public School Inequalities." In M. Carnoy, ed., *The Political Economy of Education: New Approaches to Old Problems*, forthcoming.

Mood, A., ed. *Do teachers make a difference?* Washington, D.C.: GPO, 1970.

Newbery, D. "A Theorem on the Measurement of Inequality." *Journal of Economic Theory* 2 (1970): 264–266.

Pfanzagl, J. "A General Theory of Measurement—Applications to Utility." *Naval Research Logistics Quarterly* 6 (1959): 283–294. Reprinted in R. Luce, R. Bush, and E. Galanter, eds., *Readings in Mathematical Psychology*, Vol. II. New York: John Wiley, 1965.

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

Sen, A. "On Economic Inequality." The Radcliffe Lectures at Warwick University, draft, 1972.

Stifle, J. *The PLATO IV Architecture*. Technical Report No. X-20, May, 1972. University of Illinois-Urbana, Computer-Based Education Research Laboratory.

Suppes, P., and Morningstar, M. "Computer-Assisted Instruction." *Science* 166 (1969): 343–350.

———. "Technological Innovations: Computer-Assisted Instruction and Compensatory Education." In F. Korten, S. Cook, and J. Lacey, eds., *Psychology and the Problem of Society*. Washington, D.C.: American Psychological Association, 1970.

Suppes, P.; Jerman, M.; and Brian, D. *Computer-Assisted Instruction: The 1965–66 Stanford Arithmetic Program*. New York: Academic Press, 1968.

Wells, S.; Whelchel, B.; and Jamison, D. "The Impact of Varying Levels of Computer-Assisted Instruction on the Academic Performance of Disadvantaged Students." Technical Report, Stanford University, Institute for Mathematical Studies in the Social Sciences, 1973.

Zinn, K. L. *An Evaluative Review of Uses of Computers in Instruction*. Final Report, U.S.O.E. Contract No.: OEC-5-9-32-509-0032, December, 1970, University of Michigan, Center for Research on Learning and Teaching.

5 COMMENTS

Allen C. Kelley

Duke University

For more than a decade Patrick Suppes and his associates at Stanford have been pioneering the application of computer technology to elementary and high school education. Their imaginative and ambitious projects have included basic research in learning theory, the formulation and evaluation of programmed-instruction techniques, and the development of computer and communications systems for disseminating programmed materials. Thousands of students across the United States have been "plugged into" the Stanford CAI facility; moreover, recent advances in satellite communications technology have greatly expanded the geographic dispersion of their instructional programs. Based on the extensive data resulting from these applications, it has been demonstrated that CAI and programmed instruction are effective, under selected circumstances, in increasing student achievement; furthermore, CAI is technologically feasible in a wide range of educational settings. However, resources are scarce. Technological feasibility is only one of the conditions necessary for the utilization of CAI; this technique must also be shown to be economically efficient. The present paper by Jamison, Fletcher, Suppes, and Atkinson represents one of several recent studies which examines the economic viability of CAI.

The paper before us presents a benefit-cost analysis of CAI in three subject-matter areas—arithmetic, reading, and computer programming—with particular attention to the role of CAI in compensatory education. Several notable conclusions are reached:

1. The CAI programs have both a statistically significant and quantitatively large impact on increases in student achievement.
2. The program's benefits appear greatest at relatively low achievement levels.
3. CAI tends to narrow the inequality in the distribution of educational outputs.
4. CAI is economically efficient in selected settings, e.g., compensatory education in rural schools.

Given the high failure rate of the compensatory-education programs in the United States, the CAI system reported in the present paper might be considered as a striking success story. Not only are extremely impressive gains in student achievement shown to result from the CAI format—about a 50 per cent increase in output, but these impressive benefits are obtained at only 30 per cent of the cost of even the most successful of the currently operational programs. It sounds too good to be true; unfortunately, it probably is.

For several reasons, we must be somewhat cautious in accepting without qualification the conclusions of this study. Consider first the results showing a positive impact of CAI on student achievement. The critical question in interpreting this result relates to identifying what the experiment measures; that is, what the control group represents. In the arithmetic and reading programs, students in the control and experimental groups obtained basic instruction in concepts, application, and drill from their regular classroom teachers. In addition, those in the experimental group participated in a CAI curriculum, which, in arithmetic, emphasized drill; and which, in reading, focused on phonics. The CAI curricula were a supplement to the regular classroom arithmetic and reading programs. The impact of the CAI programs, then, must be considered *relative* to the activities of the control group during the period when the experimental students were engaged in CAI. The net value of the output of CAI is thus the difference between the *value* of the

CAI output and the output obtained by students in the control group when the experiment was taking place. To the extent that the value of the control group's time was positive, then the statistical results are an overstatement of the impact of CAI on the economic value of the student's total instruction. Indeed, it is possible that if the quantity or price of the forgone output was high enough, the statistically significant positive CAI impact could be translated into a decrease in the economic value of educational output. I realize that designing and implementing an experiment which takes into account the opportunity cost of the experimental group's time is difficult. On the other hand, the benefits of this modification in experimental design could be great. For example, if an experiment were implemented where matched pairs of students were assigned to similar activities at similar times—one with CAI mathematics drill, another with nondirected mathematics drill from, say, a programmed text—we would then be in a position to interpret the economic, as distinct from only the statistical, significance of the results.

Another qualification to the results relates to the assignment of control and experimental groups. The authors report neither the assignment procedures nor do they report a test of the homogeneity of student attributes. While a random selection of control and experimental groups is not necessary, knowledge of the assignment procedures is required to interpret the results. Similarly, since their evaluation procedures do not generally take into account the impact of CAI for various types of students, it would be appropriate to examine in detail the composition of the classes involved.

From an economics point of view, the most interesting aspect of this paper relates to the distribution of the benefits of CAI. Clearly, an assessment of the economic efficiency of any public good must examine not only the output level but also its distribution. This is because the value of output will typically vary according to who receives the output. Thus, in obtaining the total value of output deriving from an educational technique, it is necessary to examine its distribution and assign or identify values for that output for alternative students. Both of these issues, seldom confronted in the education literature, are recognized by the authors. They are to be commended for stressing its importance. However, their analytical and statistical treatment of the distribution issues requires some qualification.

The main statistical model employed in this study to assess the aggregate impact of CAI is a test on the differences in the mean performance of the "typical" student in the experimental and control groups. This model, unfortunately, conceals any of the distributional effects of CAI. For example, while a positive net increase resulting from CAI was revealed, this may have been the result of positive increases for some students and negative results for others. Paradoxically, the benefits of CAI are on a *priori* grounds attributed to individualization of instruction, yet the statistical models employed for analyzing its impact deal with "average" or "typical" students and abstract from any differences in individual behavior. Regression or other statistical techniques where distributional effects are explicitly considered would appear to be more appropriate to the basic research question under consideration.

Unfortunately, in the single instance where a regression was run—a stepwise regression for the computing course, the theoretical model underlying the statistical exercise was misspecified. The regression assumes that the distribution of educational achievement is invariant to the use of CAI—that is, all interaction terms between CAI and student attributes are omitted. This model is not only inconsistent with learning theory and a substantial empirical literature in education, but also with the results of the authors in previous studies. While I suspect that the authors do not believe that CAI is an equally beneficial educational technique for all students, the hypotheses they test and the statistical models they employ abstract from the key distributional questions of who gets how much and why.

The primary method employed by the authors to examine the impact of CAI on output distribution is familiar to economists: Lorenz curves and Gini coefficients. However, the results are somewhat difficult to interpret. First, without a specific norm, one cannot determine what constitutes an important change in the Gini coefficients. Second, while the authors conclude that CAI is generally egalitarian in output distribution, I am uneasy about the finding that in the Mississippi arithmetic experiment, CAI *increased* the inequality of output distribution in three out of six cases. Moreover, CAI's distributional impact was negative in the very instance where it had its most dramatic positive effect on aggregate student achievement—the Mississippi first graders. Third, even though a comparison of the CAI and control groups shows that the direction of change toward egalitarianism was statistically significant, it should be underscored that this result is specific to the output measures of the study—arithmetic drill skills and phonics. But if we assume that the opportunity cost to the students engaged in CAI was nonzero, then clearly a complete assessment of the distributional impact of CAI would appropriately employ more comprehensive output measures, including those specific study areas relevant to the control group during the period of CAI instruction. Finally, and most important, Gini coefficients and the other measures of output distribution employed in this study are not particularly interesting from either an educational or a policy point of view. They obscure the nature of the relevant redistribution taking place. For example, a zero Gini coefficient difference, given the results presented in this paper, is fully consistent with a simultaneous reallocation of educational benefits from girls to boys and from high achievers to low achievers.

While I have qualms regarding the statistical and theoretical analysis underlying the authors' examination of the distributional impact of CAI, I do support their concern with this issue. Their interest in output distribution, as well as that of government officials and the general public, is based primarily on normative grounds. As an economist interested in issues of economic efficiency, however, I am uneasy about promoting CAI or any educational technique as a *means* for redistributing output. Put differently, CAI may or may not represent an efficient technique for redistributing output, even given preferences for such a redistribution. For example, to the extent that different students respond differently to alternative production techniques—a widespread empirical finding in the education literature—then the most efficient redistribution procedure may be one which uses several different production techniques and allocates resources among them so as to maximize output subject to a distributional goal. CAI used mainly for compensatory education, given the evidence provided, may in some settings be relatively efficient both in terms of total output and its distribution. On the other hand, we must stress again that distributional goals are also multidimensional. We may wish to redistribute over some student attributes but not others. Thus, the use of CAI for all students requiring compensatory education may be inefficient, even if CAI produces greater output and allocates this toward bright youngsters who have been low achievers, if at the same time CAI reallocates output from girls to boys (probably an unpopular outcome). Again, at the empirical level, this argues for a research design and statistical models which highlight differences in individual learning styles.

Finally, a temporal dimension might also be employed in assessing the efficiency of output redistribution. While an educational technique may distribute more output to low-achieving students, it is also relevant to identify whether this output so redistributed is retained. If, for example, retention rates are lower for low-achieving students, especially in areas such as recall and recognition where the CAI drill programs appear most effective, and if *retained* value-added is a primary objective in the education industry, then we must include in our distribution analysis not only normative judgments but also long-run technological possibilities.

The shortest section of this study, and of my comments, applies to the cost analysis of CAI. The authors have examined the cost per student of *providing* the CAI curriculum to the classroom, i.e., computer and communications installation, maintenance, and operation costs. Even though they focus on the rural school setting, I would suggest that their assumptions must be considered as generally providing a lower bound on CAI costs. While I shall not examine these costs in detail, I would urge the authors to examine and present a sensitivity analysis of the cost figures to alternative key assumptions. They assume, for example, that the terminals and central processing units are highly utilized, that floor space and rooms in schools devoted to CAI have no opportunity cost, that CAI curriculum-development costs should be excluded, that costs for CAI proctoring are low or zero through the use of volunteers, and that the opportunity cost of the student's time engaged in CAI is low or zero. With these assumptions they arrive at a per student cost of $75 which compares favorably with Kiesling's study showing per student costs of comparable compensatory-education programs ranging between $200 and $300. For at least two reasons, this comparison must be qualified. First, Kiesling's cost figures are more comprehensive than those applied to CAI. Second, Kiesling's cost function varies according to the amount of output produced. If, for the several reasons discussed above, we conclude that the CAI output estimates represent an upper bound on likely long-run production possibilities of this technique, then the appropriate Kiesling cost figures must be adjusted downward.

The paper concludes with an interesting discussion of the financial feasibility of CAI in compensatory education. Recognizing that schools are unlikely to increase significantly their budgets for compensatory education, the authors compute the increase in the student-teacher ratio required to implement a CAI program without any increase in the total per pupil cost of education. Using their cost figure of $75, they find that class sizes averaging 26 increase "only" 20 per cent. If overhead is included in their CAI calculations, the class-size increase would be 40 per cent. They conclude that an increase in class size would seem to be a "quite attractive reallocation" under the assumption that student performance is insensitive to the student-to-teacher ratio. The latter assumption is critical. On the one hand, the authors attribute to individualized instruction the impressive increase in arithmetic and reading skills provided by a short ten-to-twenty minute daily contact with CAI. On the other hand, they are willing to assume, on the basis of little evidence, that a reduction in individualized instruction through a 20 per cent to 40 per cent increase in class size for the remaining five and one-half hours of instruction per day will have a negligible impact on the size and distribution of educational output. If they are right that individualized instruction provided by small class sizes in the compensatory-education setting is unimportant—and this is a testable hypothesis—then we have a long way to go to explain why the brief exposure to the computer programs is able to bestow such significant achievements through individualized instruction. Short of such an explanation, and given the nagging possibility of Hawthorne effects, we must defer our conclusion that CAI is presently, or will be in the future, an economically viable production technique.

David E. Wiley

University of Chicago

The paper by Jamison, Fletcher, Suppes, and Atkinson has three major parts: one devoted to a description of the technology, the second to an evaluation of the performance of this technology, and the last, to an assessment of its costs of application. The second section is divided into two parts; one concerning differences in level of performance for computer-assisted and conventional instruction; and the other, differences in the distribution of performances.

SOME COMMENTS ON THE ASSESSMENT OF DIFFERENTIAL PERFORMANCE

1. Mathematics

The experiment was performed in two states (California and Mississippi) in all six elementary grades. In Mississippi, twelve schools were used (8, 3, 1);

while in California, 7 were used (2, 2, 3). The gain in grade equivalents for each treatment was compared for all six grades in both California and Mississippi. Because the assignment of treatments to schools was not random, and because the process of student selection for participation in CAI in a school is not clear, some care must be taken in the interpretation of the results. The assignment and selection process has seemingly produced some differences in the previous performances of the CAI and conventional groups. The use of gain scores (in grade equivalent units) was an attempt to eliminate some of these initial differences and to increase the sensitivity of the hypothesis test. The appropriate question relates to differences between the conditional distributions of final performance for various levels of initial performance. The gain score analysis assumes that the mean of the conditional distribution is a linear function of initial status and that the slope of the regression line is one. Both assumptions seem unreasonable for these data. The authors note that the gains in Mississippi were greater than those in California. They suggest that gains may be greater for CAI when pupils are below grade level. This would imply a nonlinear relation. Also, the way in which the CAI treatments allocate resources to pupils (more resources are allocated to slow learners in an attempt to achieve more equal outcomes) would presumably make the relation between initial and final performance smaller than for conventional instruction. The metric problem is also rather bothersome. It is not clear that equal amounts of instructional effort will produce equal amounts of gain in grade-equivalent units. This problem is distinct from the one mentioned above, since the treatments are not scaled in resource or effort units. Also the grade equivalents are based on typical performances of individuals in various grades and do not represent equal amounts of learning. Finally, the utility of various final performances would presumably vary, depending on the objectives of instruction. The problem of measurement error is also crucial for an appropriate treatment of this problem.

2. Reading

The reading study used matched pairs of individuals, one of whom had received CAI; and the other, conventional instruction. Since the CAI pupils came from two schools and the conventional pupils came from two different schools, if the school averages are different, the measurement errors in the initial performance levels have different biases for each member of the pair. Since the measurement errors are large for these achievement tests, large differences in school means could produce very different results when the errors in these variables are taken into account.

3. Computer Programming

The interesting thing about the results for computer programming is the discrepancy between the results based on the gain scores and those based on the regression analysis. The regression coefficient for the sum of the pretest scores in the gain analysis is −.26 with a standard error of .14, and not zero, as is assumed in the original analysis. Also the verbal score has a significant coefficient indicating that the groups differed in verbal ability initially. None of the other coefficients is significantly different from zero. This indicates that when initial differences are taken into account there is no detectable effect of CAI (coefficient is −.99 with a standard error of .96).

COMMENTS ON THE ANALYSIS OF DIFFERENCES IN THE DISTRIBUTION OF ACHIEVEMENT

There are conceptual problems in using test scores as quantities. The total number of items on a test is fixed in any application and the score is probably better viewed as equivalent to the proportion of items correct rather than a count variable. This is especially true in using standardized achievement tests, because the items are selected so that about half the respondents will select the correct alternative for each item. These characteristics make the relations among test scores nonlinear. Recent work in test theory has produced rather strong models for test scores, which indicate that transformations of raw scores are necessary to produce variables that have appropriate ratio scale properties. It would be useful to explore the invariance of the Gini coefficients and the Atkinson models with respect to these transformations. The raw test-score metric has some characteristics which might cause some problems with these procedures. If one of the groups is relatively close in average value to the ceiling of the test, the dispersion of that group must be less than that of the other group. This linkage between mean level and dispersion may often be removed by an appropriate transformation. These problems are related to, but are conceptually distinct from, problems of invariance with respect to classes of utility functions.

Some problems exist with respect to the sign tests performed in comparing the Gini coefficients for CAI and conventional instruction. In Table III-8 the two grade-placement differences for the computer programming experiment are correlated. This violates the assumptions of the sign test. Note that all differences are correlated in Table III-9.