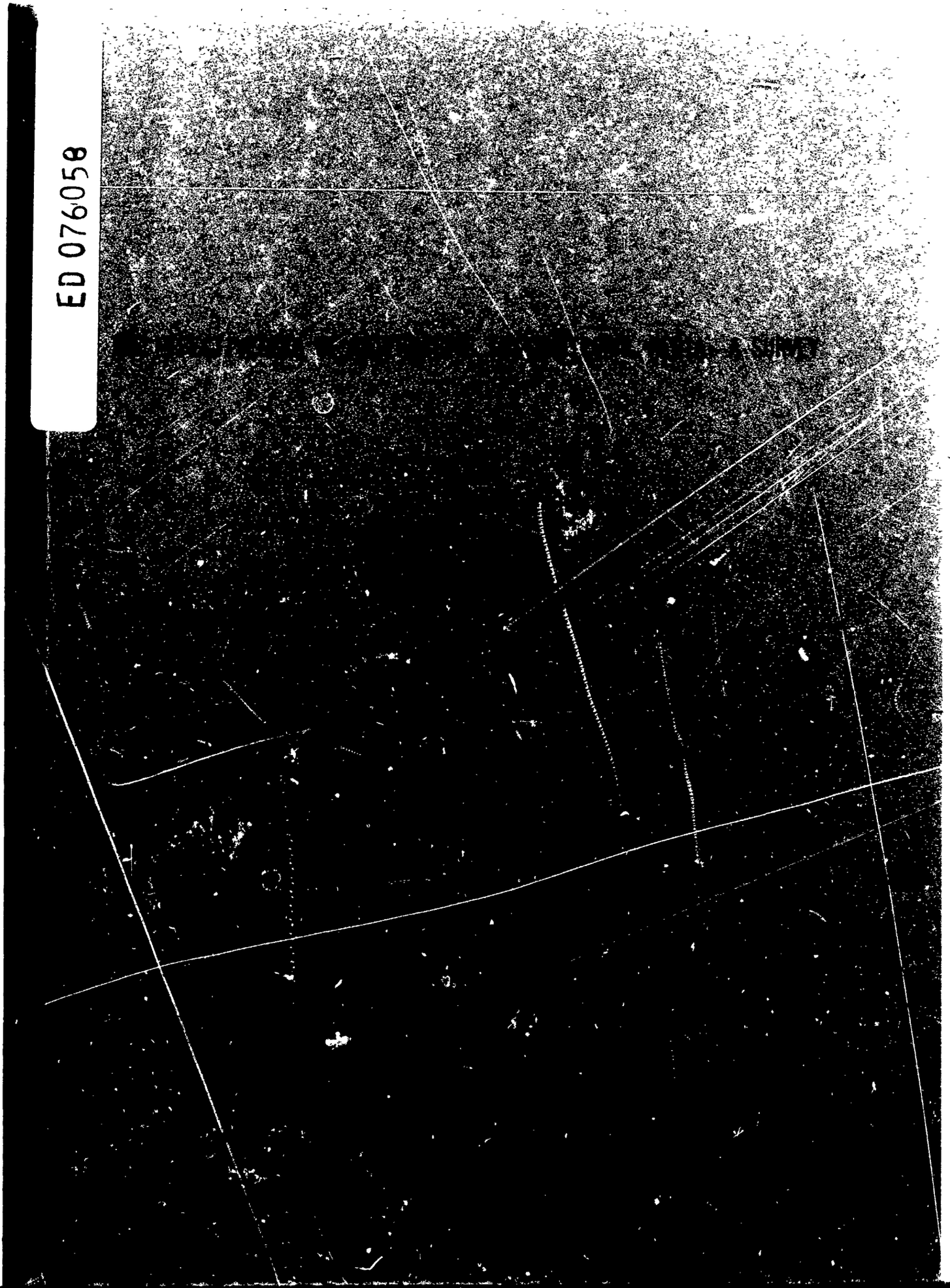
DOCUMENT RESUME

ED 076 058 EM 011 062

AUTHOR	Jamison, Dean; And Others
TITLE	The Effectiveness of Alternative Instructional Media: A Survey.
INSTITUTION	Stanford Univ., Calif. Inst. for Communication Research.
SPONS AGENCY	Agency for International Development (Dept. of State), Washington, D.C.; Federal Council for Science and Technology, Washington, D.C.
PUB DATE	Feb 73
NOTE	99p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Audiovisual Instruction; Computer Assisted Instruction; Conventional Instruction; Educational Radio; Educational Technology; Educational Television; Effective Teaching; *Evaluation; *Instructional Media; *Instructional Technology; Instructional Television; Programed Instruction; Radio; State of the Art Reviews; Technological Advancement

ABSTRACT

The results of research on the effectiveness of four alternative instructional media are reviewed: traditional classroom instruction, instructional radio, instructional television, programed instruction, and computer-assisted instruction (CAI). It was concluded that students learn effectively from all these media; relatively few studies indicate significant differences between media in effectiveness. Future research should focus on four areas: 1) determine if programed instruction and CAI can be shown to save instructional time over a longer period and with a higher percentage of students; 2) investigate long-term effects of instructional technologies on students' motivations; 3) investigate the long-term effects of the individualization and privacy made possible by some of the technologies; and 4) future investigations should consider more imaginative uses of new technologies instead of using technology to provide a simulation of some traditional method. (RH)



THE EFFECTIVENESS OF ALTERNATIVE INSTRUCTIONAL MEDIA: A SURVEY

by

Dean Jamison, Patrick Suppes, and Stuart Wells

February 1973

Institute for Communication Research
Stanford University

Support for preparation of this paper came from the Federal Council on Science and Technology's Committee on Automation Opportunities in the Service Areas, from the National Science Foundation, and from Contract No. AID/csd3284 from the Agency for International Development to the Institute for Communication Research, Stanford University. Points of view expressed herein are not necessarily those of any agency of the United States Government.

The authors are indebted to Ernest Anastasio, Joanne Leslie Jamison, Leslie Schneider, Wilbur Schramm, and Daniel Smith, for valuable comments on various aspects of this paper.

TABLE OF CONTENTS

THE EFFECTIVENESS OF ALTERNATIVE INSTRUCTIONAL MEDIA: A SURVEY

This survey provides an overview of research on the effectiveness of alternative instructional media. The media discussed are traditional classroom instruction (TI), instructional radio (IR), instructional television (ITV), programmed instruction (PI), and computer-assisted instruction (CAI). The effectiveness of these media is examined from a reasonably macroscopic point of view; the psychology of pupil-teacher interaction or the 'content variables' of ITV, to take two examples, are at a micro-level not considered. Achievement test scores constitute the measure of effectiveness most frequently used in this survey though, where available, results concerning the affective impact of the various media of instruction are included. Achievement test data, in most cases, were collected only on an annual basis, so they reveal no fine-grained detail about the learning process.

Since this survey is relatively brief and its scope broad, a few caveats are in order. First, where literature surveys are available, their results have been cited to the extent possible and, frequently, original sources remain unchecked. Second, available knowledge of the effectiveness of the various methods varies considerably; much more is known about TI and ITV than about the others. For this reason a survey such as this is inherently spotty in its conclusions. The third warning, related to the second, is that many of the evaluations fall short of (or lack entirely) scientific standards of analysis and reporting. For this reason, it was sometimes necessary to attempt to cull conclusions from essentially journalistic accounts of projects. Fourth, it should be noted that this survey is limited to instruction within a school setting. Finally, the survey excludes information on costs.

Before beginning the literature survey, we present an 'ideal' paradigm for measurement of effectiveness and then discuss several less desirable alternatives that have actually been employed. In the surveys of the individual methods where adequate prior surveys are unavailable,

results from a representative sample of individual evaluations are discussed. Where adequate surveys are available, their conclusions are presented with a description of one or a few specific project evaluations. In addition to a number of medium specific surveys there exist several reviews -- Allen [1960], Chu and Schramm [1967], and Schramm [to appear] -- that cover more than one of the topics dealt with in this review. Our objective is to attempt to bring together the overall results for all the principal media; other of the reviews mentioned here and elsewhere in our review sometimes have more detailed and specific references to the literature in some particular area than we are able to provide. The media are discussed in approximately the temporal order in which they were introduced; individuals with a particular interest in one medium are urged to skip directly to the appropriate section.

I. ASSESSING THE IMPACT OF ALTERNATIVE INSTRUCTIONAL MEDIA

An ideal study of the cognitive effectiveness of alternative instructional media would relate a vector of output measures relevant to a subject matter, including criterion-referenced measures of achievement, to the time pattern of instructional inputs. This function would include as independent variables factors not under the control of the school system so that, in its allocation of resources, the system could provide, to the extent desirable, different patterns of resource inputs to different categories of students. In order to assess the effects of different mixes of media and total amounts of time spent in learning a subject matter, we would need an experiment of vast magnitude; present survey methods are inadequate because of the current lack of substantial variation in methods of instruction. Since these methods are now virtually 100 percent TI, survey methods, as reported in the next section, can be used to assess the effect of different types of TI.

If it were to be possible to conduct an ideal experiment, the resulting function relating the educational system's outputs to its inputs would be of great value in efficient allocation of resources to and within school systems. This is primarily because the effect on output of more or less of any one input would be known as a function of the levels of all the inputs. Even with much less ambitious experimentation it is possible to obtain some idea of how output varies with input through simple multivariate regression models. For example, to assess the impact of CAI drill and practice in arithmetic (assuming CAI as an addition to and not a substitute for TI in arithmetic), let us postulate a model of the following form:

$$A_k = b_0 + b_1 A_{k-1} + b_2 C_{k-1} ,$$

where A_k is arithmetic achievement at the beginning of year k, C_k is the number of CAI sessions the student has in year k, and b_0,

b_1 , and b_2 are parameters to be estimated. Some results of CAI and TI surveyed are basically models of this form, though seldom do the TI studies have data that are either longitudinal or on a student-by-student basis. While models of this general sort, employing a variety of functional forms, give a quantitative estimate of how output varies with a few inputs, they fall short of the ideal by being inadequate for examining the impact of mixes of instructional technologies over time.

Still less informative are studies that examine whether supplementing TI with a technology or replacing it with a technology will yield achievement results that are significantly different from TI, because the magnitude of the effects, when they do exist, cannot be obtained in functional form. The vast majority of good evaluations of educational technologies are, however, of this general format.[1] The good studies provide controls by careful matching or randomizing and thus provide statistically valid results. Many more studies lack adequate controls or are in other ways flawed. That the results frequently indicate "no significant difference" is a valuable finding, not sufficiently used and appreciated in selecting a medium of instruction.

Finally, least satisfactory for purposes of assessing performance are projects whose evaluations are essentially journalistic. While much can be learned from good journalism, it is difficult to avoid feeling uneasy without supportive data, particularly if claims about substantial improvements in performance are made. It should be stressed, however, that there is no proved correlation between the effectiveness of a project and the sophistication with which it is evaluated. For this reason journalistic accounts can provide valuable screening for more detailed examination of projects that show potential for widespread use.

---

[1] This perhaps results from what the authors feel is an overemphasis on a control vs. experimental group methodology relative to a methodology that seeks to model input-output relationships. To take one example, Suchman [1967] paid almost no attention to the problem of ascertaining how the amount of effect is related to the amount of stimulus, to use his terminology.

This paper reports, then, on evaluations of varying degrees of adequacy and attempts to draw some general conclusions about the relative effectiveness of alternative instructional media. Schramm [1971] stressed the difficulties involved in making scientifically valid cross-media comparisons, and we share many of his reservations.[2] Yet a number of reasonably clear patterns do emerge from the data and these are what we report.

[2] Schramm also discussed how best to design experiments to make these comparisons. The central problem is that the number of potentially relevant variables to be controlled, or orthogonally varied, is so large that 'experiments' become substantial real-world projects over which the experimenter may end up having little control.

II. TRADITIONAL CLASSROOM INSTRUCTION

This section reviews the determinants of a student's scholastic achievement in a traditional classroom setting. Much of the work reviewed uses multiple regression analysis to relate a student's achievement test scores to attributes of his school environment (including the composition of the student body), his background and socioeconomic status, and his teachers. Many of the studies utilize the extensive data base provided by the Equality of Educational Opportunity (EEO) survey and first analyzed in Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York [1966]. Coleman et al. concluded that variation in school inputs accounts for at best a very small fraction of the variation in student achievement; socioeconomic variables, they concluded, are much more central. Other analyses, some of them also based on the EEO survey data, are summarized later in this section and a number of them find more evidence for the efficacy of the things schools provide than Coleman et al. do. However, the findings are often inconsistent and a recent thorough review of this literature [Averch, Carroll, Donaldson, Kiesling, and Pincus, 1972] listed as a basic conclusion:

> "Proposition 1: Research has not identified a variant of the existing system that is consistently related to students' educational outcomes."

In a recent reanalysis of much of the data used in the studies Averch, et al. surveyed, Jencks et al. [1972] reached much the same conclusion: "We see no evidence that either school administrators or educational experts know how to raise test scores. Certainly we do not know how to do so [p. 95]."

The reason is not that no studies have found significant input variables. Rather, Averch et al. [1972] state: "The literature contains numerous examples of educational practices that do seem to have significantly affected student outcomes. The problem is that other studies, similar in approach and method, find the same educational practices to be ineffective; and we have no clear idea why

this discrepancy exists [pp. x-xi]." The present survey accepts this basic conclusion, with only minor reservations that are stated later.

This section first presents a summary of 17 studies in a readily usable, tabular form;[3] it then provides a brief verbal summary of some of the more significant findings. After the summary of studies based on survey data, this review examines more closely the literature on what is perhaps the most economically significant variable under the control of a school system -- class size or student-to-teacher ratio. The significance of this variable lies in the observation that increasing the aggregate student-to-teacher ratio by either enlarging class size or reducing the number of periods the student spends with the classroom teacher constitutes the principal available way of substituting capital for labor in the educational system. It is thus important to ascertain what negative effects would need to be compensated for (or more than compensated for) by introducing a technology. Some of the studies we discuss are based on experimental, as well as survey data, and others report affective impact. As indicated at the outset, the present survey does not deal with relatively micro-level variables relating, for example, to the psychology of pupil-teacher interaction.[4]

---

[3] Earlier surveys of parts of this literature appear in Guthrie [1970] and Katzman [1971]; more up-to-date surveys are included in Averch et al., [1972] and Anderson and Greenberg [1972]. The studies reviewed in this literature were undertaken primarily in the United States; international data and comparisons are more difficult to find. An exception is the International Study of Achievement in Mathematics edited by Husen [1967]. This study suggests that the pattern of results found in the United States is more generally applicable.

[4] For a review of much of this literature see Rosenshine [1971]; he reviewed 51 studies of the relation between specific teacher behaviors and student achievement, and observed that studies of this sort have had a better history of finding significant influences on student achievement than have the studies of the effect of teacher personality and background variables that are the focus of the present review. Clarity of the instructor's presentational style was one of the most important variables he found.

Survey Data Studies of School Effectiveness

The review of these studies is summarized in Table 1 which is subdivided by variable. . ..cludes mainly variables under the

---

Insert Table 1 about here

---

control of school systems though Burkhead, Fox, and Holland [1967] and Coleman et al. [1966] both stressed the primary importance of the socioeconomic variables and used step-wise regressions to enter these variables first. Due to the multicollinearity problem[5] between socioeconomic group and school variables, this procedure biases the regression results in these reports in the direction of concluding that school resource variation does little to predict achievement score variation. Many studies that followed the Coleman report and used data

---

[5]The multicollinearity problem arises because, in general, higher income districts have more money to purchase higher quality resources. For example, Guthrie, Kleindorfer, Levin, and Stout [1971] examined the relationship between socioeconomic status and quality of school resources and found a positive relationship on individual, school, and district bases. The effect of entering socioeconomic variables in the regression first, as did Coleman et al. [1966] and Burkhead [1967], is that the reduction in variance attributable to socioeconomic status includes a joint effect with school resources. When school resources are entered into the equation, their importance is diminished, since only the unique contribution for school resources is measured. Mayeske [1970] evaluated the unique contributions of each set of inputs and the joint contribution of the two sets using analysis of variance. He concluded that out of the total amount of variance accounted for when both background and school variables were in the equation, 12 percent were uniquely identified with background variables, 6 percent with school variables, and 82 percent jointly. Clearly, with school resources entered second, background accounted for 94 percent and school for 6 percent of the total variance. Stratification by social class [Benson, Schmelze, Gustafson and Lange, 1965; Guthrie, et al., 1971; Hanushek, 1970; Kiesling, 1967; Michelson, 1970] is one possible method of dealing with these multicollinearity problems. In each of these studies there were some school variables which were significant.

TABLE 1

Selected School-resource Variables: Conclusions of Research Studies in Traditional Classroom Teaching

School-resource variable	Conclusions of studies finding selected school-resource variable significant[a]				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation[b,c]	Units of output variable	Units of school-resource variable	
Teacher verbal score	Adelman & Parti [unpublished]	b = .20	Verbal score scaled with mean = 27.8, s.d. = 7.2.	Score on 30-point test: m = 23.7, s.d. = 2.2.	
	Bowles [1970]	b = 1.2	Student verbal score; no units reported.	Units not reported, but probably on 30-point test.	
	Bowles & Levin [1968]	b = 1.24	Verbal score - raw score.	Score on 30-point test.	
	Guthrie et al. [1971] 24 of 30 regressions.		No units reported.	No units reported.	

[a] In studies that analyzed either many outputs or many equations, the results for each equation are reported.

[b] If available, the regression coefficient is included, where b = linear regression coefficient, β = standard regression coefficient, and σ = output elasticity. The standard regression coefficient and output elasticity are defined as follows:

$$\beta = b \frac{\text{s.d. dependent}}{\text{s.d. independent}} \quad ; \quad \sigma = \text{percentage change in output per one-percent change in input.}$$

[c] The regression coefficients are not comparable across studies.

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher verbal score (cont.)	Hanushek [1970] White-manual.	b = .09 (3rd grade teacher) b = .05 (2nd grade teacher)	Raw score: m = 55.74, s.d. = 19.1.	Score on 100-point test: m = 66.9, s.d. = 15.8; m = 68.4, s.d. = 19.	Hanushek [1970] White, non-manual; Mexican-American, manual.
	Hanushek [1968]	σ = .117 (for whites) σ = .178 (for blacks)	Verbal score units not reported.	Score on 30-point test: m = 24.8, m = 24.0.	Levin [1970] Michelson [1970] Whites, simultaneous equations; blacks, single equation.
	Michelson [1970] Whites, single equation.	b = .8	Verbal score, raw points.	Score on 30-point test.	
Teacher experience	Adelman & Parti [unpublished]	.003 - multiplier effect through recursive equations.	Verbal score scaled with mean = 27.8, s.d. = 7.2.	Years of experience: m = 13.47, s.d. = 4.78.	Bowles [1970]
	Burkhead et al. [1967] Small community, 12th grade reading.	β = .26	School mean.	6-point scale, 0-3 years to 15+.	Burkhead et al. [1967] Small community, other 3 outputs; Chicago, all outputs.

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher experience (cont.)	Carnoy [1971] 1. 3rd grade urban, Spanish reading. 2. 3rd grade rural, Spanish reading. 3. 6th grade rural, reading. 4. 3rd grade urban, general ability.	b = .384 b = .491 b = .37 b = .109	Raw score, points. Raw score, points. Raw score, points. Raw score, points.	Number of years of experience. Number of years of experience. Number of years of experience. Number of years of experience.	Carnoy [1971] 3rd and 6th grade rural for reading and 6th grade urban for reading and general ability (also stratified by SES, but those results are not included here).
	Hanushek [1968] Whites, blacks.	σ = .02 σ = .045	Verbal score units not reported.	Avg. years of exp.: m = 11.9, m = 11.3.	Guthrie [1970] Hanushek [1970] All groups.
	Katzman [1971] Reading score.	b = 1.36 σ = .184	6th grade score - 2nd grade score in grade equivalents.	% teachers > 10 years experience.	Katzman [1971] Other 5 outputs.
	Levin [1970]	b = .694	Raw score in points.	Number of years of full-time experience.	Kiesling [1969]
	Michelson [1970]	b = .6 b = .63	Verbal score, raw points.	Average years of experience.	Michelson [1970b] Black, single equation; white, simultaneous, attitude and grade aspiration.
	Thomas [1962]		18 different test scores.	m = 4.46, s.d. = 1.06.	

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher salary	Benson et al. [1965] 1. All district sizes for upper quartile salary. 2. Small district for mean salary.		5th grade median reading.	% in upper salary quartile. Mean salary.	
	Bowles & Levin [1968b]	b = 1.78	Verbal score, raw score.	Average teacher salary.	Bowles [1970]
	Burkhead et al. [1967] 1. Atlanta, dropout rate. 2. Small community, 12th grade reading.	β = -.5 β = .26	% male dropouts. School mean score.	Average teacher salary. Beginning salary, male teachers; 10-point scale, 0-1000 to $5000+.	Burkhead et al. [1967] Atlanta, all other outputs; small community, all other outputs.
	Cohn [1968]	b = .00019 r = .047	12th grade score - 10th grade score.	Median salary of high school teachers.	
	Kiesling [1969]	b = -.0097	Gain from 4th to 6th in standard grade equivalents.	% in top salary decile.	Kiesling [1969] Median teacher salary.

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher salary (cont.)	Raymond [1968] 1. Average for all teachers. 2. Average for elementary teachers.	b = .4752 b = .3895	Achievement test score. No units reported.		
	Thomas [1962]		18 different test scores.	Median starting salary, females: m = $6890, s.d. = 1240.	
Per-pupil expenditure	Benson et al. [1965] Instructional expenditure (small districts only).		5th grade median reading.	Instructional expenditure.	Benson et al. [1965] Medium and large districts.
	Burkhead et al. [1967] 1. Chicago, dropout, materials and supplies. 2. Atlanta, dropout, current expenditures. 3. Small community, reading test, instructional.	β = -.53 β = 1.23 β = .26	% dropouts, 11th grade. % male dropouts, all grades. School mean.	Materials and supplies, expenditure per pupil. Current expenditure per pupil. Total expenditure per pupil.	Burkhead et al. [1967] Chicago and Atlanta, achievement tests and all other outputs. Small community, current expenditure for all outputs.

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Per-pupil expenditure (cont.)	Kiesling [1967] 4, 5 and 6 for all occupations.	b = 4.36 for total sample.	Basic subjects, standard grade equivalent.	Per-pupil expenditure.	Kiesling [1967] Higher grade levels.
	Kiesling [1969] Urban schools, math gain.	$-.028 \leq b \leq -.0027$, range over occupations; b = -.0051, total sample.	Gain in standard grade equivalent.	Per-pupil expenditure.	
	Thomas [1962]		18 different test scores.	m = $334, s.d. = 144.	Raymond [1968] Current expenditure.
Class size	Adelman & Parti [unpublished]	-.02 multiplier through recursive equations.	Verbal score scaled with mean = 27.8, s.d. = 7.2.	Number of students in class: m = 30.3, s.d. = 6.3.	
	Carnoy [1971] 6th grade, urban, reading, general ability.	b = -.555 b = -.284	Raw score. Raw score.	Average class size by school.	Carnoy [1971] Other 6 stratifications.

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Class size (cont.)	Thomas [1962]		18 different test scores.	m = 8.09, s.d. = 1.40.	Cohn [1968]
Pupil-teacher ratio					Benson et al. [1965] Bowles [1968] Burkhead et al. [1967] Atlanta, Chicago Katzman [1971] Kiesling [1969] Raymond [1968]
ADA (average daily attendance in the school district)	Benson et al. [1965] Medium-size district.		5th grade median reading.	Average daily attendance.	Benson et al. [1965] Small and large district. Burkhead et al. [1967] Cohn [1968] Katzman [1971] Kiesling [1967] Kiesling [1969]

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher major	Adelman & Parti [unpublished]	b = .08	Verbal score scaled with mean = 27.8, s.d. = 7.2.	Proportion with elementary ed. major, m = .66, s.d. = .18.	Bowles & Levin [1970]
	Michelson [1970a] 1. White, single equation, math. 2. Black, single equation, reading.	b = 2.6 b = -7.1	Math, raw score. Reading, raw score.	Whether teacher was academic major or not.	Michelson [1970a] Both groups, verbal.
Teacher education	Carnoy [1971] 6th grade, rural, Spanish reading.	b = .88	Reading score, raw points.	Number of years beyond high school.	Burkhead et al. [1967] MA and higher, Chicago. Carnoy [1971] Number of years beyond high school. Hanushek [1970] Number of graduate units. Katzman [1971] Other 5 outputs. Michelson [1970a] Years of schooling.
	Katzman [1971] Math score.	b = -1.07	Median 5th grade score in grade equivalents.	5 teachers with MA or higher.	

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher certification	Katzman [1971] Math, special school application.	b = 1.31 σ = .286 b = .35 σ = 1.47	Median 5th grade score in grade equivalents. % taking exam for special high school.	% teachers accredited. % teachers accredited.	Burkhead et al. [1967] Carnoy [1971] Katzman [1971] Other 4 outputs. Kiesling [1969]
Teacher turnover	Katzman [1971] Attendance. Math score.	b = -.015 σ = .004 b = 1.87 σ = -.02	Rate of ADA. Median 5th grade score in grade equivalents.	Annual rate of teacher turnover.	Katzman [1971] Other 4 outputs.
	Levin [1970] Student attitude.	b = -.047	Index of a 3-point and a 2-point question.	Proportion of teachers who left.	Levin [1970] Verbal score, grade aspiration.
	Michelson [1970a] Student attitude, simultaneous equation, whites.	b = -.048	Index of student responses.	No units reported.	Michelson [1970a] Verbal score, grade aspiration.

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher attitude	Guthrie et al. [1971] Would you be a teacher again? Do you like the school you're teaching in? 23 of 30 regressions.		No units reported.		
	Levin [1970] Grade aspiration (Do you like the school you're teaching in?).	b = .693	Grade level student wishes to complete.	3-point scale for aspiration.	Levin [1970] Verbal score, student attitude.
	Michelson [1970a] 1. White, single equation (race preference of teacher).	b = 1.5	Verbal raw score.	Desired % of white students.	Michelson [1970a] White, simultaneous, verbal score, student attitude; black, single equation.
	2. White, simultaneous, grade aspiration (Do you like the school you're teaching in?).	b = .701	Grade level student wishes to complete.	No units reported.	

TABLE 1 (continued)

School-resource variable	Conclusions of studies finding selected school-resource variable significant				Studies finding selected school-resource variable insignificant
	Author(s) of study	Coefficient of variable in final equation	Units of output variable	Units of school-resource variable	
Teacher years since most recent attendance at educ. institution	Hanushek [1970] 1. White, manual, 2nd grade teacher. 2. White, manual, 3rd grade teacher. 3. White, non-manual, 2nd grade teacher. 4. White, non-manual, 3rd grade teacher.	b = -.68 b = .57 b = -.66 b = -.79	Verbal score, points.: m = 55.7, s.d. = 19.1. m = 55.7, s.d. = 19.1. m = 64.8, s.d. = 16.8. m = 64.8, s.d. = 16.8.	Number of years: m = 2.54, s.d. = 2.6. m = 1.91, s.d. = 1.6. m = 1.88, s.d. = 1.7. m = 2.02, s.d. = 1.7.	Hanushek [1970] Mexican-American, manual.
Teacher experience with SES class	Hanushek [1970] 1. White, non-manual, 2nd grade teacher. 2. White, non-manual, 3rd grade teacher.	b = .20 b = .10	Raw score, verbal test: m = 64.8, s.d. = 16.8. m = 64.8, s.d. = 16.8.	Number of years: m = 7.94, s.d. = 6.1. m = 7.85, s.d. = 8.1.	Hanushek [1970] White, manual; Mexican-American, non-manual.
Teacher tenure	Michelson [1970a] Blacks, verbal.	b = -1.1	Verbal, raw score.	No units reported.	
Teacher undergrad. institution	Michelson [1970a] White, simultaneous, verbal score.	b = 6.457	Verbal, raw score.	No units reported.	Michelson [1970a] Grade aspiration, student attitude. Levin [1970]

from it (such as Adelman and Parti, unpublished, Bowles, 1970; Levin, 1970; Michelson, 1970) do, however, show a significant relationship between various school resources and student achievement. To the education administrator or policy maker, the existence or extent of the effect of socioeconomic variables is far less important than a finding that school resources have a differential effect on children, depending on their background. Such an effect is evident in results of Carnoy [1971], Hanushek [1970], and Michelson [1970], as well as in Coleman et al. [1966].

Coleman et al. based their results on the amount of variation explained by a group of variables after socioeconomic variables were entered in the regression. Four groups of variables were used: socioeconomic, teacher, school and student body variables. For ninth and twelfth graders, the teacher characteristics added 8 percent to the explanatory power, or raised R^2 by .08 in the equation for Southern blacks, .03 for Northern blacks, .022 for Southern whites, and .015 for Northern whites. These variations might have been higher if the teacher verbal score, which according to the report bears the highest relationship with student achievement, had been included in the group of teacher characteristics. The importance of the above results is that there is a differential impact on achievement depending upon the student's race and geographic region.

Hanushek [1970] used the EEO survey data for sixth graders in the Northeast and Great Lakes region and stratified by race. He used a multiplicative model, and the regression coefficients were output elasticities (σ), that is, the percentage change in output for a 1 percent change in input. For teacher experience and teacher score on a 30-point verbal test, the results differed for blacks and whites. For teacher score, $\sigma = .117$ for whites, and .178 for blacks; for teacher experience, $\sigma = .02$ for whites, and .045 for blacks. For both teacher variables, there was a higher impact on the black achievement than on white achievement. If there were a correlation between race and socioeconomic group (with whites being from a higher socioeconomic group than blacks), these results would contrast with those of Carnoy

[1971] where teacher experience had a greater impact on high socioeconomic students than on low socioeconomic students.

In the same paper Hanushek analyzed data for third graders in one school district in California. As opposed to the EEO survey data, where average teacher characteristics by school were applied to each student or to average student achievement, Hanushek was able to match students with their second- and third-grade teachers. The students were then stratified by ethnic background (with or without Spanish surname) and by the occupation of the head of the household (manual or nonmanual labor). There are only three groups since in his sample there were no Spanish-surnamed children from a home in which the head of the household had a nonmanual job. The teacher characteristics analyzed are teacher experience, teacher verbal score (on a 100-point test), number of graduate units, teacher experience with socioeconomic class and number of years since teacher's most recent educational experience. Teacher experience and education were not significant in explaining achievement for any of the groups, and there was no teacher characteristic which explains achievement of Spanish-surnamed children. This differs from his other result that school resources have a larger effect on minority children, perhaps because of the language difficulties of Spanish-surnamed students for whom English was a second language. Hanushek [1972] provides an extensive discussion of these results.

The studies just discussed provide a sample of the type of analysis that the studies summarized in Table 1 represent. What does emerge from those studies, and from the tabular summary, is a striking lack of uniformity concerning the significance of various variables. Further, more targeted research will be required to ascertain more exactly the nature of the conditions that make significant a particular factor of instruction.

Table 1 included only studies at the elementary and secondary level; Dubin and Taveggia [1968] surveyed the results of 74 studies that compared various teaching methods at the higher education level. In most of the studies students were randomly assigned to one of two methods of teaching; the results do not give, then, regression

coefficients that could be used to examine the magnitude of the effect on output of various levels of change in input. Though individual studies may have concluded one method of teaching superior to another, Dubin and Taveggia concluded from all of the studies taken together that there was no evidence for the superior effectiveness of one teaching method over another at the college level. The methods included in their survey included lecture sections, discussion, and supervised and unsupervised independent study.

A recent regression analysis of the determinants of economics achievement, based on extensive survey data, is perhaps the best study to date of input effectiveness at the university level. Attiyeh and Lumsden [1972] summarized this long term study in a recent paper; more detailed analyses are referred to there. The output measure used was the score of the student at the end of the year on an objective examination stressing the student's ability to apply fundamental economic principles to the solution of real-life situations or problems. The independent variables included pretest score, student background variables (age, sex, year at university, general aptitude, attitudes, and field of specialization), faculty characteristics (age, experience and rank of lecturers and tutors), and course characteristics (class size, hours devoted to microeconomics, hours devoted to macroeconomics, and course materials in both lectures and tutorials). The student's attitudes toward the course and lecturer were not significantly related to posttest score[6] but the student's opinion of the "usefulness" of economics was. Of the controllable variables tutorial size was significant while lecture size (with a range of 30 to 400 students) was insignificant; rank, age, and years of experience were significant for lecturers and insignificant for tutors. The number of class hours was significant.

---

[6] In another study of student evaluations Rodin and Rodin [1972] found that "Students rate most highly instructors from whom they learn least." These findings of the invalidity of student ratings are not supported in a review paper by Costin, Greenough, and Menges [1971].

Effects of Class Size

As class size is perhaps the most economically significant variable in TI, we will deal with it in slightly greater detail at this point. In Table 1 the variables "class size" and "student to teacher ratio" were seen to be insignificant in all but 2 of the regression studies that reported using these variables. This subsection discusses a number of additional studies of the effects of class size, including several experimental studies.

A frequently cited review of the early literature on the effect of class size is Blake [1954], which is summarized in Sitkei [1968] and Varner [1968]. Blake summarized 85 studies on the effects of class size in public elementary and secondary schools. Of these, 35 favored smaller classes, 32 were inconclusive, and 18 favored larger classes. When stricter requirements were imposed on statistical procedures, 16 studies favored smaller classes, 3 were inconclusive, and 3 favored larger classes. An additional survey of early literature on class size may be found in Fleming [1959]. This is the background for more recent studies which in some cases provided regression coefficients that can be used to estimate the change in achievement to be expected with given changes in class size. In the following discussion, results are also reported in some experiments and surveys where regression coefficients or elasticities are not available, though some of these results can give an impression of the size of the effects. Several studies of the effects of class size are first summarized in the text; following that is a table summarizing these and other results reported subsequently to Blake's 1954 survey.

Frymier [1964] surveyed 12 Florida school districts and then selected all classes with more than 35 students and all with less than 30 students in the first grade. There were a total of 201 students in the larger classes and 219 in the smaller ones. The larger classes scored significantly higher at the beginning of the year on the Metropolitan Readiness Test. At the end of the school year (May) the students were given the Williams Primary Reading Test with the result that

students in the smaller class scored better at a significance level of .001. The difference in grade placement was, however, slight; for the small classes it was 1.75 and, for the large, 1.62. Though there were no controls for the many other possible factors, physical handicaps and teacher differences were checked and were not found significantly different in the two groups.

In another study at the primary (K-3) level, Balow [1969] found small classes superior to large ones; the difference was statistically significant at the .01 level, but not large in absolute terms. The classes were assigned to conditions randomly. A more detailed analysis showed that the difference was due to learning among boys. In the subsequent two years the students who had been in the small classes continued to gain more than those who were assigned to small classes after having been in a large class for the first grade; in the second year, however, the difference was not statistically significant. Balow's interpretation of the results was that small classes are important the first year; after that the difference is not significant.

In Sweden, Marklund [1963] found that in a large sample of sixth-grade classes, those with 26 to 30 students learned the most. After that came the 16-20 group (smallest) and the 31-35 group (largest). Among classes that had sixth-grade students combined with other grades, the smallest classes were favored. In comparisons among students divided into groups according to socioeconomic status, IQ, homogeneity, etc., 22 comparisons favored smaller classes, 37 favored larger, and 222 were not significantly different.

Johnson and Scriven [1967] used data from the New York Quality Measurement Program to examine the effects of class size. From the total sample only those classes within 0.3 of the mean of the class in terms of grade level on the pretest were examined. English and mathematics classes in grades 7 and 8 were the subjects of the study. Random sampling was used to derive equal numbers in cells for an analysis of variance. The results favored larger classes in 10 out of 16 comparisons. The small classes did relatively better for the seventh-grade students and for students above the mean on their pretest scores.

Table 2 summarizes a number of studies on the cognitive effects

---

Insert Table 2 about here

---

of class size that were undertaken since the time of Blake's 1954 survey. Not included in Table 2 are the extensive results of the international survey reported in Husen]1967]; volume II of that study (pp. 79-85) reports on numerous comparisons of different sized mathematics classes. The results were usually no significant difference and, where significant differences were found, they were more likely to be for older students.

While the relationship between class size and achievement is generally weak, some researchers believe that the interpersonal aspects of the classroom suffer with increased class size. Olson [1971] found an advantage for smaller classes in terms of individualism, interpersonal regard, group activity, and creativity. This survey obtained data from almost 10,000 classrooms at the elementary level and 8,600 at the secondary level. Smaller classes were favored at all levels. Using the same sample, Vincent [1968] found inconsistent relationships between class size and achievement.

In a smaller study, Cannon [1966] reported that in two kindergarten classes (one with 34-39 students, the other with 23-28 students), the smaller class was favored in terms of fewer aggressive acts, better peer relationships, more and better child-teacher contacts, more creative activities, and better feelings on the part of the teacher. The differences were not large, however.

Thus at the elementary level the quality of interaction appears to be inversely related to class size. At the secondary level the matter is not so clear. Olson [1971] reported that observational data supported less attractive styles of interaction as class size increases. Anderson, Bedford, Clark, and Schipper [1963], Ed. W. Clark High School [1968], and Williams and Koelsche [1967] reported no difference in

TABLE 2

Summary of Studies on the Cognitive Effects of Class Size

Author(s) of study	Type and level	Output measure	Input measure	Finding
Anderson [1963]	Experimental, secondary	Algebra tests	Class size	No difference with 40 or 80
Attiyeh & Lumsden [1972]	Survey, higher	Test of economics comprehension	a. Lecture class enrollment b. Tutorial section size	a. Larger favored slightly, statistically insignificant b. Smaller favored significantly
Balow [1969]	Experimental, elementary	Reading	Class size	Smaller favored in first grade, after that no difference
Burkhead et al. [1967]				
Chicago:	Survey, secondary	Various	Aggregate teacher-man-years/student	No significant regression coefficients
Atlanta:	Survey, secondary	Various	Enrollment/faculty	No significant regression coefficients
Small high school:	Survey, secondary	Various	Enrollment/faculty	No significant regression coefficients
Cohn [1968]	Survey, secondary	10th to 12th grade gain in Iowa test	Subjects/teacher	Favored fewer subjects per teacher; elasticity = -.123
			ADA/teacher	Not significant

TABLE 2 (continued)

Author(s) of study	Type and level	Output measure	Input measure	Finding
Counelis [1970]	Survey, elementary	1st grade reading scores	Class size	No significant difference
De Cecco [1964(a)]	Experimental, higher	Introductory psychology criteria test; final examination; attitude measures	Class size and organization	No significant differences
Ed W. Clark High School [1968]	Experimental, secondary	Business class tests	Class size	Two cases with no difference and one favoring smaller; approx. elasticity = -0.12
Frymier [1964]	Survey, elementary	1st grade reading scores	Class size (> 36 or < 30)	Favored smaller; approx. elasticity = -0.3
Furno & Collins [1967]	Survey, elementary	Various achievement	Class size	Favored smaller for non-white students; otherwise no difference
Guthrie et al. [1971]	Survey, secondary	Various	Classrooms/1000 students	Small but statistically significant effects in 18 of 30 cases
Haskell [1964]	Experimental, secondary	Geometric drawing	Class size	Two cases no significant difference; one case favored larger
Hopper & Keller [1966]	Experimental, higher	Writing	Class size of 28 or 56	Generally no significant differences

TABLE 2 (continued)

Author(s) of study	Type and level	Output measure	Input measure	Finding
Johnson & Lobb [1966]	Survey, secondary	Various achievement	Class size	Class of 10 favored; classes of 20, 35, 60, 70 no difference
Johnson & Scriven [1967]	Survey, secondary	English and mathematics scores	Class size (individual)	Larger classes generally favored
Katzman [1971]	Survey, elementary	2nd to 6th grade reading gain score	Students/staff Percentage of students in crowded classes	Favored larger; elasticity = 0.231 Favored smaller; elasticity = -0.06
Madden, J. [1968]	Experimental, secondary	Mathematics	Class size	Large classes favored
Mansfield [1968]	Experimental, secondary	Algebra achievement	Class size	Class size not statistically significant
Marklund [1963]	Survey, elementary	Various	Class size	Classes in the range 26-30 favored over 16-20 and 31-35; very small differences
Menniti [1964]	Survey, elementary	Reading; mathematics	Class size	Large classes favored

TABLE 2 (continued)

Author(s) of study	Type and level	Output measure	Input measure	Finding
Sorensen & Thomas [1967]	Quasi-experimental, elementary	1st and 2nd grade reading scores	Class size reduction of 26 to 22 in first grade and 31 to 27 in second grade	No significant difference
			Class size reduction as above plus added services	Favored smaller
Thomas [1962]	Survey, secondary	12th grade information	Average mathematics and science class size	Smaller favored; very small elasticity
			Average non-science class size	Smaller favored; very small elasticity
Williams & Koelsche [1967]	Experimental, secondary	Chemistry	Class size and organization	No significant differences

student-reported attitudes toward the class. The finding for secondary schools appears to hold at the community college level as well. There were no significant differences in the studies examined except that Hopper and Keller [1966] indicated that students prefer the larger classes. For students who do not particularly care to participate in discussion, the large class can be superior.

Conclusion

In concluding this survey on the effectiveness of traditional classroom instruction, it seems reasonable to agree with Averch et al. [1972] that few variables consistently make a difference in student performance. Exceptions to this general conclusion would be that teacher verbal ability appears important in a high fraction of the instances examined, and that small classes seem to improve the cognitive and affective performance of young children.[7] This conclusion does not, however, imply that schools make no difference in the cognitive development of their students; on the contrary, school attendance is clearly important in promoting academic achievement though few studies

[7] In light of this finding it is perhaps ironic that national average pupil-to-teacher ratios are substantially higher at the elementary level (24.8:1) than at the secondary level (20.0:1). These figures are for 1969 and are from the United States Office of Education [1970, p. 59]. Stevenson [1923, pp. 122-125] noted this anamolous situation a half century ago. He estimated class sizes then to average 38 at the elementary level and 25 at the senior high school level; his research concluded that the only noticeable advantages for small classes were at the elementary level, particularly for dull pupils.

seem to have examined this issue.[8] It remains to be seen that variations in school inputs are consistently related to variations in school outputs.

---

[8] Guthrie [1970] referred to a study undertaken by Green et al. [1964] on the effects of closing the schools in Prince Edward County, Virginia as a result of court-ordered desegregation. Students who attended volunteer schools scored significantly higher on achievement tests than those who did not attend school; for older students (aged 11-17) the differences were substantial. There exists more evidence on the effects of attendance or nonattendance in the literature on the effectiveness of ITV; Chu and Schramm [1967] reviewed nine examples of research that compared ITV with no instruction and in all nine those with ITV performed better. This stands in contrast to the typical "no significant difference" that predominates comparisons of ITV with face-to-face instruction. For a further discussion of the effects of school attendance see Jencks [1972, pp. 85-89].

III. INSTRUCTIONAL RADIO

Beginning in the 1920's, instructional radio was widely used in the United States, but with the advent of television and adverse regulatory decisions its use here dwindled as it did, to a lesser extent, in other developed countries. Developing countries,[9] however, make increasing use of radio and, as our evidence suggests that radio can be effective instructionally, there may be an important role for it in the developed countries as well. Its principal attraction lies, of course, in its cost, which is low when compared to television.

Early attempts to use radio for instructional purposes were rarely subject to systematic evaluation and, since IR has been used infrequently in the United States recently, available evaluation material is limited. For this reason the present review begins by providing evidence on the extent to which IR has been used in various countries as indirect evidence that it has some value. Then the conclusions of two earlier surveys on the effectiveness of IR are reported and, finally, several examples evaluating IR and audio recordings are presented in more detail.

Use of IR

Atkinson [1942 (a), 1942 (b)] provided journalistic information on a substantial number of IR projects undertaken in the United States prior to 1939; his books provide information concerning the operational

---

[9] For example, the New York Times of August 22, 1972 reported that "Shanghai is tuning in the radio daily and gleefully learning to say 'hello'." For the preceding five months, a half-hour English lesson had been broadcast three times daily and had met with great popular success. Radio has a history of use for education in China; Chang [1936] reported that its use was one of three components of a mass education program then underway in China.

problems and history of early uses of the medium in this country. Skornia [1962], Saettler [1968], and Wrightstone [1952] described the later evolution of instructional radio in the United States, and Wrightstone provided a valuable summary of early research concerning its impact. Though it is not extensively used at present, a number of school districts do continue to use radio.[10]

In Britain radio has been used extensively to provide school broadcasts. Currently 63 educational radio series are broadcast to schools in England. Almost all of these series use illustrated pupil pamphlets to support the lessons at the reception end. Within Britain, school broadcasting emphasizes collaboration between the classroom teacher and the radio teachers. Radio primarily provides lessons which the children might otherwise not receive, such as art, music, and foreign languages.

Australia also makes use of instructional radio broadcasts in its schools -- see Bull [1960] or Kinane [1967]. In 1960 over 90 percent of the schools received some radio lessons. Curriculum enrichment broadcasts, similar to those of the BBC, are used in the urban schools and even more extensively in the one-room rural schools. At the higher education level, the Radio University, of New South Wales, enrolled over 6000 students in 1965. One particularly inventive instructional radio program was originated by Miss Adelaide Miethke, a well-known educator from the state of South Australia. She arranged to use the shortwave services of the Royal Australian Flying Doctor Service to communicate with students in the isolated "outback" regions, and each outback community purchased a transmitter. So, for a limited period each day, the students are able to talk to a teacher and to each other

---

[10]The more active stations using IR in the United States at the present time include KRVM (Eugene, Oregon), WGBO (Newark, New Jersey), KSLH (St. Louis, Missouri), KANW-FM (Albuquerque, New Mexico), KBPS (Portland, Oregon), and WYNE-FM (New York, New York). Kottmeyer [1970] reported that the KSLH program in vocabulary improvement, a supplement to traditional instruction, resulted in substantial gains in IQ and spelling over controls from previous years. Evaluation material on the other programs was unavailable to the present authors.

about their correspondence lessons. The interest in IR in Australia dates back to at least the 1930's. One of the first statistically sound IR evaluations was undertaken there, by Thomas [1937], and he reported no statistically significant differences in achievement in most cases. There was a tendency for the TI students to do slightly better on an immediate post-test and for the IR students to do better after a delay. The amount of exposure to IR was, however, small.

Another country making widespread use of instructional radio is Japan. In 1935, Nippon Hoso Kyokai or the Japan Broadcasting Corporation (NHK) began a small program of radio broadcasts to the school [Hatono, 1960; NHK, 1964]. After World War II, a decision was made to modernize completely the Japanese educational system, in terms of both curriculum and teaching technique. Radio played a large role in this modernization in compensating for the many textbooks lost during the war and in rapidly disseminating the new methods of instruction. A 1958 survey by the Broadcasting Culture Research Institute of the NHK reported that 47 percent of the primary schools, 37 percent of the lower secondary schools, and 27 percent of the upper secondary schools regularly used radio broadcasts. In Japan it is possible to receive a secondary level diploma without attending a classroom through a combined program of correspondence courses and radio lessons.

One of the more successful uses of radio in a developing country has been in Thailand when broadcasts to the schools began in 1957 and by 1965 reached over 800,000 students with lessons in English, social studies, and music. Students receive an average of 10 to 30 minutes of instruction weekly in each subject, as supplements to their regular lessons. Schramm [1967] summarized the Thai experience and reported on a 1959 evaluation by the Thai Ministry of Education. The evaluation showed students who received the radio music supplements to be significantly superior ($p = .001$) along several dimensions to student who did not; the English lessons showed no such consistently positive effect and were subsequently extensively revised. Perhaps most interesting were the lessons in social studies, the purposes of which were to inculcate socially desired values. The Ministry evaluation

concluded that this objective was being met since a significantly higher percentage of radio students expressed agreement with desired attitudes and values in a questionnaire.

These descriptions provide only a sample of the instances in which radio has been utilized abroad; Williams [1950], Bereday and Lauwerys [1960], and Leslie [1971][11] described additional examples.

Surveys of IR Evaluations

Two surveys review information relevant to the effectiveness of IR. One is Section VI of Chu and Schramm's [1967] comprehensive review of learning by television. The second is a position paper by Forsythe [1970] that, in an earlier form, was prepared for the President's Commission on Instructional Technology. Sources of further information on IR may be found in a 432-entry indexed bibliography compiled by R. Madden [1968], and an early review of research undertaken primarily in the late 1930's and early 1940's may be found in Woelfel and Tyler [1945].

Chu and Schramm [1967] numbered the principal conclusions of their extensive survey. The ones most relevant to IR follow.

"53. Given favorable conditions, pupils can learn from any instructional media that are now available.

"58. The use of visual images will improve learning of manual tasks as well as other learning where visual images can facilitate the association process. Otherwise, visual images may cause distraction and interfere with learning.

"60. Student response is effectively controlled by programmed methods, regardless of the instructional medium."

Their general conclusion is that radio, particularly when appropriately supplemented by visual material, can teach effectively and, for many purposes, as well as other media.

---

[11] The present review draws to some extent on this unpublished paper by Leslie.

Forsythe [1970] reached a similar conclusion. In summarizing studies of radio's effectiveness he concluded:

> "Research clearly indicates that radio is effective in instruction. Experimental studies comparing radio teaching with other means or media have found radio as effective as the so-called 'conventional methods.' Even though radio has been criticized for being only an audio medium, studies have shown that visual elements in learning are not uniformly important. In many educational situations visuals may be more harmful than helpful. Also, the efficiency of combined audio and visual media has been challenged by studies which show that multi-channel communications may not be inherently more effective than single channel presentations."

To support his conclusions, Forsythe listed, among others, studies of Carpenter [1934], Cook and Nemzek [1939], Harrison [1932], Heron and Ziebarth [1946], Lumley [1933], Miles [1940], and Wiles [1940]. He also mentioned two experiments by NHK in Japan [NHK, 1955, 1956] that favored radio. Forsythe, along with Chu and Schramm, concluded that IR compares well with TI. It should be kept in mind, though, that most of these studies are old, and that in many of them the statistical controls were imperfect, the amount of instruction carried by IR was small, or the classroom teacher did participate in the program. Nonetheless, we believe that the overall conclusions of Chu and Schramm and of Forsythe are consistent with the available evidence. We also feel that there is substantial value, particularly for developing countries, in obtaining much more extensive evidence on the effectiveness of IR; of particular importance would be experiments using IR to carry the bulk of instruction in one or more subject matters for periods of at least one academic year.

Specific Evaluations of IR

To give a more concrete impression of the results of this research, this subsection discusses several of the better studies in more detail;

these include studies that compare IR with ITV as well as some that compare instruction by audio tape to TI.

McLuhan [1964] summarized an interesting study in which four randomized groups of university students were given the same information about the structure of preliterate languages. One group received it via radio, one by TV, one by lecture, and one read it. In all cases the information was given in a straightforward manner, unembellished with teaching aids. The first results indicated that the students learned more from TV and radio teaching than they did from lectures and print, and that the TV group stood above the radio group. However, when the experiment was repeated using improved auditory and visual aids, the relative effectiveness of the different media changed. Television and radio once again ranked above lecture and print. Unexpectedly, however, radio stood significantly above TV. In this experiment, TV seemed to fare less well as a teaching medium because of limited audience participation; better results were obtained with IR because of efforts to engage the students (asking them to look at certain illustrations, etc.).[12]

One interesting and detailed evaluation is an early study of the Wisconsin Research Project in School Broadcasting [1942] of radio lessons in music. A music course was first broadcast in 1922 and an evaluation was undertaken in 1929. This evaluation indicated that the music course was highly successful and so, in 1931, the Wisconsin School of the Air began a series of weekly broadcasts called "Journeys in Music Land," the effectiveness of which was studied during 1937 and 1938. The students who participated in these radio classes were in the fifth and sixth grades in both rural and urban schools. By March 1, 1938, there were 814 listening classes in 770 schools. The aim of the broadcasts was not only to teach music appreciation, but also to teach children to sing and read music. The broadcasts were planned around

---

[12] Without formal evaluation Skornia [1968] reported that in Holland and the Scandinavian nations IR had been found better than ITV for some subjects when exercise manuals and other student participation materials were used simultaneously with the radio lesson.

a minimum of assistance by the classroom teachers, because most of the teachers had no skill in musical instruction. The first year of the experimental music culminated in a radio music festival in May 1938 and the evaluation states:

> "It was evident to all who heard the Radio Music Festival that the children had learned to sing with clarity of diction and beauty of tone. Their enthusiasm and their enjoyment of song made the whole performance impressive."

The more systematic evaluation of effectiveness compared 12 classes that listened to the music broadcasts with 8 comparable classes whose teachers pursued the stated aims of the radio broadcasts, but did not use the broadcasts themselves. The classes were matched on the basis of number of students, teacher competence, and available musical equipment. Measures of the students' sex, grade level, previous musical training, chronological age, and mental age were also made, although these were not used initially to match the experimental and control classes. When examined, however, these factors did not affect the amount of student gain. The experimental period lasted 15 weeks during which classes in both groups received a total of 75 minutes of musical instruction each week. For the IR classes this was divided between a 25-minute broadcast once a week and 40 minutes of supplementary classroom practice.

Several tests were devised to measure the gains of the students. The measure of singing quality showed no differences between the radio and control classes except that the IR classes maintained better rhythm. On ability to sing an unfamiliar song at sight, the initial scores of the IR classes were significantly lower and their gains were significantly greater. Again, the most significant difference was in ability to maintain correct rhythm. The IR classes and control classes did not differ significantly in their initial ability on the test of technical skills. On the final test, however, the IR classes scored significantly better in their ability to recognize note values, read at sight, and recognize rhythms; there was no significant difference between the IR and control classes in ability to take musical dictation or staff dictation.

The Wisconsin Research Project evaluated six other radio series in addition to the music one we have just described, and the volume they produced remains perhaps the best single source of evaluative material on IR. While the other studies they reported were less favorable to IR than was the music evaluation, they provided ample evidence for the capability of IR to carry important segments of the curriculum.

Several more recent studies that were carefully controlled examined the effect of substituting an audio-tape presentation for live lectures. Popham [1961] divided an introductory graduate level course into two sections. In one he taught in a lecture-discussion format; in the other, he played a tape-recorded version of the lecture and then led a brief discussion period. The two sections were matched on scholastic aptitude and two achievement pretests; on several posttests, Popham found no significant differences between the two sections. In order to test the importance of having the course instructor present for the discussion sessions, Popham [1962] performed a similar experiment in which lectures presented by tapes were followed by a discussion led by a relatively untrained student. Again, no significant differences between conventionally taught and tape-taught students were found. In b xperiments students had generally favorable attitudes toward instruction by audiotape. They felt the lectures were better organized, and they felt freer from distractions. However, they were dissatisfied with their inability to question or disagree with the instructor during the lecture.

Menne, Klingensmith, and Nord [1969] extended Popham's work by providing each student with a tape recorder and a complete set of taped lectures that allowed each to work at his own pace. They recorded and edited lectures for an introductory psychology course taught every quarter at Iowa State University. The blackboard notes from the lectures were prepared in booklet form. For two academic quarters they compared students who took the course solely from audiotape with students who took it from the lecturer from whose earlier lectures the audiotapes had been prepared. A total of 290 students elected to take

the course by tape, while 408 chose the live lectures. In spite of the self selection, the two groups were closely matched in terms of their high school rank in class and measures of achievement and scholastic aptitude.

In terms of posttest scores and final grades, the two groups did not differ significantly. When comparisons were made in terms of groups ordered (into quartiles) by high school rank in class, there was a clear advantage to using tapes for the lowest quartile; for the others there was no difference. A possible explanation might be that the poorer students were able to listen more than once to lectures they had failed to understand the first time. A final interesting difference between the two groups was that only five of the students learning by tape dropped out, whereas 58 attending the lecture sections dropped out. Menne et al. [1969] speculated that it is less likely that students will fall irremediably behind if the tapes are always at hand.

Conclusions

Radio has been used extensively for formal classroom instruction in the United States (more in the past than at present) and elsewhere. There exist, however, only a limited number of good evaluations of the effectiveness of IR. These evaluations indicate that IR (supplemented with appropriate printed material) can be used to teach most subjects as effectively as a live classroom instructor or ITV. Due to the limited number and scope of good evaluations now available, and to the potential economic significance of IR for developing countries, much more research — both survey and experimental — is highly desirable.

IV. INSTRUCTIONAL TELEVISION

This section is briefer than the others, because two thorough and recent reviews of the literature on the effectiveness of ITV already exist: Chu and Schramm's [1967] Learning from Television: What the Research Says, and Dubin and Hedley's [1969] The Medium May be Related to the Message: College Instruction by TV. Conclusions of these reviews are summarized first with respect to achievement and then with respect to attitudes toward the use of the medium.[13] The present review does not cover the literature on the instructional use of film because of its close similarity to ITV; for a good overview of the research on film see Allen [1960, pp. 116-118].

ITV and Student Achievement

Chu and Schramm surveyed 421 comparisons of ITV with TI that are reported in 207 separate studies. Tables 3 and 4, reproduced from Chu and Schramm, summarize a number of their findings on the relative instructional effectiveness of the two media. Table 3 indicates that

---

[13] Two recent projects not covered in these two previous surveys are worth mentioning. During the last few years probably the most intensive evaluation of an ITV project was initiated and is now almost complete. This was a U.S. Agency for International Development funded evaluation of the educational reform and introduction of ITV into grades 7-9 in El Salvador. Schramm [1971] provided a summary of that research to date; more detailed information may be found in McAnany, Mayo, and Hornik [1970]. In a second project, at the postgraduate level, Colorado State University provides M.S. level courses to engineers at corporations and government research laboratories throughout the State of Colorado. Over 12,000 quarter hours of university credit were earned and 24 M.S. degrees awarded through this program to date. For a discussion of evaluation and costs see Baldwin, Davis, and Maxwell [1972].

Insert Tables 3 and 4 about here

students at all grade levels learn well from ITV, though this seems somewhat less true for older students than for younger ones. Table 4 indicates that the effectiveness of ITV cuts across virtually every subject matter.

Dubin and Hedley [1969] provided a more detailed survey of the effectiveness of ITV at the college level. They reported on 191 comparisons of which 102 favored ITV and 89 favored TI, although most of the differences were insignificant at standard levels of statistical significance. When data were available, Dudley and Hedley extended their comparisons to include the distribution of the t statistics of the individual comparisons of ITV and TI; in this way it was possible to weight appropriately differences in performance of differing degrees of statistical significance. The results of this analysis, applied to all their data, indicated a slight, but statistically significant difference in favor of TI. When studies of two-way[14] TV were dropped from the sample, the overall comparison yielded a small, statistically insignificant advantage for TI. Figure 1 shows the distribution of t statistics for this sample.

Insert Figure 1 about here

An unusually stringent criterion for interpretability of results was utilized by Stickell [1963] in comparing ITV to TI, and it is worth commenting on his survey here. After examining 250 comparisons of ITV

[14] Two-way TV incorporates an audio-return capability that allows students to ask questions during a live ITV broadcast. The 26 comparisons of this mode of instruction with TI yielded a highly significant advantage for TI.

TABLE 3

Results of 421 Comparisons Between ITV and TI

[Chu & Schramm, 1967]

Level	Number of cases of		
	No significant difference	ITV more effective	TI more effective
Elementary	50	10	4
Secondary	82	24	16
College	152	22	28
Adult	24	7	2
	308	63	50

TABLE 4

Relative Effectiveness of ITV and TI, by Subject Matter

[Chu & Schramm, 1967]

Subject	Number of comparisons	Percentage of comparisons in which ITV did as well or better than TI
Mathematics	56	89.2
Science	100	86.0
Social studies	77	89.6
Humanities	45	95.5
Languages	77	88.3
Skills	26	96.1
Miscellaneous	40	75.0

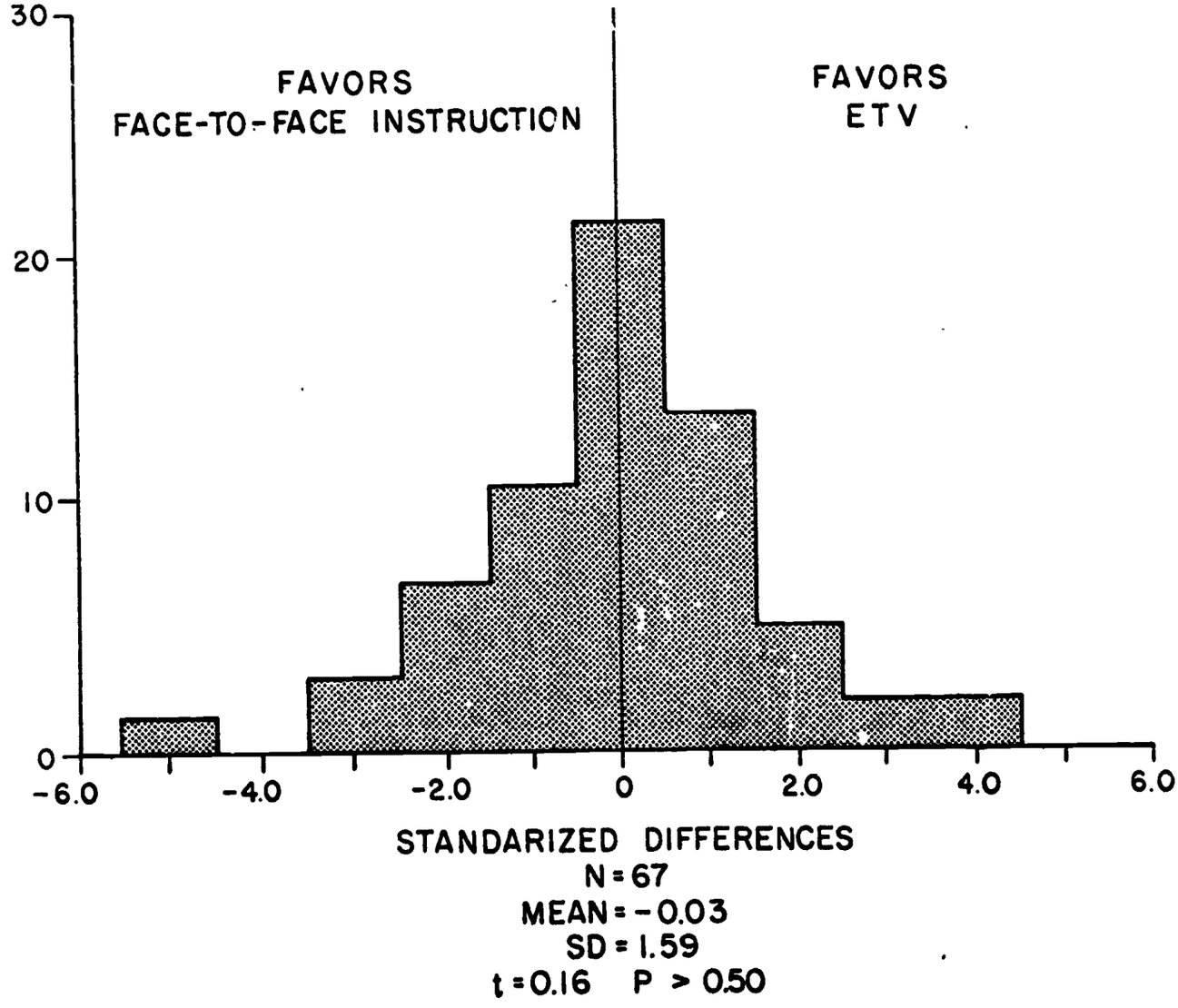


Fig. 1. One-way ITV compared to TI, independent comparisons. (Adapted from Dubin & Hedley [1969], Figure 3, p. 19. The measure on the horizontal axis is the value of the t ratio for the comparison.)

to TI Stickell found 10 studies that fully met his requirements for adequate controls and statistical method (interpretability) and 23 that partially met his requirements. Schramm [to appear] provides clear tabular summaries of these studies. None of the fully interpretable studies and 3 of the partially interpretable ones showed statistically significant differences; each of the three statistically significant cases favored the ITV group. It should perhaps be noted that when highly stringent controls are imposed on a study, the nature of the controls tends to force the methods of presentation into such similar formats that one can only expect the "no significant differences" that are in fact found. When ITV is used in a way that takes advantages of the potential the medium offers -- as, perhaps, with Sesame Street[15] -- we would expect more cases of significant differences between the experimental group and the "alternative treatment" (for it would not be a "control" in Stickell's sense) group.

Attitudes Toward ITV

Chu and Schramm summarized their conclusions in a series of numbered paragraphs. The ones relevant to attitudes are quoted below. They noted at the outset that "the research evidence makes attitudes toward instructional television seem rather more favorable than one would expect from the experience reports that circulate. Regardless of this evidence there is good reason to think that some resistance

---

[15] As a program designed for pre-school age children, and for viewing out of school, Sesame Street falls outside the scope of this survey. Evaluation of the first two years of Sesame Street -- see Bogatz and Ball [1971] -- indicated that it had a significantly positive effect on disadvantaged pre-school age children in terms of a large fraction of the specific goals the producers set for the program. A problem with the first year's evaluation was that there may have been a correlation between frequency of viewing and other variables tending to promote achievement; this was partially corrected for in the second year by facilitating and encouraging viewing by a randomly chosen half of the subjects and not doing so for the other half.

among teachers has been aroused wherever and whenever television has been introduced for purposes of direct teaching." Their numbered conclusions are as follows:

"37. Teachers and pupils are more favorable toward the use of ITV in elementary school than in secondary school and college.

"38. Administrators are more likely to be favorable toward ITV than are teachers.

"40. At the college level, students tend to prefer small discussion classes to television classes, television classes to large lecture classes.

"41. Favorable attitudes are distributed widely enough among different televised courses to cast doubt on the assumption that some academic subjects, per se, may be disliked as material for ITV.

"42. There is evidence of a Hawthorne effect among students beginning to use ITV, but no firm evidence that attitudes toward the medium necessarily improve or worsen with time.

"43. Liking ITV is not always correlated with learning from it."

Dubin and Hedley presented a slightly more optimistic view of attitudes toward ITV by college professors and students. Professors, they found, are generally favorable toward ITV though a substantial majority of them would rather send their own children to a university using TI rather than one that was otherwise similar but that used ITV for its large introductory classes. Junior faculty and faculty who have taught a number of large lecture classes tend to favor the introduction of ITV.

Dubin and Hedley also reviewed a number of studies on the attitudes of college students toward ITV. Students have more favorable attitudes toward ITV after they have experienced it than before; after exposure to ITV half to two-thirds of the students surveyed reported attitudes that were favorable (as opposed to neutral or unfavorable).

Asked whether they would choose ITV or TI, less than one-third indicated a preference for ITV (and here there is substantial variation among institutions). If, however, the choice was between ITV and TI in the form of a large lecture course, typically over half the students preferred ITV.[16] Dubin and Hedley concluded that "the college student as consumer of teaching does not exhibit any significant resistance to the introduction of educational television into his own instructional program. He will take whatever method or medium of instruction is offered, damn or praise it on its merits, and get on with the business of pursuing his college education [p. 86]."

In a particularly interesting study Greenhill, Carpenter, and Ray [1956] examined perhaps the best indicator of students' attitudes, their own free choices. In a university level chemistry class 312 students were required to attend lectures for five weeks in the large lecture hall and for five weeks in a relatively small TV classroom. The students were then given their choice concerning which way to continue the course; about one third selected TV. A large fraction of students had no strong preference.

Conclusions

ITV can teach all grade levels and subject matters about as effectively as TI, though some evidence indicates that it performs relatively better at lower grade levels. A significant fraction of teachers and students have initially negative attitudes toward ITV; these negative attitudes tend to lessen, but not necessarily disappear, with time and appropriate administrative behavior. Evaluations that report "no significant difference" between ITV and TI are usually based

---

[16] Kinane [1967] reported that students in a calculus course at the Australian Radio University (which also utilizes ITV) expressed a "strong preference" for the television over the radio version of the course.

on experimental designs that hold almost everything but the medium constant. It is plausible -- though not, to our knowledge, experimentally verified -- that attempts to use the distinctive potential of the television medium would result in more systematic findings of significant differences between ITV and alternative treatment groups.

V. PROGRAMMED INSTRUCTION

Although in recent years the intensive evaluation of PI has considerably lessened, over the past 15 years many evaluative studies have been made. We review a number of them briefly in this section. We first state the conclusions of several previous reviews of the literature then summarize a number of more recent studies. There exist several valuable anthologies of papers on PI — including Lumsdaine and Glaser [1960], DeCecco [1964 (b)], and Glaser [1965] -- and the interested reader is referred to these for useful source materials.

Previous Reviews

After a review of 35 field experiments, Silberman [1962] found that all of them showed that PI took less time to complete than TI. Furthermore, in 9 of the studies students in the PI groups scored higher than their counterparts. In the other 6 studies there was no difference between the two approaches.

Another good survey of the earlier research on programmed instruction is Schramm [1964]. Schramm introduced an annotated bibliography of approximately 190 research studies in the area of PI with a summary evaluation of those studies. Thirty-six of the studies he reviewed compared PI with TI; of these, 18 showed no significant difference in performance between the PI and TI groups, 17 showed a significant superiority for PI, and only one showed superiority for TI. In 8 of the studies PI students needed less time to complete the required materials than did the TI students. Schramm also reviewed the evidence concerning the importance of such variables as sequence ordering, step length, error rate, constructed responses, and feedback to the student on the accuracy of his answers.

In a more recent review, Lange [1972] reported that between 960 and 1964, 112 comparative studies were conducted that aimed at

matching PI and TI. Of these studies 41 percent showed PI to be superior, 49 percent found no difference, and 10 percent found PI to be worse than TI. The studies he reviewed are not, however, completely independent of those reviewed by Schramm.

Zoll [1969] provided a fairly extensive review of research in PI in mathematics. He undertook a review of 35 studies reported in the literature, many of them in the form of dissertation abstracts. He cited 7 studies that specifically evaluated one or more of the commercially produced PI programs in comparison with TI. Of these and the other studies the most common conclusion is that no significant differences were apparent. It is important to realize that such results are standard in the mathematics education literature. One explanation is that the variance in individual ability and achievement is large enough to make it difficult to establish significant differences due to different methods of instruction. Another possible explanation, consistent with Silberman's findings, is that while student achievement may not significantly differ, less student time may be required with PI. (Lumsdaine [1963, pp. 611-613] discusses the importance of time as an instructional variable.) Of the 35 studies reported by Zoll, 10 included results from attitude questionnaires on student reaction to PI. While responses were generally favorable, three studies [Alton, 1966; Little, 1967; Meadowcraft, 1966] indicated that interest decreased with time.

Along related lines, Peterson [1972, unpublished] surveyed work in the area of mastery learning. Mastery learning is a general term used to describe a programmed instructional process in which a subject matter is subdivided into many smaller units and each student attains a mastery of a specific unit before being advanced to the next unit. Advancement is based on the percentage of correct responses on a test of the current unit. A variety of materials may be used in the teaching of the subject matter including audio-visual methods, tutorial help, workbooks, games, and small group study. Peterson surveyed a total of 21 studies in mastery learning; some of the studies reported results of more than one experiment. Achievement measures included grade in

course (% A or B) and scores on post-tests. Comparisons were made with traditionally taught courses and twenty-four of the experiments favored mastery learning while there was no difference reported in four of the experiments.

Specific Studies

It is natural then to ask in what areas is PI effective and for whom? A wide range of examples on the use of PI may be cited. Brigham [1970] used programmed texts to teach woodwind fingering; Bullmer [1972] used programmed materials to teach accuracy of interpersonal perception; Ashford [1968] used PI to teach fundamental concepts of music theory and found that three years after the 11-week course, on a recall examination, students in the PI group performed better than students receiving TI.

In an extensive study, Johnson [1966] compared three different programmed textbooks and two conventional texts in 21 elementary algebra classrooms. The texts and programs were all prepared as part of the School Mathematics Study Group project. He found that one of the texts was the most satisfactory for each of the three ability levels, high, middle and low, but good achievement results were obtained by both high- and middle-ability-level students using the PI units.

In a study concerned only with low arithmetic achievers, Tanner [1966] found no differences in achievement between seventh-grade students using PI under teacher supervision and students receiving TI. In the same spirit Bobier [1965] found no significant differences among twelfth-grade students using either PI or TI to improve weaknesses in arithmetic skills.

Another area of research concerns the effects of individual differences. The finding seems to be in general that the intelligent students [Williams, 1963, 1965] and the creative students [Tobias, 1969] profit more than other students in terms of speed of learning and posttest scores. However, these results are hardly surprising for we would expect such results from almost any form of instruction.

Results of somewhat greater interest are the findings by Shrabel and Sassenrath [1970] that anxious students outperform students with low anxiety, that an easy program with short steps is better suited to persons who are low on need for achievement and high on fear of failure or text anxiety, and that a hard program with long steps is preferable for those with a high need for achievement and low fear of failure. However, Tobias and Abramson [1971] failed to replicate this anxiety finding.

In one of the largest studies to date, Attiyeh, Bach, and Lumsden [1969] reported on an experiment in introductory economics that was conducted simultaneously in 48 colleges and universities and that involved over 4000 students. The students were divided into three groups. The first group studied a programmed text during, on the average, the first 3 weeks of the term and attended no lectures. The second group supplemented regular instruction with a programmed text, and the third group served as a control. The results of the experiment were analyzed by multiple regression to control for differing characteristics of the students and the schools they attended. Of the two programmed texts used, one proved to be significantly better than the other. For the superior text, students in the first group who only read the programmed text did less well (but statistically insignificantly so) than the controls in the third group; students in the second group who supplemented their regular course with that programmed text did significantly better than controls. The first group, who read the programmed text only, experienced a substantial time saving. It is of interest that this study, by using two separately prepared sets of programmed materials, illustrates the difficulty of simply comparing media without simultaneous consideration of content. Use of the poorer programmed text as a supplement actually weakened the performance of students.

Another example of one of the better studies on the effectiveness of PI is Doty and Doty [1964]. These authors studied the effectiveness of a programmed unit on physiological psychology for 100 introductory psychology students [Kimble, 1963]. The program had 1,507 frames, was assigned as required outside classroom work and was not discussed in

any class meetings. Students were given two weeks to complete the program, after which they were tested by means of a 75-item multiple-choice achievement test. The scores on the test were used as the index of PI effectiveness.

Doty and Doty were interested in the intercorrelations between the PI achievement as measured by the test and student characteristics. The following student characteristics were studied: Academic ability as measured by cumulative GPA; achievement motivation as measured by the Edward's Personal Preference Schedule, Achievement Need Scale Scores; creativity as measured by means of Getzels' and Jackson's four tests of creativity; social need as measured from scores on the Guilford-Zimmerman Temperament Survey. Achievement on the PI unit was found to be significantly related to GPA, social need, and creativity; the correlations with creativity and social need were negative.

Conclusion

In evaluating the effectiveness of PI for use in various educational settings, the study of Doty and Doty suggests the kind of research required in the future. A better understanding is needed of how student personality variables differentially relate to achievement in PI. More generally, the current research emphasis in PI seems to have changed from direct comparative studies of effectiveness to detailed studies of how to improve the programs, how to increase student interest, and how to adapt PI to unusual educational settings. In the meantime, on the basis of the research to date, it is reasonable to conclude that PI is generally as effective as TI and may result in decreasing the amount of time required for a student to achieve specific educational goals.

VI. COMPUTER-ASSISTED INSTRUCTION

Among the alternative instructional media considered in this survey, not only is CAI the newest, but in terms of the initial cost of instruction per hour, it is also the most expensive. At the same time, however, this technology provides the richest and most highly individualized interaction between student and curriculum of any of the methods of instruction yet developed.

Prior to the early 1960's projects in CAI were virtually unknown. Until 1970 or 1971, almost all the projects were developed in university research settings, especially in universities with rich computer resources. In the last few years, however, a number of school districts have begun to run their own CAI courses, and even though it is not presently possible, extensive evaluations of CAI, separated from the stimulus and supervision of a computer-based research center, should be available in the near future.

Given the data collecting and analyzing power of computers, it is surprising that more recorded evaluative studies on the effectiveness of CAI are not available in the literature. Part of the explanation is probably that during the first years of developing this new method of instruction the main efforts have gone toward solving the technical problems associated with the ongoing operations and only in the last several years have there been adequate time and opportunity to make systematic evaluative studies. Evaluations of the effectiveness of CAI programs have, nonetheless, been conducted for most levels of education. The most intensively researched area is that of the effectiveness of drill and practice programs in elementary mathematics and reading, and we begin our survey with a review of that research. We then turn to a number of studies conducted at the college level.

Elementary-school Drill and Practice

The available CAI drill-and-practice programs provide a supplement to the elementary student's regular instruction in mathematics or language. Several times a week the student receives sessions at the CAI terminal that last about 10 minutes; these sessions provide intensive drill in the concepts he is learning in his regular class. Vinsonhaler and Bass [1972] recently surveyed over 30 separate experiments (involving a total of about 10,000 students) that compared TI to TI augmented by CAI drill and practice at the elementary level. They concluded that "... there appears to be rather strong evidence for the effectiveness of CAI over traditional instruction where effectiveness is measured by standardized achievement tests." In this survey we review several of those experiments and then summarize several other studies that attempt to relate amount of achievement gain to amount of CAI.

Suppes and Morningstar [1969] reported the results of the evaluation of drill-and-practice programs for schools in California for the 1966-67 and 1967-68 academic years and for schools in McComb, Mississippi for 1967-68. The programs they discussed and analyzed for supplementary drill and practice were given to the students on a daily basis. Students spent not more than 10 minutes a day at teletype terminals connected by phone line to the computer at Stanford.

To evaluate the effectiveness of the drill-and-practice programs, they administered the arithmetic portion of the Stanford Achievement Test to both experimental and control classes, using different forms for the pretest and the posttest. Tests were given in four California schools for the 1966-67 evaluation. The pretest, posttest, and differences for experimental and control groups on the computation sections are shown in Table 5. Table 6 shows similar data for the students from Mississippi for the 1967-68 school year.

Insert Tables 5 and 6 about here

TABLE 5

Average Grade-placement Scores on the Stanford Achievement Test: California, 1966-67

[Suppes & Morningstar, 1969]

Grade	Pretest		Posttest		Posttest-pretest		t	Degrees of freedom
	Experimental	Control	Experimental	Control	Experimental	Control		
School A versus School B								
3	2.9 (51)	3.0 (63)	3.9	3.6	1.0	0.6	2.50*	112
4	3.9 (60)	3.9 (75)	4.7	5.3	0.9	1.4	-2.93*	133
5	4.6 (66)	4.6 (81)	5.2	6.3	0.7	1.7	-4.74*	145
6	4.9 (50)	5.2 (70)	7.1	7.1	2.1	1.9	0.95	118
School C versus School D								
4	3.7 (61)	3.8 (63)	5.4	4.8	1.7	1.0	4.50*	122
5	5.4 (63)	4.9 (77)	6.3	5.4	0.8	0.6	1.32	138
6	5.8 (58)	6.0 (56)	7.4	7.1	1.6	1.1	2.19**	112

[a]Values in parentheses are numbers of students.

* $p < .01$.

** $p < .05$.

TABLE 6

Average Grade-placement Scores on the Stanford Achievement Test: Mississippi, 1967-68

[Suppes & Morningstar, 1969]

Grade	Pretest[a]		Posttest		Posttest-pretest		t	Degrees of freedom
	Experimental	Control	Experimental	Control	Experimental	Control		
1	1.41 (52)	1.19 (62)	2.55	1.46	1.14	0.26	3.69*	112
2	1.99 (25)	1.96 (54)	3.37	2.80	1.42	0.84	5.23*	77
3	2.82 (22)	2.76 (56)	4.85	4.04	2.03	1.26	4.64*	76
4	2.26 (58)	2.45 (77)	3.36	3.17	1.10	0.69	2.63*	131
5	3.09 (83)	3.71 (134)	4.46	4.60	1.37	0.90	3.43*	215
6	4.82 (275)	4.36 (160)	6.54	5.48	1.72	1.13	5.18*	433

[a]Values in parentheses are numbers of students.

* $p < .01$.

We shall not discuss all the results here, but several points of interpretation are significant. At the end of the school year the investigators learned that at one control school in California teachers and administrators had added 25 minutes per day of classroom instruction and practice in arithmetic for grades 4 and 5. Data from this control school are responsible for the negative t value at grade 4 in Table 5. What is important is the demonstration that with a sufficiently intensive effort the effects of classroom drill by the teacher can be as effective as drill and practice on a computer. Drill and practice on the computer, however, took less time and did not require an additional effort from the teacher. A second point of interest is that the CAI results for Mississippi (Table 6) are substantially more impressive than those for California. This is an example of the generally noticed result that CAI drill and practice is more effective with students who start below grade level.[17]

A different approach tried in the New York City Schools is the Dial-A-Drill program in which students are called at home and given 5 minutes' practice in oral arithmetic problems. The oral exercises are generated from digitized word recordings stored on a computer disk, and the students respond by using a touch-tone dial. Students in grades 2-6 participated in the demonstration project. Except at the third-grade level, students received the program at most three days a week. An intensive program for third graders required their receiving 5 minutes of drill and practice six days a week. Because the project was supported by an Urban Education Grant, the students participating were mainly from disadvantaged environments.

Evaluation of the Dial-A-Drill is reported in Beech, McClelland, Horowitz, and Forlano [1970]. The results may be summarized briefly as follows. Experimental and control groups were both given the Metropolitan Achievement Test of Arithmetic Computation and a specially

---

[17] For further discussion of measurement methods and empirical results concerning inequality-reducing aspects of CAI see Jamison, Fletcher, Suppes, and Atkinson [1971] or Fletcher and Jamison [1973].

designed Oral Arithmetic Test in October 1969 and May 1970. A least-squares analysis of covariance of the 1970 arithmetic achievement data failed to produce statistically significant differences between the experimental and control students at any grade. Further analysis of the data showed that some students in the program did not actively participate.

A separate analysis was performed on students in the experimental group who had more than 32 sessions (approximately one per week), and those selected students were matched with control group students. Three tests for correlated means were performed and only third-grade students exhibited a statistically significant difference. This difference was on the arithmetic test, in which the experimental students performed better than the control students. One inference to be made from this study is that 15 minutes a week, that is, three sessions a week of 5 minutes each, are not sufficient to produce a measurable difference.

Beech et al. [1970] also investigated extensively the attitudes of parents and students to the program. The results are of some significance for two reasons. The terminals were located in the homes and not in the school, and the children were in all cases drawn from poverty areas. A survey of the attitudes of the parents toward this kind of program showed generally positive attitudes. The results of a questionnaire directed to the students also indicated a favorable response. While positive attitudinal responses to this experiment must be interpreted as preliminary, they do suggest that further research on bringing instruction into the home via telephone is worth investigating.

In another study relating to attitudes Smith and Hess [1972] examined non-cognitive effects of CAI in their research. The measures of student attitudes included the Sears Self-Concept Inventory, Coopersmith Self-Esteem Inventory, Crandall Locus of Control Instrument and items from the Coleman report. All of these measurement instruments are based on student responses and question the student's attitudes relating to general control over environment, responsibility for mathematics failures and successes, aptitude in mathematics and

social relations. The sample used consisted of 159 students with CAI and 161 without in grades 7, 8 and 9. The students were using the mathematics strands drill-and-practice program. The general result was no difference in means between CAI and non-CAI groups and no difference in means for pre-test and post-test measurements of attitudes for the CAI group.

Two studies have related arithmetic achievement to amount of CAI, using regression models of the sort described in Section I. Wells, Whelchel, and Jamison [to appear] analyzed data for 446 fifth- and sixth-grade students. The analysis was done on an individual student basis and it was possible to match students with their teachers. Data were separated by grades of students and then stratified by sex on the assumption that a differential effect of school resource variables might be observed.

The dependent variable in the regression model was the score of the student on the mathematics portion of the California Test of Basic Skills at the end of the experimental year (MA). Independent variables included test score at the beginning of the year (MB), years of teacher experience (TEAEXP), score of teacher on a 100-point verbal test (TEAVER), teacher degree level (TEADEG), student self-efficacy (SELFEX), and the number of sessions of CAI for each student during the course of the year (CAISES). Students were not assigned a number of CAI sessions randomly and, with the exception of sixth grade boys, there was a slight positive correlation between MB and CAISES. The test scores were measured in grade equivalents. Multicollinearity was a problem only with the teacher characteristic variables and separate models were specified for each of the teacher variables. Both linear and Cobb-Douglas (log-log) models were tested. The equation reported below for fifth-grade males with CAI is representative of the results obtained for the various stratifications of students. The t values are in parentheses.

Fifth-grade males with CAI, Cobb-Douglas model (all variables are in logs):

$$MA = .1408 + \underset{(13.08)}{.8052}\ MB + \underset{(3.60)}{.0572}\ CAISES - \underset{(1.61)}{.0643}\ SELFEX + \underset{(1.81)}{.0195}\ TEAEXP,$$

$$R^2 = .7427.$$

The Cobb-Douglas model measures output elasticity; a 1% increase in the number of CAI sessions, for example, would result in a 5.72% increase in mathematics achievement. For this model the gain from 100 sessions would depend on the initial level of achievement. Among students who had some CAI in this sample the average number of CAI sessions ranged from 59.25 for sixth grade girls (with a standard deviation of 36.44) to 85.00 for fifth grade girls (with a standard deviation of 37.68). The number of CAI sessions were statistically significant in both models for fifth-grade males and the Cobb-Douglas models for fifth-grade females; they were statistically insignificant in the linear models for fifth grade-females and in both models for sixth-grade students.

Suppes, Fletcher, Zanotti, Lorton, and Searle [to appear] reported a 1971-72 study dealing with the effects of drill and practice in elementary mathematics on elementary-school children in residential schools or day classes for 312 deaf students. The number of sessions students were to receive in a 5-month period was randomly assigned, eliminating multicollinearity problems. A number of different models in addition to the linear regression model were tested. Application of the linear model yielded the following regression equation:

$$E(T_{i2}) = 1.116 + .793\ T_{i1} + .084\ N_i\ ,$$

where T_{i1} is the pretest score of student i on a modified Stanford Achievement Test, T_{i2} is the posttest score on a second form of the same test, and N_i is the number of CAI sessions of student i divided by 10. The multiple correlation obtained was .811. It should be noted that if linearity held in 150 CAI sessions, a gain of $15 \times .084 = 1.26$ grade-placement years would be expected. In fact, in the experiment the average number of sessions of the group with the most intensive treatment was 75.84 with a standard deviation of 29.15, and linearity held over this range reasonably well.

An extensive analysis of detailed student learning and performance data on elementary-school CAI mathematics may be found in a recent

book by Suppes and Morningstar [1972]. The results are too extensive to survey here. Much of the volume is devoted to analysis of the sort of micromodels we have excluded from consideration in this review.

In the case of beginning reading, a number of CAI studies have been reported by Atkinson and his collaborators. Results of a tutorial reading program in 1966-67 in which students were given approximately 20 minutes a day on terminals are reported in Atkinson [1968]. The results of this experiment are interesting, because while the experimental group received tutorial reading via CAI the control group in this study received tutorial mathematics via CAI; therefore, both groups were being exposed to CAI. The experimental and control groups had similar characteristics; they constituted the approximately 100 students in the first grade in the school in which the experiment was conducted (approximately 100 because the number enrolled varied slightly during the school year). The posttest results for the experimental and control groups on the California Achievement Test and the Hartley Reading Test are shown in Table 7. As can be seen, most of the results are favorable

---

Insert Table 7 about here

---

to the experimental group. The computer system used in this experiment was an expensive one, for the student stations had not only a cathode-ray tube and keyboard terminal, but also an audio and visual display unit as well as a light pen for the cathode-ray tube.

Evaluation of a recent CAI program in initial reading using only teletype terminals and audio (but computer-generated audio) is reported in Fletcher and Atkinson [1972]. The curriculum was conceived as supplementary drill-and-practice rather than tutorial. The efficacy of the program was tested by using a group of 50 matched pairs of students. Prior to receiving exposure to CAI, 25 pairs of first-grade boys and 25 pairs of first-grade girls were matched on the basis of the Metropolitan Readiness Test, which was administered in November 1969. Three posttests

TABLE 7

Posttest Results for Experimental and Control Groups
(Atkinson [1968])

Test	Experimental group	Control group	p value
California Achievement Test			
Vocabulary	45.91	38.10	< .01
Comprehension	41.45	40.62	-
Total	45.63	39.61	< .01
Hartley Reading Test			
Form class	11.22	9.00	< .05
Vocabulary	19.38	17.05	< .01
Phonetic discrimination	30.88	25.15	< .01
Pronunciation			
Nonsense word	6.03	2.30	< .01
Word	9.95	5.95	< .01
Recognition			
Nonsense word	18.43	15.25	< .01
Word	20.61	16.60	< .01

were administered in late May and early June 1970. Four subtests of the Stanford Achievement Test were used. The California Cooperative Primary Reading Test, as well as a special test developed at Stanford, were also administered. The average grade placement on the Stanford Achievement Test and on the California Cooperative Primary Test, which were used as posttests, is shown in Table 8. While the results are

---

Insert Table 8 about here

---

significant in favor of the CAI groups, what is especially interesting is the unusually good performance of the boys. Similar results where boys did about as well as girls in a CAI reading environment were also reported in Atkinson [1968]. These results are contrary to those ordinarily obtained in TI for initial reading performance of boys and girls.

College Level CAI

A variety of evaluations have been conducted at the college level, mainly in connection with courses operated as part of research and development projects in CAI. Although it is not possible to give a complete summary here, major efforts made at a number of institutions are summarized and include studies conducted at Florida State University, the State University of New York at Stony Brook, University of Illinois, University of Texas, and Stanford University.

Hansen, Dick, and Lippert [1968] of Florida State University reported results of implementing collegiate instruction in physics by means of CAI, that is, problem sessions were handled in a CAI environment. In the fall of 1967 three groups of students were compared: (a) students receiving the bulk of instruction by CAI, (b) students receiving partial CAI and partial TI, and (c) students receiving only TI. Correlated t

TABLE 8

Average Grade Placement on the Stanford Achievement Test (SAT)

and the California Cooperative Primary Test (COOP)

[Fletcher & Atkinson, 1972]

		SAT	COOP
Boys	CAI	2.2	2.5
	non-CAI	1.8	1.8
Girls	CAI	2.4	2.6
	non-CAI	2.0	2.2

tests on the sum of the midterm and final examination scores showed that the autonomous CAI group was statistically superior to the other groups, but the difference between students who received partial CAI and students who received only TI was not significant.

In the spring of 1970 three more groups of students were studied. One was an autonomous CAI group, a second was a group of students receiving TI only, and a third was a group of students receiving TI plus a 4-hour-examination review on a computer system. The mean scores for the midterm examination, the final examination, and final grade showed no significant differences among the three treatment groups. The effect of CAI seemed to truncate the distribution of lower grades. The investigators applied a Kolmogorov-Smirnov test to the distribution of scores from the CAI group and the TI group receiving the review on the computer system. The results of this test indicated that the two distributions were significantly different ($D = 8.48$; $p \leq .05$), with the CAI group receiving fewer low grades than the TI group.

The CAI Center at Florida State University has also conducted several studies on computer-managed instruction (CMI). CMI differs from CAI in that students do not interact on line with the computer system, but rather they receive from the computer program directions of what unit to do next, possibly diagnostic testing, and remedial information, if necessary. Hagerty [1970] reported the results of a CMI course in techniques of PI conducted in the fall of 1969 with 59 graduate students. Students worked at their own pace by scheduling time on the computer terminal as needed. The results indicated that the CMI students performed as well as students taught the previous year by conventional classroom lecture methods. What is interesting about this study, which did not produce a significant difference in the two groups, was that the costs of operating the course for terminal time and personnel were $3,074, which is lower than the costs for TI.

Lawler [1971] investigated the differential effects of instructional strategies in CMI, using 167 undergraduates in a health education course at Florida State University. Forty-one of the students received TI; the remaining students were randomly assigned to one of three CMI

treatments. The three CMI treatment groups were varied in their pace through the course. The results showed superiority of the CMI groups over the TI group on final examination performances. Again, the results of the extensive analysis of variance are too detailed to summarize h. e, but the general conclusion just stated is supported by extensive statistical analysis. Concerning the different CMI treatments, there seems to be some advantage to requiring students to reach mastery at each stage or level of the course.[18]

Adams [1969] and Morrison and Adams [1969] described results of experiments conducted over two years at the State University of New York, Stony Brook. The subjects were students in introductory German, and both CAI and control groups received 3 hours of instruction per week in regular classes. The control group received, in addition to class time, the standard 1 hour per week of language laboratory; the CAI group received instead 1 hour per week of CAI in reading and writing. At very slight (if any) sacrifice to their performance in listening and speaking, the CAI students performed substantially better than the control students on tests of reading and writing achievement. The CAI and control groups were well matched on the Modern Language Aptitude Test. The experimenters reported a generally favorable student attitude to CAI.

Using the PLATO system at the University of Illinois, Grandey [in press] studied the use of computers to aid instruction in beginning chemistry. Thirty-one students with weak high school chemistry backgrounds were exposed to varying amounts of material presented by PLATO. Comparisons were made between 18 students who used PLATO for instruction and 13 students who used PLATO only for review. No significant differences between the two groups were established. Axeen [1967] studied CAI in the use of the library by undergraduates. Comparing 32 students who received a sequence of PLATO lessons, which entirely replaced an introductory college course on library use, with 34 students taught by

---

[18] Baker [1971] reviewed five additional CMI projects but reported no data on the instructional effectiveness of any of them. Kelley [1972], in a later paper, did present detailed results on the effectiveness of one of the projects described by Baker.

ordinary TI, Axeen found that students in the experimental class took less time to learn the same material. However, no significant differences were noted in performance as measured by the Library Orientation Test for college freshmen.

Bitzer and Boudreaux [1969] used the PLATO system for a CAI course in nursing. One hundred forty-four nursing students spent an average of 20 hours in a course on maternity nursing and 18 hours in a course on pharmacology for nurses. Perhaps the most striking result obtained in this study was the savings in time. All members of one group of 38 nurses for which complete time measures were kept finished the maternity nursing material in 50 hours or less at terminals. The same material required 84 hours of standard lecture presentation.

Coombs and Peters [1971] used the PLATO system to study CAI in role-playing games. One hundred six students in an introductory American government course spent 18 class hours at terminals. Comparisons were made with a like number of students who received TI during 18 hours in small-group discussions. No significant differences in the two groups were obtained.[19]

At the University of Texas a number of experiments and demonstration projects have been completed in the Computer Assisted Instruction Laboratory. Castleberry and Lagowski [1970] reported on a CAI chemistry course. Fifteen CAI modules were developed as supplementary material for the introductory course in general chemistry. The following results were obtained. In both semesters during the academic year 1968-69, students who took advantage of the available CAI modules scored significantly higher than the control group on the parts of the final examination covered by the modules. In addition, during the first semester, students using the CAI modules also scored significantly higher than the control group on the parts of the final examination not covered by

---

[19] The PLATO system is currently being expanded and a detailed evaluation of that expanded system, as well as a system being constructed by the MITRE Corporation, will be conducted over the next four years. Anastasio [1972] described the plans for the evaluation and Lyman [1972] provided a listing of previous PLATO research and curriculum efforts.

the modules. These two results alone suggest that a selective process was at work rather than any instructional advantages of the CAI modules, since students used the CAI modules on a voluntary basis. However, during the second semester, no significant differences were observed between the experimental and control groups on the final examination in the material not covered by the modules. The results taken together support the hypothesis that the CAI modules were a useful addition to the course. Evidence is presented in Table 9 that the voluntary experimental group

---

Insert Table 9 about here

---

using the CAI modules was not necessarily more able than the control group. For example, the Stanford Achievement Test scores in both semesters were no better for the experimental group than for the control group; if anything, they were perhaps slightly lower. The same is true of the results on the chemistry placement score.

Judd, Bunderson, and Bessent [1970] investigated the effects of learner control in a CAI course in precalculus mathematics. They did not compare the student performance with TI classes, but they did compare learner-controlled strategies with program-controlled strategies. No striking differences were found. One general conclusion did emerge from their analysis. Student control of progress through a course seems to be successful in subjects in which the student has competence and is definitely less successful when the student's competence is low, or he has little familiarity with the course material on the basis of past experience. This study illustrates how difficult it is to obtain strong conclusions about how learner control should be built into CAI courses. As in other areas of research on the effectiveness of instructional methods, interaction between the cognitive and affective states of the student and the structure of instruction will certainly be a major focus of investigations in the next few years.

TABLE 9

Group Means [Castleberry & Lagowski, 1970]

Variable	Experimental group	Control group	Experimental dropouts
First Semester			
Final Exam Score:			
Items related to modules	86.7	68.7	74.9
Final Exam Score:			
Items not related to module	83.5	74.1	75.7
SAT-Math	518	530	480
SAT-Verbal	572	545	542
Chemistry placement score[a]	19.0	19.7	17.2
Second Semester			
Final Exam Score:			
Items related to modules	81.3	71.8	76.8
Final Exam Score:			
Items not related to modules	42.6	42.6	42.6
SAT-Math	480	515	518
SAT-Verbal	517	537	537
Chemistry placement score[a]	15.5	20.7	16.4

[a]The Chemistry Placement Examination is required of all students before they register in general chemistry. The maximum score on this examination is 50.

Homeyer [1970] reported the results of comparing a CAI with a TI course in computer programming. The students taking either the CAI version or the lecture version were required to have had at least one previous course in computer science and some experience in programming. Two sections of a course were used: one as a CAI group and the other as a lecture group. There were ten students in each group. Although the number of students was small, the study explicitly tested the following hypotheses:

H1. The CAI group can complete course instruction significantly faster than the lecture group. This hypothesis was accepted; the CAI group completed course instruction about twice as fast in terms of number of hours (an average of 13.75 hours for the CAI group compared with 24 hours for the lecture group).

H2. The CAI group makes significantly fewer personal visits to the instructor. This hypothesis was rejected. Both groups made about the same number of personal visits to the office of the instructor.

H3. There is no significant difference between the CAI and lecture groups with respect to mean scores on examinations. This hypothesis was accepted. The performance of the students was not significantly different at the .05 level.

H4. There is no significant difference between CAI and lecture groups with respect to mean grades on computer programs written. This hypothesis also was accepted, with about equal performance from the two groups.

Edwards and Judd [1972] reported on the evaluation of a course in special education for undergraduates at the University of Texas. Students in the course were assigned to one of three groups. One group received a course handbook and participated in a discussion section; the second group joined only a discussion section; and a third group received the handbook and CAI, but did not participate in a discussion section. The results were somewhat mixed, but the evidence favored the test performance of the group receiving CAI rather than discussion sections, which indicates that in this kind of course, CAI can successfully replace small group sections of large lecture courses.

At Stanford University, Joseph Van Campen has developed a full two-year tutorial course in Introductory Russian. This project, begun in 1967, teaches the standard aspects of a first-year course at the college level, that is to say, comprehension of written Russian, comprehension of spoken Russian, and mastery of grammar and syntax. Of the three main components of a college-level language course -- regular classroom sessions on a daily basis, time spent in the language laboratory, and regular homework assignments -- only the functions of the tutorial classroom sessions are assumed by the CAI course. In addition to their time at computer consoles, students spend time in the language laboratory and do off-line homework assignments. What is important about this example is that the regular 5 hours a week of classroom instruction were completely replaced by daily work for a comparable time at computer terminals.

An evaluation of the course for 1968-69 is presented in Suppes and Morningstar [1969]. First of all, the CAI course showed superior holding power in comparison with TI. Of the 30 students originally enrolled in the CAI course, 73 percent finished all three quarters of the first year, whereas of the 38 students in the two regular classes only 32 percent finished the year's curriculum. Approximately 66 percent of the content of the final examinations for the autumn and winter quarters were identical for the CAI and regular Russian courses; the final examination for the spring quarter was identical for the two groups. The average number of errors was lower for the CAI students in all three quarters and was statistically significant for the fall quarter (Mann-Whitney U test, $p < .001$) and the spring quarter ($p < .05$), but not for the winter quarter. Since the selection process resulting from more of the poorer students' leaving the regular course biased results against the CAI group, the superiority of the CAI group on the spring examination is more impressive than the statistical analysis indicates.

Conclusion

As in other methods of instruction surveyed in this report, no simple uniform conclusions can be drawn about the effectiveness of CAI. At the elementary-school level, CAI is apparently effective as a supplement to regular instruction. What we do not have are the sorts of experiments required for a complete productivity analysis. There are no examples yet of CAI's being introduced with a concomitant change in student-teacher ratio, which would, for example, cover the costs of CAI. At the present time, we can only conclude that CAI can be used in some situations to improve achievement scores, particularly for disadvantaged students.

At the secondary school and college levels, a conservative conclusion is that CAI is about as effective as TI when it is used as a replacement. It may also result in substantial savings of student time in some cases. Since the equal-effectiveness conclusion seems to be broadly correct for most alternative methods of instruction at the college level, there should be in the future increasing opportunities to experiment with selecting the method of instruction in terms of costs, and real opportunities should exist for substituting capital for labor, especially as the relative costs of technology in comparison to labor decline over the next decade.

VII. CONCLUSION

In this paper we have surveyed research on the effectiveness of traditional instruction (TI), instructional radio (IR), instructional television (ITV), programmed instruction (PI), and computer-assisted instruction (CAI). Students learn effectively from all these media, and relatively few studies indicate a significant difference in one medium over another or of one variant of a medium over another. The studies taken together suggest that alternative methods of TI are approximately equally effective, although several studies indicated that different variables are significantly correlated with student achievement. Teacher verbal ability was important in many of these studies, and the evidence suggests that smaller class size may consistently, if slightly, improve the performance of primary-grade students.

Though there is a substantial past history in the use of IR, few studies of its effectiveness exist. A number that do exist were, however, carefully done and they indicate that IR, supplemented with appropriate printed material, is about as effective as TI. There is much more extensive research literature on the effectiveness of ITV, and excellent surveys of that literature already exist. There is strong evidence that ITV, used in a way that closely simulates TI, is as effective, on the average, as TI for all grade levels and subject matters. There is very little evidence concerning the effectiveness of ITV used in ways that utilize the unique capabilities of the medium. A reasonable fraction of the student and teacher populations has a somewhat unfavorable attitude toward ITV, although the incidence of unfavorable attitudes tends to diminish as institutions gain experience with the medium. After such experience a majority of students have neutral or favorable attitudes toward ITV.

Both PI and CAI attempt to improve the quality of instruction by providing for its individualization along one or more dimensions. Nonetheless, findings of "no significant difference" dominate the research literature in this area. Though there are often no significant

differences in achievement some of the studies do report a saving in student time, and this is an index of success. When small amounts of CAI are used as a supplement to regular classroom instruction (as with the elementary-school drill-and-practice programs) substantial evidence suggests that it leads to an improvement in achievement, particularly for slower students. Models exist that relate the amount of achievement gain to the number of CAI sessions a student receives.

In broad terms, the many studies we have surveyed suggest that the costs of alternative technologies, with capital investment amortized over an appropriate number of years, should always be given serious consideration in planning an educational program or evaluating proposed changes in current programs. On the other hand, there are enough differences in the studies in terms of achievement measures to suggest that a policy of strict minimization of costs in the choice of a technology for teaching is too simple a criterion. At least four considerations will probably be of importance in the future. Each will need more extensive study:

First, we must examine if the savings in time exhibited in some of the studies using PI or CAI can be shown to be significant over longer periods and for a higher percentage of the total instructional program of students.

Second, we do not yet have an appropriately detailed evaluation of the impact of the various technologies on the long-term motivation of students.

Third, the long-term effects of individualization and privacy of learning characteristic of some of the technologies also needs more extensive evaluation. We do not know, for example, whether students who are given highly individualized programs in the elementary school for most of their instruction will strongly prefer the continuation of such methods in secondary school and college or whether they will desire to return as they grow older to more traditional forms of instruction.

Fourth, it has been indicated at a number of points in this review that most evaluations, particularly those considered well controlled, compare TI to a form of IR, ITV, or CAI that closely emulates

the TI. It is at least plausible that many of the conclusions of this survey would be overturned were more imaginative uses of the media explored, that yet permitted comparative evaluation.

Most of the educational technologies we have surveyed in this article have a relatively recent history. Even though there is already a fairly extensive literature on their evaluation, it would be a mistake to view the present state of that literature as anything but preliminary in nature. It will be many years before we have an adequately deep assessment of the strengths and weaknesses of the technological alternatives to traditional instruction that have been considered in this survey.

REFERENCES

Adams, E. Field evaluation of the German CAI Lab. In R. C. Atkinson & H. Wilson (Eds.), Computer-assisted instruction: A book of readings. New York: Academic Press, 1969.

Adelman, I., & Parti, M. The determinants of student achievement: A simultaneous equation approach. Unpublished manuscript.

Allen, W. H. Audio-visual communication. In C. W. Harris (Ed.), Encyclopedia of educational research. (3rd ed.) New York: Macmillan, 1960.

Alton, E. V. An experiment using programed material in teaching a noncredit algebra course at the college level (with supplement). (Doctoral dissertation, University of Michigan) Dissertation Abstracts, 1966, 126, 4488.

Anastasio, E. Evaluation of the PLATO and TICCIT computer-based instructional systems--A preliminary plan. Princeton, New Jersey: Educational Testing Service, 1972.

Anderson, B., & Greenberg, E., with McClung, R. Educational production functions for teacher technology mixes: Problems and possibilities. Memorandum 72/2, 1972, Washington University.

Anderson, F. H., Bedford, F., Clark, V., & Schipper, J. A report on an experiment at Camelback High School. Mathematics Teacher, 1963, 56, 155-159.

Anderson, K. E. The relationship between teacher load and student achievement. School Science and Mathematics, 1950, 50, 468-470.

Ashford, T. H. Some long-range effects of programmed instruction in music. Journal of Research in Music Education, 1968, 16, 339-344.

Atkinson, C. Public school broadcasting to the classroom. Boston: Meador, 1942. (a)

Atkinson, C. Broadcasting to the classroom by universities and colleges. Boston: Meador, 1942. (b)

Atkinson, R. C. Computerized instruction and the learning process. American Psychologist, 1968, 23, 225-239.

Attiyeh, R., Bach, G., & Lumsden, K. The efficiency of programmed learning in teaching economics: The results of a nationwide experiment. American Economic Review, 1969, 59, 217-224.

Attiyeh, R., & Lumsden, K. G. Some modern myths in teaching economics: The U.K. experience. American Economic Review, 1972, 62, 429-433.

Averch, H., Carroll, S., Donaldson, T., Kiesling, H., & Pincus, J. How effective is schooling? A critical review and synthesis of research findings. (R-956-PCSF/RC) Santa Monica, Calif.: Rand Corp., 1972.

Axeen, M. E. Teaching the use of the library to undergraduates: An experimental comparison of computer-based instruction and the conventional lecture method. Report No. R-361, 1967, University of Illinois, Coordinated Science Laboratory.

Baker, F. B. Computer-based instructional management systems: A first look. Review of Educational Research, 1971, 41, 51-70.

Baldwin, L. V., Davis, P., & Maxwell, L. M. Innovative off-campus educational programs of Colorado State University. Special report to the President's Science Advisory Committee Panel on Educational Research and Development. April 1972.

Balow, I. H. Longitudinal evaluation of reading achievement. Elementary English, 1969, 46, 184-187.

Barrow, L. C., & Westley, B. H. Comparative teaching effectiveness of radio and television. AV Communication Review, 1959, 7, 14-23.

Barrow, L. C., & Westley, B. H. Exploring the news: An experiment on the relative effectiveness of radio and TV versions of a children's news program. In W. Schramm (Ed.), The impact of educational television. Urbana: University of Illinois, 1960.

Beech, R. P., McClelland, S. D., Horowitz, G. R., & Forlano, G. Final report: An evaluation of the Dial-a-Drill program. New York: State Education Department, 1970.

Benson, C. S., Schmelze, W. K., Gustafson, R. H., & Lange, R. A. State and local fiscal relationships in public education in California. Report of the Senate Fact Finding Committee on Revenue and Taxation. Sacramento: Senate of the State of California, 1965.

Bereday, G., & Lauwerys, J. A. (Eds.) The year book of education 1960. London: Evans, 1960.

Bitzer, M., & Boudreaux, M. Using a computer to teach nursing. Nursing Forum, 1969, 8.

Blake, H. J. A summary of selected studies in elementary and secondary schools. Unpublished doctoral dissertation, Teachers College, Columbia University, 1954.

Bobier, D. T. The effectiveness of the independent use of programmed textbooks in aiding students to overcome skill weaknesses in English, mechanics, and arithmetic. (Doctoral dissertation, University of Michigan) Dissertation Abstracts, 1965, 25, 3424.

Bogatz, G. A., & Ball, S. The second year of Sesame Street: A continuing evaluation. Princeton, N. J.: Educational Testing Service, November 1971.

Bowles, S. S. Towards an educational production function. In W. L. Hansen (Ed.), Education, income and human capital. New York: National Bureau of Economic Research, 1970.

Bowles, S. S., & Levin, H. M. The determinants of scholastic achievement--An appraisal of some recent evidence. Journal of Human Resources, 1968, 3, 3-24. (a)

Bowles, S. S., & Levin, H. M. More on multicollinearity and the effectiveness of schools. Journal of Human Resources, 1968, 3, 393-400. (b)

Brewer, L. Radio as an aid to instruction in elementary science. Science Education, 1939, 23, 63-68.

Brigham, W. C. A comparison of two response modes in learning woodwind fingering by programmed text. Journal of Research in Music Education, 1970, 18, 49-56.

Bull, C. R. Australia. In G. Bereday & J. Lauwerys (Eds.), The yearbook of education 1960--Communication media and the school. London: Evans, 1960.

Bullmer, K. Improving accuracy of interpersonal perception through a direct teaching method. Journal of Counseling Psychology, 1972, 19, 37-41.

Burkhead, J., Fox, T. G., & Holland, J. W. Input and output in large-city high schools. Syracuse, N. Y.: Syracuse University Press, 1967.

Cannon, G. M. Kindergarten class size--A study. Childhood Education, 1966, 43, 9-11.

Carnoy, M. A systems approach to evaluating education. Stanford School of Education, 1971. (Mimeographed)

Carpenter, H. A. Teaching science by radio. Junior-Senior High School Clearing House, March 1934.

Carpenter, H. A. Science in the Rochester School of the Air. Science Education, April 1937, pp. 79 ff.

Castleberry, S., & Lagowski, J. J. Individualized instruction using computer techniques. Journal of Chemical Education, 1970, 47, 91-96.

Chang, C. Educational broadcasting in China. In C. S. Marsh (Ed.), Educational broadcasting 1936. Chicago, Ill.: University of Chicago Press, 1937. Pp. 183-185.

Chu, G. C., & Schramm, W. Learning from television: What the research says. Washington, D. C.: National Association of Educational Broadcasters, 1967.

Cohn, E. Economies of scale in Iowa high school operations. Journal of Human Resources, 1968, 3, 422-434.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. Equality of educational opportunity. Washington, D. C.: U. S. Office of Education, 1966.

Cook, D. C., & Nemzek, C. L. The effectiveness of teaching by radio. Journal of Educational Research, 1939, 33, 105-109.

Coombs, F. S., & Peters, J. G. Student explorations of political processes: The use of computer-based education. Proceedings of the Tenth National Gaming Symposium, 1971.

Costin, F., Greenough, W. T., Menges, R. J. Student ratings of college teaching: Reliability validity, and usefulness. Review of Educational Research, 1971, 41, 511-533.

Counelis, J. S. First grade students in the Hunters Point-Bayview SEED project: A diagnostic review. San Francisco: University of San Francisco, 1970. ERIC: ED 052 905.

De Cecco, J. P. Class size and co-ordinated instruction. British Journal of Educational Psychology, 1964, 34, 65-74. (a)

De Cecco, J. P. (Ed.) Educational technology: Readings in programmed instruction. New York: Holt, Rinehart, & Winston, 1964. (b)

Doty, B. A., & Doty, L. A. Programmed instructional effectiveness in relation to certain characteristics. Journal of Educational Psychology, 1964, 55, 334-338.

Dubin, R., & Hedley, R. A. The medium may be related to the message: College instruction by TV. Eugene: University of Oregon Press, 1969.

Dubin, R., & Taveggia, T. C. The teaching-learning paradox: A comparative analysis of college teaching methods. Eugene: Center for the Advanced Study of Educational Administration, University of Oregon, 1968. ERIC: ED 026 966.

Edwards, A., & Judd, W. A. Evaluation of CARE 1 for projected use at the University of Texas. Tech. Rept. No. 12, 1972, The University of Texas at Austin, Computer Assisted Instruction Laboratory.

Edward W. Clark High School. The influence of class size on academic attainment and student satisfaction. Las Vegas: Edward W. Clark High School, 1968. ERIC: ED 021 252.

Fleming, C. M. Class size as a variable in the teaching situation. Educational Research, 1959, 1, 35-48.

Fletcher, J. D., & Atkinson, R. C. An evaluation of the Stanford CAI program in initial reading (grades K through 3). Journal of Educational Psychology, 1972, 63, 597-602.

Fletcher, J. D., & Jamison, D. Evaluation of equality of educational output: Computer-assisted instruction as an example. Presented at the annual meeting of the American Educational Research Association, New Orleans, February 1973.

Forsythe, R. O. Instructional radio: A position paper. Unpublished paper. Stanford University, 1970. ERIC: ED 044 933.

Frymier, J. R. The effect of class size upon reading achievement in the first grade. The Reading Teacher, 1964, 18, 90-93.

Furno, O. F., & Collins, G. J. Class size and pupil learning. Baltimore, Md.: Baltimore City Schools, 1967. ERIC: ED 025 003.

Gage, N. L. Teacher effectiveness and teacher education: The search for a scientific basis. Palo Alto, Calif.: Pacific Books, 1972.

Getzels, J., & Jackson, P. Creativity and intelligence explorations with gifted students. New York: Wiley, 1962.

Glaser, R. (Ed.) Teaching machines and programmed learning II: Data and directions. Washington, D. C.: National Education Association, 1965.

Grandey, R. C. The use of computers to aid instruction in beginning chemistry. Journal of Chemical Education, in press.

Green, R., et al. The educational status of children in a district without public schools. 1964, Michigan State University, College of Education.

Greenhill, L. P., Carpenter, C. R., & Ray, W. S. Further studies of the use of television for university teaching.

Guthrie, J. W. A survey of school effectiveness studies. In A. M. Mood (Ed.), Do teachers make a difference? Washington, D. C.: U. S. Office of Education, 1970.

Guthrie, J. W., Kleindorfer, G. B., Levin, H. M., & Stout, R. T. Schools and inequality. Cambridge, Mass.: Massachusetts Institute of Technology Press, 1971.

Hagerty, N. K. Development and implementation of a computer-managed instruction system in graduate training. Tech. Rept. No. 11, 1970, Florida State University, Computer-Assisted Instruction Center.

Hansen, D. N., Dick, W., & Lippert, H. T. Research and implementation of collegiate instruction of physics via computer-assisted instruction, Vol. 1. Tech. Rept. No. 3, 1968, Florida State University, Computer-Assisted Instruction Center.

Hanushek, E. A. The production of education, teacher quality and efficiency. In A. M. Mood (Ed.), Do teachers make a difference? Washington, D. C.: Office of Education, 1970.

Hanushek, E. A. Education and race: An analysis of the educational production process. Lexington, Mass.: D. C. Heath, 1972.

Harrison, M. Measures of the effect of radio programs in rural schools. In Education on the air. Columbus: Ohio State University, 1932.

Haskell, S. Some observations on the effects of class size upon pupil achievement in geometrical drawing. Journal of Educational Research, 1964, 58, 27-30.

Hatano, K. Japan. In G. Bereday & J. Lauwerys (Eds.), The yearbook of education 1960--Communication media and the school. London: Evans, 1960.

Heron, W. T., & Ziebarth, E. W. A preliminary experimental comparison of radio and classroom lectures. Speech Monographs, 1946, 13, 54-57.

Homeyer, F. C. Development and evaluation of an automated assembly language teacher. Tech. Rept. No. 3, 1970, The University of Texas at Austin, Computer Assisted Instruction Laboratory.

Hopper, H. H., & Keller, H. Writing skills--Are large classes conducive to effective learning. Ft. Pierce, Fla.: Indian River Junior College, 1966. ERIC: ED 012 583.

Husén, T. (Ed.) International study of achievement in mathematics: A comparison of twelve countries. New York: Wiley, 1967. 2 vols.

Jamison, D., Fletcher, J. D., Suppes, P., & Atkinson, R. C. Cost and performance of computer-assisted instruction for compensatory education. Research Paper 20, Graduate School of Business, and Tech. Rept. No. 175, Institute for Mathematical Studies in the Social Sciences, 1971, Stanford University. In R. Radner & J. Froomkin (Eds.), Education as an industry. New York: Columbia University Press, in press.

Jencks, C., with Smith, M., Acland, H., Bane, M., Cohen, D., Gintis, H., Heyns, B., & Michelson, J. Inequality: A reassessment of the effect of family and schooling in America. New York: Basic Books, Inc., 1972.

Johnson, D. C. Programed learning: A comparison of the School Mathematics Study Group programed and conventional textbooks in elementary algebra. (Doctoral dissertation, University of Michigan) Dissertation Abstracts, 1966, 26, 5294.

Johnson, M., & Scriven, E. Class size and achievement gains in seventh- and eighth-grade English and mathematics. The School Review, 1967, 75, 300-310.

Johnson, R. H., & Lobb, M. D. Jefferson Co. Colo. completes three-year study of staffing, changing class size, programming and scheduling. Bulletin of the National Association of Secondary School Principals, 1966, 45, 57-77.

Judd, W. A., Bunderson, C. V., & Bessent, E. W. An investigation of the effects of learner control in computer-assisted instruction prerequisite mathematics. MATHS. Tech. Rept. No. 5, 1970, The University of Texas at Austin, Computer Assisted Instruction Laboratory.

Katzman, M. The political economy of urban schools. Cambridge, Mass.: Harvard University Press, 1971.

Kelley, A. TIPS and technical change in classroom instruction. American Economic Review, 1972, 62, 422-428.

Kiesling, H. J. Measuring a local government service: A study of school districts in New York State. Review of Economics and Statistics, 1967, 49, 356-367.

Kiesling, H. J. The relationship of school inputs to public school performance in New York State. (P-4211) Santa Monica, Calif.: Rand Corp., 1969.

Kimble, D. Physiological psychology. Palo Alto, Calif.: Addison-Wesley, 1963.

Kinane, K. Australia's correspondence schools with support broadcast programmes and Radio University. In New educational media in action: Case studies for planners--I. Paris: UNESCO, International Institute for Educational Planning, 1967. Pp. 169-203.

Kottmeyer, W. Radio in language growth for study and discussion. St. Louis, Mo.: St. Louis Public Schools Vocabulary Development Project, 1970. ERIC: ED 046 223.

Lange, P. C. Today's education. National Education Association, 1972, 61, 59.

Lawler, R. M. An investigation of selected instructional strategies in an undergraduate computer-managed instruction course. Tech. Rept. No. 19, 1971, Florida State University, Computer-Assisted Instruction Center.

Leslie, P. J. Instructional radio for developing countries. Unpublished paper, Stanford University, 1971.

Levin, H. M. A new model of school effectiveness. In A. M. Mood (Ed.), Do teachers make a difference? Washington, D. C.: Office of Education, 1970.

Little, C. E. An experimental study of programed instruction in college algebra at Colorado State College. (Doctoral dissertation, University of Michigan) Dissertation Abstracts, 1967, 27, 2948-A.

Lumley, F. H. Rates of speech in radio speaking. Quarterly Journal of Speech, June 1933.

Lumsdaine, A. A., & Glaser, R. (Eds.) Teaching machines and programmed learning: A source book. Washington, D. C.: National Education Association, 1960.

Lumsdaine, A. A. Instruments and media of instruction. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Lyman, E. R. A summary of PLATO curriculum and research materials. Computer-based Education Research Laboratory, University of Illinois, Urbana, Ill., August 1972.

Madden, J. V. Experimental study of student achievement in relation to class size. School Science and Mathematics, 1968, 68, 619-622.

Madden, R. Educational radio bibliography 1954-67. Educational Broadcasting Review, 1968, 2, 66-79.

Mansfield, H., Jr. A report on experimentation in the teaching of the first course in algebra at El Camino College (CA). 1968. ERIC: ED 019 053.

Marklund, S. Scholastic attainment as related to size and homogeneity of class. Educational Research, 1963, 6, 63-67. (Abstract) The original work is: Skilklassens storlek och struktur. Stockholm: Almqvist & Wiksell, 1962.

Mayeske, G. W. Teacher attributes and school achievement. In A. M. Mood (Ed.), Do teachers make a difference? Washington, D. C.: Office of Education, 1970.

McAnany, E., Mayo, J., & Hornik, R. Television and educational reform in El Salvador. Research Report No. 4, 1970, Institute for Communication Research, Stanford University.

McLuhan, M. Understanding media: The extensions of man. New York: McGraw-Hill, 1964.

Meadowcraft, B. A. An experiment with programed materials in seventh grade arithmetic. (Doctoral dissertation, University of Michigan) Dissertation Abstracts, 1966, 26, 1514.

Menne, J. W., Klingensmith, J. E., & Nord, D. L. The feasibility of using taped lectures to replace class attendance. Paper presented at the American Educational Research Association Convention, Los Angeles, 1969. ERIC: ED 027 748.

Menniti, D. J. A study of the relationship between class size and pupil achievement in the Catholic elementary school. (Doctoral dissertation, Catholic University of America) Dissertation Abstracts, 1964, 25, 2854-2855.

Michelson, S. The existentialist reality of educational production functions (Their existence has preceded their essence). Washington, D. C.: Office of Education, 1970.

Miles, J. R. Radio and elementary science teaching. Journal of Applied Psychology, 1940, 24, 714-720.

Mood, A. M. Do teachers make a difference? In A. M. Mood (Ed.), Do teachers make a difference? Washington, D. C.: Office of Education, 1970.

Morrison, H., & Adams, E. Pilot study of a CAI laboratory in German. In R. C. Atkinson & H. Wilson (Eds.), Computer-assisted instruction: A book of readings. New York: Academic Press, 1969.

NHK Radio-Television Cultural Research Institute. The listening effects of radio English classroom, April 1954 - March 1955. Tokyo: NHK, 1955.

NHK Radio-Television Cultural Research Institute. The effects of educational radio music classroom, April-December 1956. Tokyo: NHK, 1956.

NHK Radio-Television Cultural Research Institute. Proceedings of the second international conference of broadcasting organizations on sound and television school broadcasting. Tokyo: NHK, 1964.

Olson, M. N. Ways to achieve quality in school classrooms: Some definitive answers. Research note. Phi Delta Kappan, 1971, 53, 63-65.

Peterson, P. A survey of studies in mastery learning. Unpublished paper, School of Education, Stanford University, 1972.

Popham, W. J. Tape recorded lectures in the college classroom. AV Communication Review, 1961, 9, 109-118.

Popham, W. J. Tape recorded lectures in the college classroom--II. AV Communication Review, 1962, 10, 94-101.

Raymond, R. Determinants of the quality of primary and secondary public education in West Virginia. Journal of Human Resources, 1968, 3, 450-470.

Rodin, M., & Rodin, B. Student evaluations of teachers. Science, 1972, 177, 1164-1166.

Rosenshine, B. Teaching behaviors and student achievement. Published for the International Association for the Evaluation of Educational Achievement by the National Foundation for Educational Research in England and Wales, 1971.

Saettler, P. H. A history of instructional technology. New York: McGraw-Hill, 1968.

Schramm, W. The research on programmed instruction: An annotated bibliography. Washington, D. C.: U. S. Department of Health, Education and Welfare, 1964.

Schramm, W. Educational radio in Thailand. In New educational media in action: Case studies for educational planners--I. Paris: UNESCO, International Institute for Educational Planning, 1967. Pp. 83-103.

Schramm, W. Notes on instructional cross-media comparisons. Working paper, 1971, Stanford University, Institute for Communication Research.

Schramm, W. Big media, little media. Unpublished manuscript, Stanford University, 1972. To appear.

Shrable, K., & Sassenrath, J. M. Effects of achievement, motivation and test anxiety on performance in programmed instruction. American Educational Research Journal, 1970, 7, 209-220.

Silberman, H. F. Characteristics of some recent studies of instructional methods. In J. E. Coulson (Ed.), Programmed learning and computer-based instruction. New York: Wiley, 1962.

Sitkei, E. G. The effects of class size: A review of the research. Research Series, 1967-1968. Los Angeles: Los Angeles County Superintendent of Schools, 1968. ERIC: ED 043 124.

Skornia, H. Educational radio: Its past and its future. In W. Schramm (Ed.), Educational television--The next ten years. Stanford: Stanford Institute for Communication Research, 1962.

Skornia, H. J. Television, radio, and other new media in education. IEEE International Conference on Communications Conference Record, 1968, 4, 854-858.

Smith, I. D., & Hess, R. D. The effects of computer-assisted instruction on student self-concept, locus of control, and level of aspiration. Research and Development Memorandum No. 89, 1972, Stanford University, Stanford Center for Research and Development in Teaching.

Smith, M. S. Equality of educational opportunity: The basic findings reconsidered. In F. Mosteller & D. P. Moynihan (Eds.), On equality of educational opportunity. New York: Vintage Books, 1972.

Sorensen, P. H., & Thomas, T. C. Evaluation of the compensatory education program of the San Francisco Unified School District, 1966-1967. Menlo Park, Calif.: Stanford Research Institute, 1967.

Stevenson, P. R. Smaller classes or larger: A study of the relation of class-size to the efficiency of teaching. Bloomington, Ill.: Public School Publishing Co., 1923.

Stickell, D. W. A critical review of the methodology and results of research comparing televised and face-to-face instruction. Unpublished doctoral dissertation, Pennsylvania State University, 1963.

Suchman, E. Evaluative research. New York: Russell Sage Foundation, 1967.

Suppes, P., Fletcher, J. D., Zanotti, M., Lorton, P., & Searle, B. Evaluation of computer-assisted instruction in elementary mathematics for hearing-impaired students, in press.

Suppes, P., & Morningstar, M. Computer-assisted instruction. Science, 1969, 166, 343-350.

Suppes, P., & Morningstar, M. Computer-assisted instruction at Stanford, 1966-68: Data, models, and evaluation of the arithmetic programs. New York: Academic Press, 1972.

Tanner, G. L. A comparative study of the efficacy of programed instruction with seventh grade low achievers in arithmetic. (Doctoral dissertation, University of Michigan) Dissertation Abstracts, 1966, 26, 6458.

Thomas, J. A. Efficiency in education: A study of the relationship between selected inputs and mean test scores in a sample of senior high schools. Unpublished doctoral dissertation, Stanford University, 1962.

Thomas, M. E. The efficacy of broadcasts to schools. Melbourne, Australia: Melbourne University Press, Educational Research Series, No. 48, 1937.

Tobias, S. Effect of creativity, response mode, and subject matter familiarity on achievement from programmed instruction. Journal of Educational Psychology, 1969, 60, 453-460.

Tobias, S., & Abramson, T. Interaction among anxiety, stress, response mode and familiarity of subject matter on achievement from programmed instruction. Journal of Educational Psychology, 1971, 62, 357-364.

United States Office of Education. Projections of educational statistics to 1979-80. Washington, D. C.: Author, 1970.

Varner, S. E. Class size. Washington, D. C.: National Education Association, 1968.

Vincent, W. S. Further clarification of the class size question. IAR Research Bulletin, 1968, 9, 1-3. ERIC: ED 034 303.

Vinsonhaler, J., & Bass, R. A summary of ten major studies of CAI drill and practice. Educational Technology, 1972, 12, 29-32.

Wells, S., Whelchel, B., & Jamison, D. The impact of varying levels of computer-assisted instruction on the academic performance of disadvantaged students. Stanford University: Graduate School of Business Research Paper, to appear.

Wiles, M. K. The evaluation of school news broadcasts. Unpublished doctoral dissertation, Ohio State University, 1940.

Williams, H. R., & Koelsche, C. L. Organization of chemistry classes: A comparative study of two patterns in the teaching of high school chemistry. Science Teacher, 1967, 34, 52-54.

Williams, J. G. Radio in fundamental education in undeveloped areas. Paris: UNESCO, 1950.

Williams, J. P. Comparisons of several response modes in a review program. Journal of Educational Psychology, 1963, 54, 253-260.

Williams, J. P. Effectiveness of constructed-response and multiple-choice programming modes as a function of test mode. Journal of Educational Psychology, 1965, 56, 111-117.

Wisconsin Research Project in School Broadcasting. Radio in the classroom. Madison: The University of Eisconsin Press, 1942.

Woelfel, N. E., & Tyler, I. K. (Eds.) Radio and the school: A guidebook for teachers and administrators. Yonkers-on-Hudson, N. Y.: World Book Co., 1945.

Wrightstone, J. W. Radio education. In W. S. Monroe (Ed.), Encyclopedia of educational research. New York: Macmillan, 1952. Pp. 953-961.

Zoll, E. J. Research in programed instruction in mathematics. The Mathematics Teacher, 1969, 62, 103-110.