DOCUMENT RESUME

ED 263 132 TM 850 602

AUTHOR Waxman, Hersholt C.; And Others

TITLE Adaptive Education and Student Outcomes: A Quantitative Synthesis. 1985/11.

INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.

SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE 85
NOTE 32p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Effect Size; Elementary Secondary Education; *Individualized Instruction; Learning; *Meta Analysis; *Outcomes of Education; *Research and Development Centers; Research Methodology; Student Characteristics
IDENTIFIERS Learning Research and Development Center; *Quantitative Research

ABSTRACT

To estimate the effects of adaptive education on cognitive, affective, and behavioral outcomes of learning, 309 effect sizes were calculated using statistical data from 38 studies that contained a combined sample of approximately 7,200 students. The substantial mean of the study-weighted effect sizes is .45, suggesting that the average student in adaptive programs scores at the 67th percentile of control group distributions. The effects appear constant across grades, socioeconomic levels, races, private and public schools, and community types. In addition, the effects are not significantly different across the categories of adaptiveness, student outcomes, social contexts, and methodological rigor of the studies. (Author)

1985/11

A QUANTITATIVE SYNTHESIS

HERSHOLT C. WAXMAN, MARGARET C. WANG,
KENNETH A. ANDERSON, AND HERBERT J. WALBERG

University of Pittsburgh

ADAPTIVE EDUCATION AND STUDENT OUTCOMES: A QUANTITATIVE SYNTHESIS

Hersholt C. Waxman
Education Research Center
University of Houston

Margaret C. Wang and Kenneth A. Anderson
Learning Research and Development Center
University of Pittsburgh

Herbert J. Walberg
University of Illinois at Chicago

Learning Research and Development Center
University of Pittsburgh

1985

To be published in Journal of Educational Research, a publication of the Helen Dwight Reid Educational Foundation. Reprinted with permission.

The material reproted herein was supported by the Learning Research and Development Center, funded in part by the National Institute of Education, U.S. Department of Education. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred.

Abstract

To estimate the effects of adaptive education on cognitive, affective, and behavioral outcomes of learning, 309 effect sizes were calculated using statistical data from 38 studies that contained a combined sample of approximately 7,200 students. The substantial mean of the study-weighted effect sizes is .45, suggesting that the average student in adaptive programs scores at the 67th percentile of control group distributions. The effects appear constant across grades, socioeconomic levels, races, private and public schools, and community types. In addition, the effects are not significantly different across the categories of adaptiveness, student outcomes, social contexts, and methodological rigor of the studies.

ADAPTIVE EDUCATION AND STUDENT OUTCOMES: A QUANTITATIVE SYNTHESIS

Hersholt C. Waxman, Margaret C. Wang,
Kenneth A. Anderson, and Herbert J. Walberg

Although instruction that is adapted to individual differences has a long history (Whipple, 1925), it recently has been of great interest in current efforts aimed at improving educational opportunities for school-aged children and young adults (Birch & Reynolds, 1982; Fenstermacher & Goodlad, 1983). Executive, judicial, and legislative branches of government have called for increasing schools' responsiveness to the diverse needs of their students (National School Public Relations Association, 1981; Reynolds & Wang, 1981). School administrators and teachers indicated in a recent survey their preference for assessment of each student's capabilities and instruction that builds upon each student's skills (Rothrock, 1982). Furthermore, adapting instruction to the individual needs of the increasingly diverse student population in regular classes has come to be a major topic in effective schooling literature (Wang & Lindvall, 1984).

Adaptive programs make use of a variety of curricula and techniques that have proven effective in many types of classroom settings and with a diversity of students. These include mastery learning, cooperative teams, and individual tutorials, as well as large- and small-group instructional approaches that have shown moderate to large effect sizes in previous research syntheses (Walberg, 1984). Adaptive programs are intended to make use of these techniques in a manner best suited to each teacher, class, and student,

and there is considerable variety across programs. Within any program, moreover, teachers are expected to vary in their use of materials and procedures; and students, of course, are expected to be treated differentially according to their educational needs (see Wang & Lindvall, 1984). Adaptive educators contend that the use of such programs provides learning experiences that meet the needs of special and general education students in regular classroom settings and, therefore, results in opportunities for most students to succeed in their school learning.

Despite increased interest in research and development activities aimed at ensuring success for each student, there are some inconsistent and difficult-to-interpret results on the educational impact of adaptive instruction. The present synthesis investigates these results by addressing the following questions: First, how extensive is the empirical evidence on the relationship between adaptive education and student outcomes? Second, what is the magnitude and direction of the relationship between adaptive education and student outcomes? Third, are there certain social contexts or student characteristics that affect the relationship? Fourth, are there particular methodological characteristics that affect the relationship? Finally, are there some specific characteristics of adaptive education that affect its relationship with student outcomes?

To answer these questions, the present study quantitatively synthesized the experimental and quasi-experimental, published and unpublished research on the effects of adaptive education on student outcomes in naturalistic

settings. The techniques of research synthesis that were applied were derived from the work of Glass, McGaw, and Smith (1981) and Hunter, Schmidt, and Jackson (1982) on meta-analysis, as well as from contributions of Hauser-Cram (1983), Cooper and Rosenthal (1980), and Walberg and Haertel (1980).

Method

Search and Selection Procedure

A systematic collection of a decade of research published from 1973 through 1982 investigating the effects of adaptive education on student outcomes was conducted by searching Current Index to Journals in Education, Resources in Education, Review of Educational Research, American Educational Research Journal, Journal of Educational Psychology, Journal of Educational Research, and Journal of Experimental Education. Studies cited in the references listed in these sources also were examined.

Several criteria were established for inclusion in the present synthesis. The study had to have been conducted in a regular classroom setting in a public or private elementary or secondary school. It had to have either contrasted-groups or correlational results. In addition, the study had to have sufficient quantitative data bearing on the sample population to calculate effect sizes (Glass, McGaw, & Smith, 1981). An effort was made to include not only published articles but also dissertations and unpublished reports showing quantitative effects on student outcomes.

In order to be considered a study of adaptive education, and, therefore, be included in the synthesis, the program investigated had to contain at least one of the following characteristics (Wang & Lindvall, 1984):

1. Instruction based on the assessed capabilities of each student.

2. Materials and procedures that permit each student to make progress in mastering the curricula at a pace suited to his or her abilities and interests.

3. Periodic evaluations of progress to inform the student of his or her mastery.

4. Students' assumption of self-responsibility for diagnosing their current needs and abilities, for planning and pursuing individual learning activities, and for evaluating mastery.

5. Alternative learning activities and materials for aiding students' acquisition of academic skills and content.

6. Student choice in selecting educational goals, outcomes, and activities.

7. Students' assistance to one another in pursuing individual goals and cooperation in achieving group goals.

In those cases where the programs were not completely described in the article, additional sources were consulted to accurately assess whether any of these characteristics of adaptive education were present.

Certain types of studies and reports were excluded from the synthesis. Special education populations (e.g., students classified as mentally retarded, learning-disabled) were left out. Reports of computer-assisted instructional programs and other types of programmed instruction also were omitted. Studies comparing several different instructional approaches (e.g., Becker, 1978; Bosco, Harring, & Bandy, 1976) were excluded, as were studies such as Slavin and Tanner (1979) and Johnson, Johnson, and Scott (1978), which compared various components of adaptive education (e.g., cooperative versus individualized instructional programs), because in such instances it was difficult to determine which program should be used as the control. In addition, only the most recent of multiple reports of the same population and results was included in order to maintain statistical independence. This may have occurred when the investigator conducted an empirical study to satisfy the requirements of a Ph.D. dissertation and subsequently published an article reporting the same results.

The search and selection procedures resulted in a collection of 38 studies. Of these, 24 were published articles, 10 were unpublished reports, and four were dissertations.

Analysis

To calibrate the studies' results, or place them on a common scale, "effect sizes" were calculated. These consisted of the treatment group mean minus the control mean divided by the control standard deviation. Effect sizes can be considered a standardized estimate of where the treatment group stands in comparison to the control group distribution. In the present case, a positive effect size indicated that the adaptive education group received higher (i.e., more desirable) scores than the control group. Glass, McGaw, and Smith's (1981) formulas were employed for studies that did not report group means or standard deviations but contained F or t values, correlations, or other statistics from which effect sizes could be calculated.

For the present synthesis, one investigator recorded 39 codable characteristics and other data for each of the 309 effect sizes from 38 studies. (The coding categories are listed in the Appendix.) A second investigator independently coded five studies. The intercoder agreement for each study reviewed exceeded the 85% criterion. The methodological threats to validity were adopted from Cook and Campbell (1979). The 39 categorical variables were employed as factors in an analysis of variance (ANOVA).

The studies varied by the number of comparisons they reported. Therefore, those studies with a greater number of comparisons (e.g., those that reported separate results by ability level, sex, or race) would have been weighted more heavily than others if each comparison had been given equal

weight. To give all studies the same unit weight in the analysis, each comparison was weighted in inverse proportion to the number of comparisons in the study from which it was taken (i.e, 1/n where n = number of comparisons in the study). Each of Goldner's (1973) four comparisons, for example, received a weight of .25.

Results and Discussion

Characteristics of the 7,200 students in the 38 studies covered a broad spectrum. The median number of students in the sample of studies was 108 and the range was from 19 to 3,018 students. Sixteen studies involved primary students (grades K-3), 11 included intermediate students (grades 4-6), and 11 focused on junior-high and high-school students. Eleven studies focused on math as the subject area and 10 studies involved reading. Language arts and science were each the focus of four studies, while the remaining studies included a variety of other subject areas.

The 309 comparisons are displayed in a stem-and-leaf diagram (Tukey, 1977) in Figure 1. The stem-and-leaf diagram is useful for exhibiting such results because it provides all the information of a histogram as well as displaying the actual value of each effect size to two decimal places. The first digit, or stem, appears to the left of the vertical line, and subsequent digit(s), or leaves, appear to the right. For example, the leaves 0, 1, and 2 to the right of the stem 1.5 represent three effect sizes: 1.50, 1.51, and 1.52. Figure 1 also reports the five-point summary of the distribution, consisting of minimum (*), first quartile or hinge (H), median (M), third quartile or hinge (H), and maximum (*).

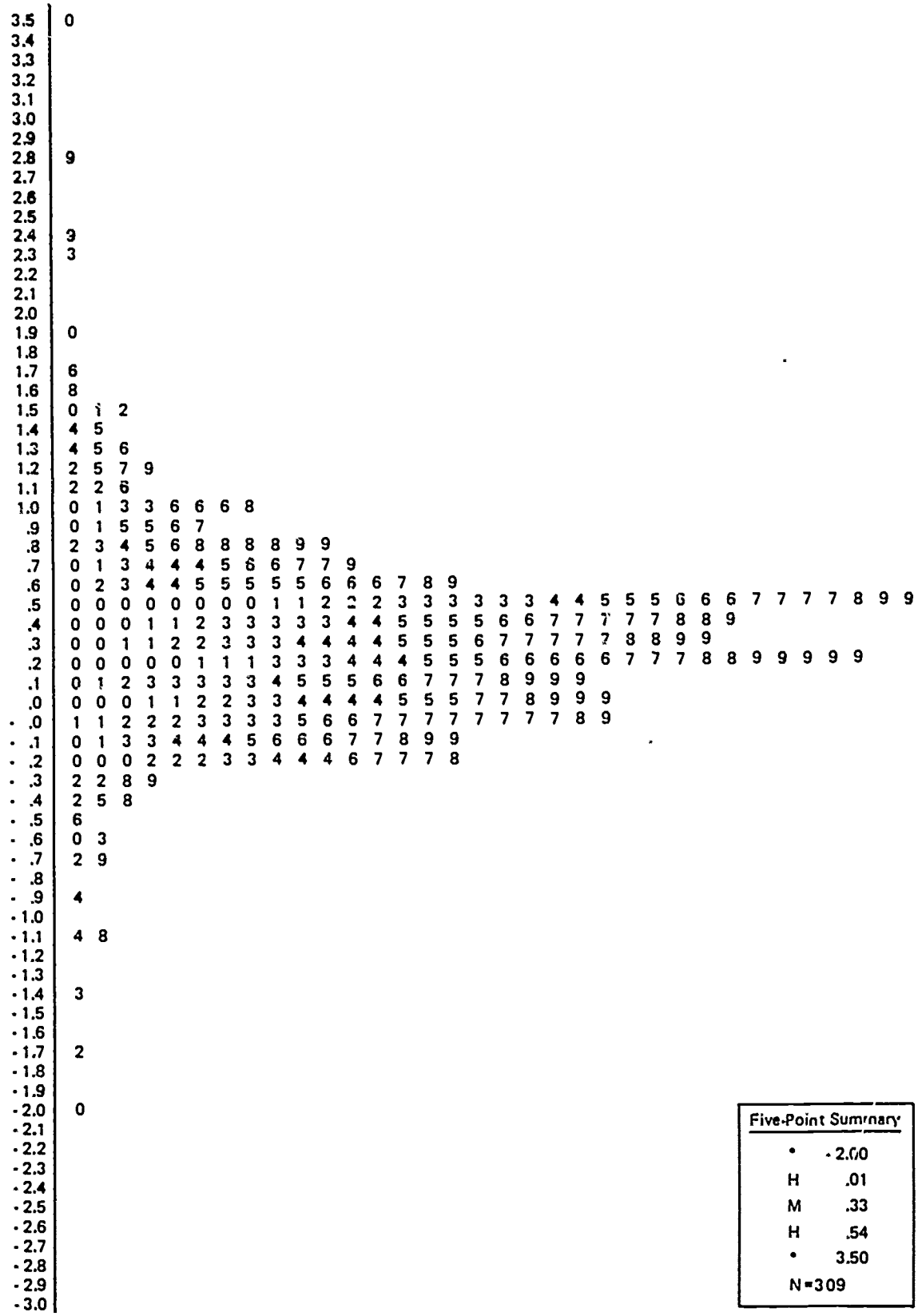

Figure 1. Stem-and-leaf and five-point summary of effect sizes.

Figure 1 shows that the majority of comparisons lie between -.4 and 1.5, but that the range extends from -2.00 to 3.50. Fifty percent are between .02 and .50. Seventy-seven percent of the effect sizes are positive and favor the adaptive education groups.

The 38 study-weighted effect sizes between adaptive education and student outcomes are displayed in Figure 2. The majority of the study-weighted effect sizes lie between .1 and .7. The range extends from -1.18 to 1.73. Thirty-four of the 38 comparisons, or 89%, are positive. The binomial probabilities of the evidence, given the hypothesis of an even 50-50 or chance split, is less than .01 (Dixon & Massey, 1957, p. 417).

Table 1 lists the 38 studies synthesized, the number of comparisons in each study, the mean study-weighted effect size for each study, and a breakdown of the mean study-weighted effect size for each student outcome. Effect sizes were converted to percentiles by obtaining the item in the column of percentiles corresponding to the Z-score equal to the effect size in a standard table of the normal distribution (see, for example, Talmage, 1976, pp. 60-61). The mean study-weighted effect size for all 309 comparisons for the 7,200 students is .45. This suggests percentiles on student outcomes of 67 and 50 for the experimental and control groups respectively. The mean study-weighted effect size for the 208 comparisons containing cognitive outcomes is .39, suggesting percentiles on student cognitive outcomes of 65

1.7	3
1.6	
1.5	
1.4	
1.3	
1.2	1 2
1.1	
1.0	1 2
.9	
.8	5
.7	4 4
.6	5 9 9
.5	2 3 4 4 5 7 7
.4	4 8 8
.3	0 5 7 7 8
.2	4 4
.1	1 2 4 7 9
0	9
- 0	8
- .1	
- .2	1
- .3	
- .4	0
- .5	
- .6	
- .7	
- .8	
- .9	
- 1.0	
- 1.1	8
- 1.2	
- 1.3	
- 1.4	
- 1.5	
- 1.6	
- 1.7	

Five-Point Summary	
*	- 1.18
H	.12
M	.48
H	.69
*	1.73
N = 38	

Figure 2. Stem-and-leaf and five-point summary of study-weighted effect sizes between adaptive education and student outcomes.

Table 1

Effect Sizes for Each Study

Study	N	Study Mean	S.D.	Means by Outcome		
				Cognitive	Affective	Behavioral
Abelson (1974)	44	.144	.384	.201	-.113	
Abrams et al. (1979)	1	.190		.190		
Bar-Tal et al. (1978)	9	1.210	1.240		1.210	
Breuning (1978)	3	.119	.430	.119		
Broussard (1975)	1	.536		.536		
Campbell (1978)	1	.572		.572		
Clasen (1974)	8	.543	.504	.543		
Cohen & Rodriquez (1980)	2	.528	.063	.528		
Corley & Lewis (1975)	9	.350	.446	.350		
Dornseif (1976)	1	.569		.569		
Fennell (1973)	3	.172	.206	.172		
Fesler et al. (1976)	3	.685	.302	.685		
Goldner (1973)	4	.444	.123	.444		
Hales (1976)	7	.366	.194	.366		
Hales (1980)	16	.380	.347	.380		
Hales (1981)	18	.477	.340	.477		
Hales (1982)	12	.548	.281	.548		
Halliwell (1975)	3	.108	.022	.108		
Jerman (1973)	12	- .210	.224	- .210		
Keim-Abbott & Abbott (1977)	1	-1.182		-1.182		
Kourilsky & Baker (1975)	20	- .398	.731		- .398	
Krebs (1982)	24	.370	.380	.551	.189	
Leinhardt & Engel (1981)	2	.295	.375	.295		
Levin (1979)	2	1.005	.629	1.005		
Linn et al. (1976)	6	.741	.538	.741		
Marquez (1976)	8	.239	.195	.300	.178	
McDuffie & DeRose (1982)	8	1.022	.397	1.060	.757	
Parr et al. (1981)	15	- .096	.231	- .141	.198	
Rayder et al. (1978)	4	.647	.311	.647		
Slavin et al. (1983)	7	.237	.099	.237		
Slavin et al. (in press)	13	.477	.310	.227	.492	.576
Stefanides (1976)	2	1.728	1.069	.972	2.484	
Stone & Vaughn (1976)	2	1.215	.081	1.215		
Wang (1979)	4	.849	.451		.948	.816
Wang & Stiles (1976 a)	8	.691	.817			.691
Wang & Stiles (1976 b)	10	.738	.571		.738	
Wendt (1980)	4	.094	.092	.094		
Whitley (1979)	12	.523	.250		.523	
$\bar{X}$		.445	.383	.394	.601	.694
S.D.		.488	.275	.435	.746	.120

Note. Total N of studies = 38
Total N of comparisons = 309

and 50 for the experimental and control groups respectively. The mean study-weighted effect size for the 81 comparisons that focused on student affective outcomes is .60, while the mean study-weighted effect size for the 20 comparisons that contained behavioral outcomes is .69. These correspond to percentiles of 73 for affective outcomes and 75 for behavioral outcomes.

The relationship of each independent variable to the mean study-weighted effect size was tested for significance using ANOVA. Table 2 lists the F-values for the 39 conditioning variables for all 38 studies. The results indicate that none of the variables had a statistically significant impact on the mean study-weighted effect size. This suggests that no one conditioning variable affected the overall student outcomes or the gains attributable to the adaptive education program, and that any one produced robust, consistent, and moderately large effects. In other words, the results do not differ significantly across categories of adaptiveness, student outcomes, social contexts, methodological rigor and characteristics of the study, and subject characteristcs.

Conclusion

The results of the synthesis show strong and consistent effects of adaptive education. The mean effect size of .45, or nearly half a standard deviation, is considerably larger than the average of several dozen syntheses of productive factors in classroom learning compiled by Walberg (1984). In fact, it is more than twice the average effect size of .20 standard deviations in classroom research of the past few decades that has been recently

Table 2

F-Values for the Impact of All Conditioning Variables on the Mean Study-Weighted Effect Sizes

Variable	F-Value
Adaptiveness Characteristics	
Determination of initial capabilities	1.57
Continuous assessment	.64
Individualized progress	.00
Periodic evaluations	1.43
Student self-management	.80
Alternative routes and materials	.14
Student choice of goals	.93
Peer assistance	.81
Outcomes	
Cognitive	.20
Affective	.31
Behavioral	.36
Methodological Rigor	
Contamination and compensation	.22
Control method	1.65
Generalizability	.59
History	.73
Instrument type	1.34
Maturation	.54
Method of observation	1.25
Mortality	1.77
Operational irrelevancy	1.01
Pretest equivalency	1.41
Reliability of measures	.24
Reliability of treatment	.73
Selection bias	1.71
Statistical power	.32
Statistics	.30
Social Contexts	
Community type	.31
Region	.17
School type	.12
Study characteristics	
Publication features	.19
Sample size	1.25
Subject area	.87
Unit of analysis	.64
Year	1.19
Subject Characteristics	
Ability level	.21
Grade level	.27
Race	.12
Sex	.60
SES	.11

synthesized. The relatively strong effects are also robust and consistent across contexts of learning and categories of students. Such robustness and consistency are scientifically valuable because the results are generalized across a wide variety of conditions that have been investigated, as well as across grade levels and student characteristics such as ability, socioeconomic status, race, and sex.

In contrast to the results from earlier reports of inconsistent and low or negative effects of adaptive instruction, the consistent outcomes reflected in the past decade's reports of research on the effects of adaptive instruction are particularly noteworthy. Tailoring instruction to respond to the learning characteristics and needs of individual students seems to be an educational alternative that can be effective in obtaining intended social and academic outcomes.

References

Abelson, W.D., Zigler, E., & De Blais, C.L. (1974). Effects of a four-year Follow Through program on economically disadvantaged children. Journal of Educational Psychology, 66, 756-711.

Abrams, A.I., & Body, B., Rayder, N.F. (1979). Problems and solutions in evaluating child outcomes of large-scale educational programs. Journal of Experimental Education, 48, 155-164.

Bar-Tal, D., Leinhardt, G., & Bar-Tal, Y. (1978). Effect of an innovative science program on perceptions of locus of control and satisfaction. Science Education, 62, 349-357.

Becker, W.C. (1978). The national evaluation of Follow Through: Behavior-theory-based programs come out on top. Education and Urban Society, 10, 431-458.

Birch, J., & Reynolds, M. (1982). Special education as a profession. Exceptional Education Quarterly, 2(4), 1-13.

Bosco, J., Harring, R., & Bandy, R. (1976). Management systems for individualized instruction: An analysis of five approaches. Education, 96, 251-264.

Breuning, S.E. (1978). Precision teaching in the high school classroom: A necessary step towards maximizing teacher effectiveness and student performance. American Educational Research Journal, (15), 125-140.

Broussard, V. (1975). A personalized-individualized instructional approach on achievement in mathematics. California Journal of Education Research, 26, 233-237.

Campbell, R.L. (1978). Effects of teacher training in the individualized science materials on achievement of first grade students. Science Education, 62, 523-527.

Clasen, R.E. (1974). Evaluation report for strategies in early childhood education. Madison, WI: University of Wisconsin, Instructional Research Laboratory.

Cohen, S.A., & Rodriquez, S. (1980). Experimental results that question the Ramirez-Castaneda model for teaching reading to first grade Mexican Americans. Reading Teacher, 34(1), 12-18.

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

Cooper, H.M., & Rosenthal, R. (1980). A comparison of statistical and traditional procedures for summarizing research. Evaluation in Education, 4, 33-36.

Corley, G.B., & Lewis, C.W. (1975). The impact of individualized instruction on low achieving youth. Urban Education, 10, 321-328.

Dixon, W.J., & Massey, F.J. (1957). Introduction to statistical analysis. New York: McGraw-Hill.

Dornseif, A. (1976). Dissemination review panel: Matteson Four-Dimensional Reading Program. Matteson, IL: Matteson School District 162.

Fennell, F.M. (1973). The effect of diagnostic mathematics instruction on the achievement and attitude of low achieving sixth grade pupils. Unpublished doctoral dissertation, Pennsylvania State University.

Fenstermacher, G.D., & Goodlad, J.I. (Eds.). (1983). Individual differences and the common curriculum. Eighty-second yearbook of the National Society for the Study of Education (Part I). Chicago: National Society for the Study of Education.

Fesler, E., Guidubaldi, J., & Kehle, T.J. (1976). Effects of Follow Through model: Primary Education Project-Individually Prescribed Instruction (PEP-IPI) on children's academic competence. Psychology in the Schools, 13, 181-184.

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

Goldner, L. (1973). Don't give up on compensatory education: Just make it more relevant to individual needs. Urban Education, 8, 311-331.

Hales, C.M. (1976). Report for submission to Joint Dissemination Review Panel: Utah System Approach to Individualized Learning (U-SAIL), Levels 1-6 in mathematics and reading. Salt Lake City: U-SAIL Center.

Hales, C.M. (1980). Report for submission to Joint Dissemination Review Panel: Utah System Approach to Individualized Learning (U-SAIL), Levels 7-9. Salt Lake City: U-SAIL Center.

Hales, C.M. (1981). U-SAIL elementary school achievement data. Salt Lake City: U-SAIL Center.

Hales, C.M. (1982). Report for re-submission to Joint Dissemination Review Panel: Utah System Approach to Individualized Learning (U-SAIL), Levels 1-9. Salt Lake City: U-SAIL Center.

Halliwell, D.D. (1975). A study of the effect of different approaches to the teaching of vocabulary on reading achievement and vocabulary development at the fifth grade level. Unpublished doctoral dissertation, Syracuse University.

Hauser-Cram, P. (1983). Some cautions in synthesizing research studies. Educational Evaluation and Policy Analysis, 5, 155-162.

Hunter, J.E., Schmidt, F.L., & Jackson, G.B. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.

Jerman, M. (1973). Individualized instruction in problem solving in elementary school mathematics. Journal for Research in Mathematics Education, 4, 6-19.

Johnson, D.W., Johnson, R.T., & Scott, L. (1978). The effects of cooperative and individualized instruction on student attitudes and achievement. Journal of Social Psychology, 104, 207-216.

Keim-Abbott, S., & Abbott, R.D. (1977). Moderation of achievement prediction in an elementary school metric curriculum by trait X instructional method interactions. Educational and Psychological Measurement, 37, 481-486.

Kourilsky, M., & Baker, E.L. (1975). An empirical comparison of open and nonopen structured classrooms. California Journal of Education Research, 26, 238-245.

Krebs, L.L. (1982). Summary of research on an Individual Education (IE) school. Individual Psychology, 38, 245-252.

Leinhardt, G., & Engel, M. (1981). An iterative evaluation of NRS: Ripples in a pond. Evaluation Review, 5, 579-601.

Levin, T. (1979). Instruction which enables students to develop higher mental processes. Evaluation in Education, 3, 173-220.

Linn, M.C., Chen, B., & Thier, H.D. (1976). Personalization in science: Preliminary investigation at the middle school level. Instructional Science, 5, 227-252.

Marquez, A.A. (1976). The effect of individualized instruction on academic growth and attitudes toward school in low achieving 6th grade students. Unpublished doctoral dissertation, University of Alabama.

McDuffie, T.E., Jr., & DeRose, J.V. (1982). Five years of achievement in ISCS. Science Education, 66, 35-43.

National School Public Relations Association (1981). Good schools: What makes them work? Arlington, VA: National School Public Relations Association.

Parr, G.D., Baca, F., & Dixon, P. (1981). Individualized versus group instruction in bilingual education: A two-year study. Elementary School Journal, 81, 223-227.

Rayder, N., Boady, B., & Nimnicht, G. (1978). Assessing Follow Through: Changes in intelligence test scores over two and three years of experience in the Responsive Program. Journal of Experimental Education, 47, 60-66.

Reynolds, M.C., & Wang, M.C. (1981). Restructuring school programs: A position paper (LRDC Publications Series 1981/24). Pittsburgh: University of Pittsburgh, Learning Research and Development Center.

Rothrock, D. (1982). The rise and decline of individualized instruction. Educational Leadership, 39, 528-531.

Slavin, R.E., Leavey, M., & Madden, N.A. (1983, April). Combining student teams and individualized instruction in mathematics: An extended evaluation. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Slavin, R.E., Leavey, M., & Madden, N.A. (in press). Combining cooperative learning and individualized instruction: Effects on student mathematics achievement, attitudes, and behaviors. Elementary School Journal.

Slavin, R.E., & Tanner, A.M. (1979). Effects of cooperative reward structures and individual accountability on productivity and learning. Journal of Educational Research, 72, 294-298.

Stefanides, G.A. (1976). An evaluation of achievement and attitudes of students and teachers using an individualized spelling approach. Unpublished doctoral dissertation, University of Pittsburgh.

Stone, R., & Vaughn, L. (1976). Implementation and evaluation of the self-schedule system in an adaptive school learning environment. In M.C. Wang (Ed.), The self-schedule system for instructional-learning management in adaptive school learning environments (LRDC Publications Series 1976/9) (pp. 55-88.) Pittsburgh: University of Pittsburgh, Learning Research and Development Center.

Talmage, H. (1976). Statistics as a tool for educational practitioners. Berkeley, CA: McCutchan.

Tukey, J.W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.

Walberg, H.J. (1984). Synthesis of research on teaching. In M.C. Wittrock (Ed.), Handbook of research on teaching. Washington, DC: American Educational Research Association.

Walberg, H.J., & Haertel, E.H. (Eds.). (1980). Evaluation in Education, 4, 1-44.

Wang, M.C. (1979). Maximizing the effective use of school time by teachers and students. Contemporary Educational Psychology, 64, 87-201.

Wang M.C., & Lindvall, C.M. (1984). Individual differences and school learning environments: Theory, research, and design. In E.W. Gordon (Ed.), Review of research in education (Vol. 11). Washington, DC: American Educational Research Association.

Wang, M.C., & Stiles, B. (1976a). Effects of the self-schedule system on teacher and student behaviors. In M.C. Wang (Ed.), The self-schedule system for instructional-learning management in adaptive school learning environments (LRDC Publication Series 1976/9) (pp. 7-54.) Pittsburgh: University of Pittsburgh, Learning Research and Development Center.

Wang, M.C., & Stiles, B. (1976b). An investigation of children's concept of self-responsibility for their school learning. American Educational Research Journal, 13, 159-179.

Wendt, R.N. (1980). Project evaluation report for strategies in early childhood education in Keene, N.H. Union School District. Toledo, OH: University of Toledo, College of Education.

Whipple, G.M. (Ed.). (1925). Twenty-fourth yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

Whitley, T.W. (1979). The effects of individualized instruction on the attitudes of middle school pupils. Journal of Education Research, 72, 188-193.

APPENDIX

Information Coded for Each Study

Adaptiveness Characteristics	Coding Specifications
Determination of initial capabilities	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.
Continuous assessment of capabilities	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.
Individualized progress	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.
Periodic evaluations of progress	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.
Student self-management	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.
Alternative routes and materials	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.
Some student choice of goals	Unspecified = 0; Inadequately specified or implied = 0; Adequately specified = 2.

Peer assistance and group projects	Unspecified = 0; Inadequately specified or implied = 1; Adequately specified = 2.

Outcomes

Cognitive outcomes	Unspecified = 00; Academic achievement, factual knowledge = 01; Academic achievement, higher level, (comprehension, understanding, concept attainment) = 02; I.Q. tests = 03; Grades = 04; Discrimination learning = 05; Accuracy = 06; No. of problems attempted = 07; Behavior modification = 08; Accuracy task-general = 09; Correctly completed-general = 10; Correctly completed task-general = 11; Accuracy-reading/writing, English = 12.
Affective outcomes	Unspecified = 00; Self-concept = 01; School concept (attitude toward school, teacher, etc.) = 02; Creativity = 03; Curiosity = 04; Independence = 05; Adjustment, security = 06; Locus of control = 07; Cooperation 08; Persistence = 09; Motivation = 10; General affective gain = 11; Teacher attitude = 12.
Behavioral outcomes	Unspecified = 00; Time-on-task = 01; Socialization = 02; Opportunity to learn = 03.

Methodological Rigor

Contamination and compensation	Do untreated control groups imitate or somehow gain the benefits of a treatment, or work harder to compete with experimental group (evidence of teacher interchange or other bias)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.
Control method	Unspecified = 0; Stratification (e.g., %, means, Chi square) = 1; Partial correlation = 2; Beta weights in regression = 3; Raw or metric weights in regression = 5; Factorial analysis of variance = 6; Analysis of covariance = 7; None = 8.
Generalizability	Can the cause-effect be generalized to other times, units, or settings (random or typical samples of students, classes, or schools)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.
History	Have external factors in the environment, other than the presumed treatment causes, brought about the resultant change (evidence of differential teacher assignment or other bias)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.

Instrument type for dependent variable	Unspecified = 0; Local = 1; Standardized = 2.
Maturation	Have factors within the units, rather than the presumed treatment causes, brought about changes or differences (extended treatment or without random assignment to treatments and control conditions)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.
Method of observation of independent variable	Unspecified = 0; Student questionnaire = 1; Teacher report, observer report, etc. = 2; Other = 3.
Mortality	Do differential dropout rates in the treatment, or control groups, account for differences on the dependent variable obtained at the end of the experiment (self-evident)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.
Operational irrelevancy	Does the treatment, as administered, contain Hawthorne, experimenter-bias, placebo, or other unintended causes that produce the effects? Unspecified = 0; Adequate-self-evident in the report = 2.

Pretest equivalency	Have the initial differences between the two groups been accounted for? Unspecified = 0; Statistical control (e.g., ANCOVA, regression) = 1; Random assignment = 2; Statistical control and random assignment = 3; Gain scores = 4.
Reliability of measures	Are the outcome measurement instruments internally consistent or stable enough to detect changes during, or differences in, treatments? Unspecified = 00; Adequate = 01; Actual reliability figure = NN.
Reliability of treatment	Have the units in a treatment actually undergone the same conditions as the one in the control group? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.
Selection bias	Do pre-existing differences among the groups account for later differences on the dependent variable (e.g., different I.Q. levels)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.

Statistical power	Is the sample size large enough to reject the null hypothesis at a given level of probability or are estimate coefficients with reasonably small margin of error (60 or more classes; 100 or more individuals)? Unspecified = 0; Inadequate indication = 1; Adequate indication = 2.
Statistics	t-value = 1; F-value = 2; Correlations = 3; Chi-square = 4; Means = 5.

Social Contexts

Community type	Unspecified = 0; Urban = 1; Rural = 2; Suburban = 3; Other = 4.
Region	Unspecified = 0; East = 1; Midwest = 2; West = 3; South = 4; More than one = 5; Other country = 6.
School type	Unspecified = 0; Public = 1; Private = 2.

Study Characteristics

Publication features	Unpublished = 1; Published = 2; Ph.D. dissertation = 3.
Sample size	Unspecified = 0000; 4 spaces for exact N.

Subject area	Unspecified = 0; Reading = 1; Math = 2; Language arts = 3; Science = 4; Social studies = 5; Other = 6.
Unit of analysis	Unspecified = 0; Individual = 2; Class = 3; Grade = 4; Pod, module, etc. = 5; School = 6; Varied = 7.
Year of study	4 spaces for each year.

Subject Characteristics

Ability level	Unspecified = 0; All low = 1; All middle = 2; All high = 3; Special education students = 4; Mixed/varied = 5.
Grade level	Unspecified = 00; First grade = 01; Second grade = 02; Mixed primary (K-3) = 22; Mixed middle (4-6) = 23; Mixed upper (7-8) = 24; Mixed high school = 25; Mixed other = 26.
Race	Unspecified = 0; Black = 1; White = 2; Latino = 3; Oriental = 4; Mixed = 5; Other Other ethnic = 6.
Sex	Boys = 1; Girls = 2; Mixed = 3.
Socioeconomic status	Unspecified = 0; Lower = 1; Middle = 2; Upper = 3; Mixed = 4.