DOCUMENT RESUME

ED 242 559 SE 044 360

AUTHOR	Slavin, Robert E.; Karweit, Nancy L.
TITLE	Mathematics Achievement Effects of Three Levels of Individualization: Whole Class, Ability Grouped, and Individualized Instruction.
INSTITUTION	Johns Hopkins Univ., Baltimore, Md. Center for Social Organization of Schools.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
REPORT NO	JHU-CSOS-349
PUB DATE	Jan 84
GRANT	NIE-G-83-0002
NOTE	42p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Ability Grouping; Educational Research; Elementary Education; *Elementary School Mathematics; Grouping (Instructional Purposes); *Individualized Instruction; Intermediate Grades; *Large Group Instruction; *Mathematics Achievement; *Mathematics Instruction; Student Attitudes
IDENTIFIERS	*Mathematics Education Research

ABSTRACT

This research evaluated mathematics achievement and attitudinal effects of three instructional methods directed in varying degrees toward accommodating diversity in students' prior achievement. Two randomized field experiments of 16 and 18 weeks' duration, respectively, compared an individualized model, Team Assisted Individualization (TAI); an ability grouped model, Ability Grouped Active Teaching (AGAT); a group-paced model, the Missouri Mathematics Program (MMP); and, in Experiment 2 only, untreated control classes. The first experiment involved pupils in grades 4-6, while the second experiment was conducted with pupils in grades 3-5. Nested analysis of variance on CTBS Computations scores adjusted for pretests indicated that in both experiments, TAI and AGAT exceeded MMP. In Experiment 2, TAI, AGAT, and MMP also exceeded control. No interactions with prior achievement were found. Effects on Liking of Math Class and (in Experiment 1) Self-Concept in Math favored TAI. (Author/MNS)

Report No. 349
January 1984

MATHEMATICS ACHIEVEMENT EFFECTS OF THREE LEVELS OF INDIVIDUALIZATION: WHOLE CLASS, ABILITY GROUPED, AND INDIVIDUALIZED INSTRUCTION

Robert E. Slavin and Nancy L. Karweit

The Johns Hopkins University

STAFF

EDWARD L. MCDILL, CO-DIRECTOR
JAMES M. MCPARTLAND, CO-DIRECTOR

KARL L. ALEXANDER
HENRY J. BECKER
JOMILLS H. BRADDOCK, II
RUTH H. CARTER
ROSLYN L. CHESTER
MICHAEL COOK
ROBERT L. CRAIN
DORIS R. ENTWISLE
JOYCE L. EPSTEIN
JAMES FENNESSEY
DENISE C. GOTTFREDSON
GARY D. GOTTFREDSON
LINDA S. GOTTFREDSON
EDWARD J. HARSCH
JOHN H. HOLLIFIELD
BARBARA J. HUCKSOLL
RAMONA M. HUMPHREY
LOIS G. HYBL
NANCY L. KARWEIT
HAZEL G. KENNEDY
MARSHALL B. LEAVEY
GRETCHEN M. LUEBBE
NANCY A. MADDEN
ALEJANDRO PORTES
ROBERT E. SLAVIN
VALARIE SUNDERLAND
GAIL E. THOMAS
L. THOMAS WEBB
JANICE G. ZDANIS

MATHEMATICS ACHIEVEMENT EFFECTS OF THREE LEVELS OF INDIVIDUALIZATION: WHOLE CLASS, ABILITY GROUPED, AND INDIVIDUALIZED INSTRUCTION

Grant No. NIE-G-83-0002

Robert E. Slavin
Nancy L. Karweit

Report No. 349

Janaury 1984

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the National Institute of Education (NIE), U.S. Department of Education. The opinions expressed in this publication do not necessarily reflect the position nor policy of the NIE, and no official endorsement by the Institute should be inferred.

Center for Social Organization of Schools
The Johns Hopkins University
3505 North Charles Street
Baltimore, MD 21218

Printed and assembled by VSP Industries, Baltimore MD

Introductory Statement

The Center for Social Organization of Schools (CSOS) has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through three research programs to achieve its objectives:

The School Organization Program investigates how school and classroom organization affects student learning and other immediate outcomes of schooling. Current studies focus on parental involvement, microcomputers in schools, use of time in schools, cooperative learning, and other organizational strategies that alter the task, reward, authority and peer group structures in schools and classrooms.

The Education and Work Program examines the relationship between schooling and students' later-life occupational and educational successes. Current projects include studies of the competencies required in the workplace, the sources of training and experience that lead to employment, college students' major field choices, and employment of urban minority youth.

The Schools and Delinquency Program studies the problems of crime, violence, vandalism, and disorder in schools and the role that schools play in delinquency. Ongoing projects address the development of a theory of delinquent behavior, school effects on delinquency, and the evaluation of delinquency prevention programs in and out of schools.

CSOS also supports a Fellowships in Education Research program that provides opportunities for talented researchers to conduct and publish significant research in conjunction with the three research programs.

This report, prepared by the School Organization Program, presents results of two randomized field experiments that compared the effects on students of three types of instructional processes.

ABSTRACT

This research evaluated mathematics achievement and attitudinal effects of three instructional methods directed in varying degrees toward accomodating diversity in students' prior achievement. Two randomized field experiments of 16 and 18 weeks' duration, respectively, compared an individualized model, Team Assisted Individualization (TAI); an ability grouped model, Ability Grouped Active Teaching (AGAT); a group-paced model, the Missouri Mathematics Program (MMP); and, in Experiment 2 only, untreated control classes. Nested analysis of variance on CTBS Computations scores adjusted for pretests indicated that in both experiments, TAI and AGAT exceeded MMP. In Experiment 2, TAI, AGAT, and MMP also exceeded control. No interactions with prior achievement were found. Effects on Liking of Math Class and (in Experiment 1) Self-Concept in Math favored TAI.

Acknowledgments

We would like to thank Thomas Good, Kathy Glyshaw, Marshall Leavey, Peter Idstein, and William Wright for their invaluable assistance with this research, and Jere Brophy and Nancy Madden for comments on earlier drafts.

Introduction

One of the most troublesome and enduring problems of mathematics instruction is accomodating heterogeneity in student preparedness and learning rate. If students do not have the prerequisite skills to learn a lesson, or if they have already mastered it before the lesson began or do so in the first minutes of the lesson, then instructional time for them is wasted. For example, it is of little use for students to sit through a lesson on dividing with two digit divisors if they: a) did not master one-digit division, b) already know two-digit division, or c) learn the concept in a few minutes. Yet in most mathematics classes in which the teacher is teaching one lesson to the entire class, we can be certain that some students fall into one or another of these categories. In fact, it could be argued that if a teacher teaches a three-period lesson, the only students for whom time is used effectively are those who require neither more nor less than three periods to master the skill or concept being taught (see Slavin, in press a).

The most common means of dealing with the problem of heterogeneity are various forms of ability grouping of classes, such as tracking/curriculum placement. special education, and gifted classes. The purpose of such programs is to reduce classroom heterogeneity so that students' needs can be more efficiently met. However, decades of research on tracking have failed to find achievement benefits of this practice for students as a whole (Esposito, 1973; Good & Marshall, in press; Rosenbaum, 1980). Special education programs tend to be no more effective than regular classes for the achievement of students with mild academic handicaps (Madden & Slavin, 1983), and randomized studies of between-class ability grouping (i.e., tracking) and of gifted programs for high-ability students find few if any benefits for their achievement (Kulik & Kulik, 1982; Slavin, in press b).

Another means of accomodating student heterogeneity is __within-class__ ability grouping (e.g., reading or math groups), in which students are taught in ability-homogeneous subgroups within a larger, more heterogeneous classroom. Within-class ability grouping is virtually universal in elementary school reading instruction, but is less common (though still widely used) in mathematics. In contrast to between-class ability grouping, studies comparing within-class ability grouping to whole-class instruction most often find greater learning in the ability-grouped classes (e.g., Begle, 1975; Dewar, 1963; Jones, 1948; Heathers, 1969.), although these effects are not entirely consistent. If true, the disparity between the achievement effects of within-class and between-class ability grouping is of great importance. One reason for such a difference in effects is suggested in a recent study of within- and between-class ability grouping of fourth graders by Rowan and Miracle (1983), who found that, controlling for student ability, students in lower-track classes received a slower pace of instruction than students in higher-track class, but students in low reading groups received a __faster__ pace of instruction than those in high reading groups (i.e., they covered more reading levels per unit time). Faster pace was found in this study and others (see Brophy, 1974) to be associated with greater learning. Also, low-track classes may be difficult to teach because of concentrations of behavior problems, low morale, and a perception on the part of students that their classmates do not value learning. In contrast, low ability groups in heterogeneous classes may be superior on these counts to low-track classes because students in the low ability groups are still members of a class that has norms for appropriate behavior, is high in morale, and values learning (see Slavin & Karweit, 1984).

A third prevalent means of accomodating diverse learning needs is individualized instruction (Glaser, 1965), in which students work on materials at their own level and rate. Evaluations of individualized instruction in mathematics have not generally found benefits of these strategies for student achievement (Miller, 1976; Schoen, 1976). Slavin (in press a) has argued that these disappointing results are due to problems with management, motivation, and lack of direct instruction characteristic of individualized systems. An individualized instruction model directed at solving these problems by the use of cooperative learning teams and regular teacher-led instruction in small groups, Team Assisted Individualization (TAI), has been found in three recent studies to be effective in increasing mathematics achievement more than traditional group-paced instructional methods (Slavin, Leavey, & Madden, in press; Slavin, Madden, & Leavey, in press).

The present research. The purpose of the present research was to investigate the mathematics achievement effects of three commonly proposed methods of dealing with student heterogeneity: individualized instruction, within-class ability grouping, and whole-class instruction. The strategy adopted was to compare the effects of models typifying each of these levels of individualization, each of which had been found in previous research to be more effective than traditional instruction. The whole-class model was the Missouri Mathematics Program (MMP), an instructional program emphasizing a high ratio of active teaching to seatwork, frequent feedback, smooth transitions between activities, and other features derived from the practices of outstandingly effective traditional teachers (Good, Grouws, & Ebmeier, 1983). The MMP has been found to be more effective than control methods in increasing student achievement (Good & Grouws, 1979). The ability-grouped method, Ability-Grouped Active Teaching (AGAT), was designed for the present study (Slavin & Karweit, 1983) to incorporate the major principles of the MMP in the context of a system employing two ability-homogeneous math groups. Finally, the individualized instruction model was Team Assisted Individualization (TAI), discussed earlier.

The principal purpose of the present research was to compare the three methods described above in terms of their overall effects on student mathematics achievement, and their differential effects in different settings and for students of different levels of prior achievement. The research involved two related experiments conducted at the same time: one in urban classrooms in which heterogeneous class assignments were mandated as part of a desegregation plan, and the second in relatively homogeneous rural classrooms that used between-class ability grouping to reduce the heterogeneity of mathematics classes. It was hypothesized that the instructional programs most directed toward accomodating student heterogeneity, TAI and AGAT, would be most effective in heterogeneous settings and for students furthest from the class mean in past performance, while the group-paced instructional program (MMP) would be most effective in relatively homogeneous settings and for students closest to the class mean in prior performance.

Experiment 1: Methods

Subjects. The subjects in Experiment 1 were 354 students in 16 grade 4-6 classes in one of the Wilmington, Delaware school districts formed as a consequence of an extensive desegregation plan. Approximately 71% of the students were white, 26% were black, and 3% were Asian-American.

Treatments. Classes and their teachers were randomly assigned to three experimental treatments. Teachers in each of the treatments received three hours of training, and were then assisted with implementation in the early weeks of the program, which took place over a period of 18 weeks in Spring, 1983. The treatments were as follows:

1. Missouri Mathematics Program (MMP). The MMP (Good, Grouws, & Ebmeier, 1983) is a whole-class, group-paced mathematics program whose principal features were derived from studies contrasting the teaching strategies used by whole-class teachers whose classes consistently performed well to those used by teachers whose classes did poorly. It consists of a regular sequence of teaching, controlled practice, independent seatwork, and homework, with an emphasis on a high ratio of active teaching to seatwork, teaching mathematics in the context of meaning, frequent questions and feedback, rapid pace of instruction, and management strategies intended to increase student time on-task. The training for the MMP was conducted by Dr. Thomas Good, its principal developer.

2. _Ability Grouped Active Teaching_ (_AGAT_). AGAT was developed for the present study as a means of applying the main principles of the MMP to an ability-grouped method (Slavin & Karweit, 1983). On the basis of an initial test, students in each AGAT class were divided into a high group (about 60% of the students in each class) and a low group (40% of the students). Teachers were instructed to differentiate pace and materials for the two groups, in particular to push the pace for the high group. In most other respects, AGAT was quite similar to the MMP. It emphasized a high ratio of active instruction to seatwork (in each group), teaching mathematics in the context of meaning, frequent questions and feedback, and management strategies (derived in part from the work of Anderson, Evertson, and Brophy, 1979, and Clements & Evertson, 1982) designed to minimize the management problems characteristic of ability-grouped instruction and maintain high time on-task. The training for AGAT was done together by the first author and Dr. Good.

3. _Team Assisted Individualization_ (_TAI_). The individualized model used was Team Assisted Individualization. TAI has been described in detail elsewhere (Slavin, Leavey, & Madden, in press). Briefly, students in TAI worked in heterogeneous four or five-member learning teams on individualized mathematics materials at their own levels and rates. Students within the teams helped one another with problems and took responsibility for almost all checking, routing, and other management tasks inherent in an individualized program. This student management freed the teacher to work with three regularly constituted teaching groups composed of students (drawn from many teams) performing at the same level in the materials. At the end of each week, students on teams that met certain pre-set criteria received attractive certificates. Team rewards of this type have been found in many previous studies (Slavin, 1983a, 1983b) to increase student motivation and achievement.

TAI resembles earlier individualized models (see, for example, Talmage, 1975) in its use of individualized materials that students complete at their own rates. It differs in its emphasis on direct instruction (in homogeneous teaching groups), student management, cooperative learning teams, and cooperative incentives.

Measures

1. Mathematics Achievement. The Mathematics Computations and Concepts and Applications subscales of the Comprehensive Test of Basic Skills (CTBS) were the achievement criterion measures. Fourth graders took Level 2, Form S, while fifth and sixth graders took Level H, Form U. District-administered California Achievement Test scores were used as covariates for their respective CTBS scores. That is, CAT Computations was used as a covariate for CTBS Computations, and CAT Concepts and Applications was a covariate for CTBS Concepts and Applications. Because of the different tests used at different grade levels, all scores were transformed to T scores (Mean = 50, S.D. = 10), and then CTBS scores were adjusted for their corresponding CAT scores using separate linear regressions for each grade. These adjusted scores were used in all subsequent analyses. Note that this adjustment removes any effect of grade level, as the mean for all tests was constrained to be 50 at each grade level.

2. Attitudes. Two eight-item attitude scales were given as pre- and posttests. They were Liking of Math Class (e.g., "This math class is the best part of my school day") and Self-Concept in Math (e.g., "I'm proud of my math work in this class;" "I worry a lot when I have to take a math test"). For each item, students marked either YES!, yes, no, or NO! Scores of negatively phrased items were reversed, so that high scale scores indicated more positive attitudes. Coefficient alpha reliability estimates on these scales were computed in an earlier study (Slavin, Leavey, & Madden, in press) and found to be .86 and .77. respectively. The range of possible scores on both scales was 8 to 32, with a midpoint of 20.

3. Behavioral Observation. Two forms of behavioral observation were used, primarily to determine adequacy of implementation of the various methods. First, TAI, AGAT, and MMP classes were observed to see that the main components of each treatment were in place. However, many of the components of the MMP and AGAT treatments were too subtle for simple implementation checks, so for this reason, more systematic observations were made in these classes. These observations took place for an average of three full mathematics periods per class distributed over the course of the experiment. Behaviors observed included amounts of time spent in active teaching, seatwork, and test-taking, and the percent of time students spent on-task. The observation system was derived from Karweit and Slavin (1981).

Experiment 1: Results

Implementation and Time Use. All teachers were found to be implementing the major components of their methods. AGAT and MMP students spent similar proportions of their instructional time on-task (92.0% and 90.1%, respectively), but as expected, AGAT students spent much more time in seatwork: 54.0% vs. 22.3% for MMP. AGAT teachers spent more time teaching the class (82.3% vs. 73.2%), but of course each student in AGAT received an average of half this amount of teacher instruction.

Achievement. The adjusted CTBS scores were analyzed by means of random-effects nested analysis of variance, similar to analysis of variance using class means. The factors in the analysis were treatment and class/teacher within treatment. The mean square for treatment was tested against that for class/teacher within treatment, which was compared in turn to the within-cells error mean square. If the overall nested analysis of variance was statistically significant ($p < .10$ or better), individual-level planned comparisons between treatment means were computed using a modified Bonferroni procedure (Keppel, 1982) to reduce the possibility of Type II error due to many tests being made on the same data.

Table 1 about here

Table 1 summarizes the means and standard deviations of the CAT and CTBS scales and the adjusted CTBS scores in T scores and grade equivalents. The grade equivalent scores are presented as supplementary information; only the T-scores were used in the analyses. Note that because the grade equivalents were combined across grade levels in the same manner as the T-scores, they are adjusted for differing numbers of students in each grade/treatment cell, and their standard deviations are computed as within-grade standard deviations.

Initial tests for pretest differences using the same statistical procedures as in the main analyses revealed no pretest differences for Computations, but despite random assignment, there were marginally significant differences for Concepts and Applications ($F(2,13) = 2.91$, $p<.09$), due to high pretest scores for TAI classes and low scores for MMP classes. Analyses involving Concepts and Applications must be interpreted cautiously in light of these initial differences.

Table 2 About Here

Table 2 presents the results of the nested analyses of variance for the achievement measures. For Computations, the nested analysis was statistically significant ($F(2, 13) = 6.27$, $p<.012$), but there were no effects on the Concepts and Applications scale. As noted earlier, these random-effects nested analyses are virtually identical to class-level analyses of variance using class means. For example, a class-level analysis for Computations produced $F(2, 13)=7.22$, $p<.008$. Individual comparisons among treatment Computations means revealed that TAI and AGAT means were nearly identical, but both were substantially higher than MMP. Table 2 presents the mean differences both in effect sizes (difference between adjusted means divided by the standard deviation) and in grade equivalent differences. As the table indicates, both TAI and AGAT classes exceeded MMP classes by 75% of a standard deviation, or approximately 1.17 grade equivalents.

Tables 3 and 4 About Here

Attitudes. Table 3 presents the means and standard deviations of the two questionnaire scales. The data were analyzed as for the achievement analyses. There were marginally significant pretest differences on Liking of Math Class ($F(2, 13) = 3.08$, $p < .08$), due to low pretest scores in the MMP classes. No differences were found on Self-Concept pretest scores.

Table 4 summarizes the results of the attitude analyses. For Liking of Math Class, the overall nested analysis was statistically significant ($F(2,13) = 4.06$, $p < .043$). Modified Bonferroni comparisons revealed that the differences were due to low scores in the MMP classes; TAI and AGAT did not differ. On the Self-Concept in Math scale, the nested analysis was also significant ($F(2, 13) = 4.15$, $p < .040$), but in this case, TAI students scored much higher than AGAT and MMP students, who did not differ from one another. On both scales, it is interesting to note that none of the treatment groups improved in attitudes over time; the treatment effects came about more as preventing a negative trend than causing a positive one.

Experiment 2: Methods

Subjects. The subjects in Experiment 2 were 480 students in 22 grade 3-5 classrooms in and around Hagerstown, a town in Western Maryland. Ninety-one percent of the students were white, 7% were black, and 2% were Asian-American.

Treatments. Classes and their teachers were randomly assigned to four experimental treatments. Three of these were the MMP, AGAT, and TAI methods described above; the fourth was an untreated control group. Experiment 2 began two weeks after Experiment 1 and continued for 16 weeks in Spring, 1983. In all other respects the treatments, training, observations, and other procedures were the same as in Experiment 1, except that Dr. Good was unable to participate in the training.

Measures. The measures were the same as in Experiment 1, with students in grades 3 and 4 taking Level 2, Form S of the CTBS and fifth graders taking Level H, Form U. Also, because CAT's are given in the fall in grades 3 and 5 but not grade 4 in Maryland, CAT scores for third and fifth graders were recent, but those used for fourth graders were their third grade scores. Because posttest scores were adjusted for the CAT's separately for each grade level, this should make little difference in the analyses.

Experiment 2: Results

Implementation and Time Use. All TAI, AGAT, and MMP teachers were found to be implementing the major components of their methods. Time use in the AGAT, MMP, and Control classes corresponded closely with expectations. Time on-task was greatest in MMP (94.3% of instructional time), less in Control (87.2%), and least in AGAT (84.4%). Seatwork time was least in MMP (22.6% of instructional time), intermediate in Control (34.6%), and greatest in AGAT (52.1%). AGAT teachers spent the largest amount of time teaching (89.6%), followed by MMP (77.7%) and Control (59.6%), but as noted for Experiment 1, AGAT students each received only about half this amount of teacher instruction.

Achievement. Analyses for Experiment 2 were the same as for Experiment 1. Table 5 summarizes the means and standard deviations of the achievement measures. Pretest differences were not statistically significant for Computations, but despite random assignment, there were statistically significant differences on Concepts and Applications ($F(3, 18) = 3.89$, $p<.026$), due to high pretest scores in AGAT classes and low scores in Control clases.

Tables 5 and 6 about here

As is shown in Table 6, the overall nested analysis of variance was marginally significant for Computations ($F(3,18)=2.71$, $p <.076$). Modified Bonferroni comparisons indicated that as in Experiment 1, AGAT and TAI did not differ in effects on Computations, but both were superior to MMP. All three experimental conditions exceeded Control. Also, as in Experiment 1, the nested analysis of variance for Concepts and Applications did not approach statistical significance ($F(3,18)=1.94$, n.s.).

Tables 7 and 8 about here

Attitudes. Tables 7 and 8 present descriptive statistics and analyses, respectively, for the attitude variables in Experiment 2. There were no statistically significant pretest differences on either attitude scale. For Liking of Math Class, the overall nested analysis of variance was statistically significant ($F(3,18)=5.41$, $p<.008$), with modified Bonferroni comparisons indicating that TAI students significantly exceeded all others, but AGAT and MMP did not differ. MMP students exceeded Control students, but

AGAT did not differ from Control. On Self-Concept in Math, the nested analysis of variance was not significant ($F(3,18)=0.52$, n.s.).

Interactions

It was expected that the different programs evaluated in Experiments 1 and 2 would have different achievement effects for students of different levels of past performance (either absolute or relative to their class averages), race, or sex. Individual-level analyses of covariance were conducted to test for such interactions with treatment. Students were trichotomized on: a) absolute past performance (the mean of their standardized CAT pretests), and b) performance relative to their own class averages. The resulting analyses of variance thus tested both linear and curvilinear effects of past performance (that is, the pretest by treatment interaction would have been significant if high or low achievers benefitted disproportionately from one or another treatment or if high <u>and</u> low achievers benefitted more than average achievers). Because of the great power of the individual-level analyses, an alpha criterion of .01 was chosen for tests for interactions. Contrary to expectations, no interactions between treatment and absolute or relative achievement level, sex, or race were found in either experiment (race by treatment interactions were tested in Experiment 1 only).

Discussion

There is a remarkable degree of commonality of findings regarding student achievement across the two experiments reported here. In both, Team Assisted Individualization (TAI) and Ability Grouped Active Teaching (AGAT) increased computational skills markedly more than the Missouri Mathematics Program (MMP) and, in Experiment 2, an untreated control condition. In neither experiment were any differences between TAI and AGAT found, and in neither were there statistically significant differences in Concepts and Applications. The similarity in achievement effects was particularly striking in light of the differences between the urban, integrated, untracked schools studied in Experiment 1 and the rural, mostly white, tracked schools involved in Experiment 2. Along the same lines, it was surprising to find that the achievement effects were main effects; no statistically significant interactions were found between treatment and either absolute levels of prior achievement or prior achievement relative to class means. These findings contradict the expectation that the effects of TAI and AGAT, programs designed to accomodate diverse achievement levels, would be most positive for students performing farthest from their class means and in settings with the greatest degree of student heterogeneity.

One difficulty with the data from the present study is pretest differences. Despite random assignment, pretest differences were statistically significant for Concepts and Applications in both studies, and though not statistically significant for Computations in either study, they were large enough to warrant caution in interpreting results. When pretests are unequal, statistical controls for pretest tend to underadjust (Lord, 1960). However, in the present studies the <u>magnitude</u> of treatment effects on Computations might be overstated because of pretest differences, but the effects themselves cannot be entirely ascribed to pre-test differences. For example, the difference between TAI and MMP increased over the course of Experiment 1 from .36 standard deviation units to .81 units, and in Experiment 2 from .15 units to .41. The AGAT-MMP difference increased from .14 standard deviation units to .70 in Experiment 1, and from .33 to .62 in Experiment 2. These effects appear too large to be attributable to pretest differences alone.

The findings of the present studies with regard to computations skills have important implications for several research traditions. The success of TAI in increasing student achievement replicated findings of earlier studies evaluating this method (Slavin, Leavey, & Madden, in press; Slavin, Madden, & Leavey, in press), lending further support to Slavin's (in press a) prediction that if inherent problems of management, motivation, and lack of direct instruction could be solved, individualized instruction could be made instructionally effective. The positive effects of TAI also support the utility of student learning teams for student motivation and instructional management (Slavin, 1983b).

The success of the AGAT program was more surprising. While previous research on within-class ability grouping has found more positive than negative effects of this practice on student achievement, these effects have been neither strong nor consistent (Begle, 1975). The substantial positive effects of AGAT seen in the present studies may be due to the specific method of implementing within-class ability grouping, which specified class management strategies derived from the work of Anderson, Evertson, and Brophy (1979) and Clements and Evertson (1982). Within-class ability grouping may be seen as producing gains in instructional effectiveness by providing appropriate levels of instruction that are offset to some degree by losses in instructional effectiveness due to the difficulty of managing multiple ability groups. If these management problems can be solved, within-class ability grouping may be a particularly effective procedure (see Slavin & Karweit, 1984). Finally, the success of the MMP relative to untreated control classes in Experiment 2 replicates earlier work by Good et al. (1983), reemphasizing the importance of active teaching and effective management strategies in the context of group-paced instruction.

The failure to find any statistically significant interactions was unexpected and intriguing. If AGAT and TAI are effective even in part because they increase appropriate levels of instruction for students, they should have especially strong positive effects on the achievement of students farthest from the class mean in prior performance. The fact that the positive effects of these programs were equal for all students might suggest that they are effective not because they accomodate heterogeneity in student preparation and learning rate, but because they provide more effective instruction in general. For example, it is important to note that both AGAT and TAI are highly structured instructional models. Teachers and students know what they are to do at any given moment. In contrast, the MMP involves subtle (though important) changes in teaching practices, such that MMP and Control classes could not be easily discriminated in observations. It may be that the structured nature of the AGAT and TAI treatments contributed to their effectiveness. In fact, it might be argued that these methods, because of the explicit directions for teachers and well-specified routines for students, might be more faithful operationalizations of the principles on which the MMP is based than the MMP itself, and the more positive achievement effects observed for these methods might validate rather than repudiate these principles.

The results of these experiments should give pause to any who might overemphasize the importance of time on-task or of teacher-directed instruction as opposed to seatwork. The issue is more complex than might have been assumed earlier. Comparing MMP to Control, greater teaching time and slightly higher time on-task for the MMP may be at least partially responsible for the superiority of the MMP on Computations in Experiment 2. That is, a high percent of time on active instruction and high time on-task may make a difference _in the context of group-paced instruction_. Outside of that context, though, the effects of time on-task are much less clear. In both experiments, AGAT students spent more than twice as much of their time in seatwork than did MMP students, and in Experiment 2 were much less on-task (probably because of the high amount of seatwork time, which has been associated with low time on-task; see Brophy, 1979). TAI students (who were not observed) probably spent even more of their time in seatwork. Yet TAI and AGAT students learned more than did MMP or Control students. It might well be the case that among AGAT classrooms or among TAI classrooms, the amount of active teaching and time on-task would make a difference in the predicted direction, but comparing across methods is another matter. In traditional group-paced teaching, seatwork may be less than optimally effective not because it reduces time on-task but because it is rarely well adapted to the needs of various students, many of whom are either practicing errors or working problems they already know how to do (see Anderson, 1981). In AGAT and especially TAI, seatwork is likely to be more productive, as it is closely keyed to students' individual needs.

The failure in both studies to find any treatment effects on Concepts and Applications contrasts sharply with the findings for Computations. This discrepancy, which is not uncommon in research on mathematics instruction, is probably due to relatively low correspondence between what is taught in school and what is tested by the Concepts and Applications scales. While virtually all skills tested on the Computations scale are taught in school, Concepts and Applications tests include many items that appear to tap general aptitude rather than school learning, as well as many word problems that depend as much on reading skill as on mathematics knowledge.

At least as far as computational skills are concerned, the results of the present studies provide striking evidence that methods which include means of adapting instruction to diverse needs can be considerably more effective than methods that do not. However, the success of TAI and AGAT should emphatically not be interpreted as justification of individualized instruction or within-class ability grouping in general. Previous research has clearly indicated that individualized instruction as usually structured is no more effective than traditional methods in mathematics (Miller, 1976; Schoen, 1976), and within-class ability grouping is only inconsistently more effective than traditional methods (Begle, 1975). What the results of the present studies do mean is that if problems of management and motivation inherent in attempts to accomodate student heterogeneity can be solved, then such methods may be able to achieve the outcomes for which they were designed--enhanced achievement for all students.

In both experiments, TAI was associated with more positive student attitudes toward math class than was MMP or (in Experiment 2) untreated control methods, and in Experiment 2, TAI was also significantly superior to AGAT on this variable. TAI students reported more positive self-concepts in mathematics than AGAT or MMP students in Experiment 1, but there were no differences on this variable in Experiment 2. The positive effects of TAI on student attitudes replicate findings for cooperative learning methods in general (Slavin, 1983a) and for TAI in particular (Slavin, Leavey, & Madden, in press). In addition to positive attitudes expressed by students, teachers also responded favorably to the TAI program. On questionnaires given at the end of Experiments 1 and 2, more than half of all TAI teachers indicated that "TAI was by far the best method I've used to teach math." Not one of the AGAT teachers and only one-third of the MMP teachers agreed with similar statements about their methods. At the end of the experiments, teachers were allowed to choose any method other than the one they had used before in which to receive training and materials. Every eligible teacher chose TAI, and every TAI teacher has continued to use the program during the next (1983-84) school year. None of the AGAT teachers continued to use the program. One possible reason for the relative unpopularity of the AGAT treatment is the amount of teacher work required; 88% of the AGAT teachers felt that AGAT required more work from them than their usual methods, while only 40% of TAI teachers and 25% of MMP teachers responded this way.

The outcomes of the present studies suggest many directions for further research. First, the effects themselves (particularly the effects of AGAT) should be replicated in field experiments involving larger numbers of teachers and classes at each grade level--combining across grade levels, necessitated by the small number of classes involved, may have obscured important developmental trends. Also, the unfortunate pretest differences on some measures (despite random assignment) could account for a portion of the effects observed. Adjustment of posttests for pretest scores only partially solves this problem; further replication is clearly needed. Second, more detailed and extensive observations of treatment implementation and collection of data on such variables as student motivation and perceptions would allow for better understanding of how and why the various methods produce their effects. Finally, component analyses of the complex TAI and AGAT programs are needed to establish which elements of these programs affect student achievement.

One major drawback of recent research on teaching (e.g., Brophy, 1979) is that it largely consists of correlational (process-product) studies relating teaching practices to student achievement outcomes in traditionally taught classrooms. Although this line of research has added much to our understanding of effective teaching practices, its correlational nature makes it subject to errors in understanding direction of causality, and its restriction to traditionally taught classrooms limits its prescriptions to the range of current widespread practice. Slavin (in press a) has called for a movement in research on teaching toward experimental studies evaluating alterable components of instruction. The present research represents one step in this direction, focusing on means of accomodating heterogeneity in mathematics instruction. Much work of this kind remains to be done before we will have a true scientific basis for instructional practice.

References

Anderson, L.M. (1981). Student response to seatwork: Implications for the study of students' cognitive processing. Paper presented at the Annual Convention of the American Educational Research Association, Los Angeles.

Anderson, L.M., Evertson, C., & Brophy, J. (1979). An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 79, 193-223.

Begle, E. (1975). Ability grouping for mathematics instruction: A review of the empirical literature. Stanford University, Stanford Mathematics Education Study Group.

Brophy, J.E. (1979). Teacher behavior and its effects. Journal of Educational Psychology, 71, 733-750.

Clements, B., & Evertson, C. (1982). Orchestrating small group instruction in elementary school classrooms. (Technical Report No. 6053). University of Texas, Research & Development Center for Teacher Education.

Dewar, J. (1963). Grouping for arithmetic instruction in sixth grade. Elementary School Journal, 63, 266-269.

Esposito, D. (1973). Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. *Review of Educational Research*, *43*, 163-179.

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

Good, T., & Grouws, D. (1979). The Missouri Mathematics Effectiveness Project: An experimental study in fourth grade classrooms. *Journal of Educational Psychology*, *71*, 355-362.

Glass, G., & Stanley, J. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Good, T., Grouws, D., & Ebmeier, H. (1983). *Active mathematics teaching*. New York: Longman.

Good, T., & Marshall, S. (in press). Do students learn more in heterogeneous or homogeneous groups? In P.E. Peterson, L.C. Wilkinson, & M.T. Hallinan (Eds.), *Student diversity and the organization, process, and use of instructional groups in the classroom*. New York: Academic Press.

Heathers, G. (1969). Grouping. In R. Ebel (Ed.) Encyclopedia of Educational Research (4th ed.). New York: Macmillan.

Hopkins, K. (1982). The unit of analysis: Group means versus individual observations. American Educational Research Journal, 19, 5-18.

Jones, D. M. (1948). An experiment in adaptation to individual differences. Journal of Educational Psychology, 39, 257-273.

Karweit, N., & Slavin, R.E. (1981). Measurement and modeling choices in studies of time and learning. American Educational Research Journal, 18, 157-171.

Keppel, G. (1982). Design and analysis. Englewood Cliffs, NJ: Prentice-Hall.

Kulik, C. L., & Kulik, J. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. American Educational Research Journal, 19, 415-428.

Lord, F. (1960). Large-sample covariance analysis when the control variable is fallible. Journal of the American Statistical Association, 55, 307-321.

Madden, N.A., & Slavin, R.E. (1983). Mainstreaming students with mild academic handicaps: Academic and social outcomes. _Review of Educational Research_, _53_, 519-569.

Miller, R. L. (1976). Individualized instruction in mathematics: A review of research. _Mathematics Teacher_, _69_, 345-351.

Rosenbaum, J. (1980). Social implications of educational grouping. _Review of Research in Education_, _8_, 361-401.

Rowan, B., & Miracle, A. (1983). Systems of ability grouping and the stratification of achievement in elementary schools. _Sociology of Education_, _56_, 133-144.

Schoen, H.L. (1976). Self-paced mathematics instruction: How effective has it been? _Arithmetic Teacher_, _23_, 90-96.

Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. _Review of Educational Research_, _50_, 241-271.

Slavin, R.E. (1980). _Cooperative learning_. New York: Longman.

Slavin, R. E. (1983a). When does cooperative learning increase student achievement? *Psychological Bulletin*, *94*, 429-445.

Slavin, R.E. (in press a). Component building: A strategy for research-based instructional improvement. *Elementary School Journal*.

Slavin, R.E. (in press b). Meta-analysis in education: How has it been used? *Educational Researcher*.

Slavin, R.E., & Karweit, N. (1983). *Ability grouped active teaching* (AGAT): *Teacher's manual*. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Slavin, R.E., & Karweit, N. (1984). *Within-class ability grouping and student achievement: Two field experiments*. Paper presented at the Annual Convention of the American Educational Research Association, New Orleans.

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., Madden, N.A., & Leavey, M. (in press). Effects of Team Assisted Individualization on the mathematics achievement of academically handicapped and non-handicapped students. *Journal of Educational Psychology*.

Talmage, H. (Ed.). (1975). *Systems of individualized education*. Berkeley, CA: McCutchan.

![ERIC logo]

Table 1

Means and Standard Deviations in T-Scores and Grade Equivalents, Mathematics Achievement Measures, Experiment 1

		Computations			Concepts and Applications		
		CAT(Pre)	CTBS(Post)	Adjusted	CAT(Pre)	CTBS(Post)	Adjusted
TAI	T	52.03	53.50	52.66	54.23	53.00	49.84
	S	(10.73)	(9.83)	(8.86)	(10.78)	(10.84)	(6.87)
	G.E.	5.51	7.16	7.01	6.07	7.12	6.57
	S	(1.12)	(1.74)	(1.56)	(1.53)	(1.87)	(1.19)
	N		122			123	
AGAT	T	49.77	52.48	52.67	49.11	49.77	50.41
	S	(10.21)	(9.60)	(8.55)	(8.65)	(8.77)	(7.11)
	G.E.	5.28	6.98	7.01	5.34	6.56	6.67
	S	(1.07)	(1.69)	(1.51)	(1.23)	(1.51)	(1.23)
	N		89			89	
MMP	T	48.40	45.44	46.04	46.90	47.55	49.88
	S	(8.85)	(8.51)	(7.60)	(8.71)	(9.24)	(6.51)
	G.E.	5.13	5.74	5.84	5.03	6.18	6.58
	S	(0.92)	(1.50)	(1.34)	(1.24)	(1.59)	(1.13)
	N		142			142	
TOTAL	T	50.00	50.00	50.00	50.00	50.00	50.00
	S	(10.00)	(10.00)	(8.89)	(10.00)	(10.00)	(6.77)
	G.E.	5.30	6.54	6.54	5.47	6.60	6.60
	S	(1.04)	(1.76)	(1.57)	(1.42)	(1.72)	(1.17)
	N		353			354	

Note: Table entries are T scores (Mean=50, S.D.=10) computed separately for each grade level. Adjusted scores are CTBS (Post) scores adjusted for CAT (Pre) scores separately for each grade level.

Table 2

Results of Nested Analyses of Variance, Adjusted Mathematics Achievement, Experiment 1

ce of ation	Computations				Concepts and Applications			
	d.f.	M.S.	F	p <	d.f.	M.S.	F	p <
tment	2	18.61	6.27	.012	2	0.10	0.13	n.s.
s/Teacher in Treatment	13	2.97	4.95	.001	13	0.77	1.73	.054
r (within s)	337	0.60			338	0.45		

erences between Adjusted Means in ct Sizes and (Grade Equivalents), Computations

	TAI	AGAT	MMP
	---	.00 (.00)	+ .75*** (+1.17)
		---	+ .75*** (+1.17)

.001 Using Modified Bonferroni Procedure

Table 3

Means and Standard Deviations, Attitude Scales, Experiment 1

		Liking of Math Class			Self-Concept in Math		
		Pre	Post	Adjusted	Pre	Post	Adjusted
TAI	$\overline{X}$	25.15	25.16	24.67	24.10	25.01	24.70
	S	5.35	4.78	4.06	4.65	4.11	3.35
	N		120			117	
AGAT	$\overline{X}$	24.99	24.48	24.06	24.15	23.31	22.98
	S	5.04	4.90	4.41	4.12	4.44	3.57
	N		86			80	
MMP	$\overline{X}$	22.71	21.65	22.35	22.86	22.68	23.13
	S	5.96	6.01	5.42	4.66	5.10	4.09
	N		135			139	
TOTAL	$\overline{X}$	24.14	23.60	23.60	23.60	23.64	23.64
	S	5.63	5.55	4.83	4.56	4.72	3.80
	N		341			336	

Note: Table entries are raw scores from 8-item attitude scales. Range of possible scores: 8-32. Adjusted scores are posttests adjusted for pretests.

Table 4

Results of Nested Analyses of Variance,
Adjusted Attitude Scales, Experiment 1

Source of Variation	Liking of Math Class d.f.	M.S.	F	$p <$	Self-Concept in Math d.f.	M.S.	F	$p <$
Treatment	2	183.74	4.06	.043	2	101.86	4.15	.040
Class/Teacher Within Treatment	13	45.28	2.11	.013	13	24.52	1.82	.039
Error (within cells)	325	21.43			320	13.45		

Differences between Adjusted Means in Effect Sizes, Liking of Math Class

	TAI	AGAT	MMP
TAI	---	+.13	+.48***
AGAT		---	+.35***
MMP			---

Differences between Adjusted Means in Effect Sizes, Self-Concept in Math

	TAI	AGAT	MMP
TAI	---	+.45***	+.41**
AGAT		---	-.04
MMP			---

$p <$ .01 Using Modified Bonferroni Procedure

$p <$ .001 Using Modified Bonferroni Procedure

Table 5

Means and Standard Deviations in T-Scores
and Grade Equivalents, Mathematics Achievement Measures, Experiment 2

		Computations			Concepts and Applications		
		CAT(Pre)	CTBS(Post)	Adjusted	CAT(Pre)	CTBS(Post)	Adjusted
TAI	T	51.35	52.92	51.83	52.21	51.07	49.27
	S	(9.05)	(9.64)	(7.94)	(8.96)	(9.15)	(7.13)
	G.E.	4.64	6.31	6.14	5.36	6.03	5.71
	S	(0.79)	(1.51)	(1.24)	(1.16)	(1.62)	(1.27)
	N		112			114	
AGAT	T	53.16	54.99	53.49	55.40	57.05	53.31
	S	(8.81)	(10.12)	(8.58)	(10.46)	(9.90)	(7.26)
	G.E.	4.79	6.63	6.40	5.77	7.09	6.43
	S	(0.77)	(1.58)	(1.34)	(1.35)	(1.76)	(1.29)
	N		98			98	
MMP	T	49.84	48.80	48.92	48.19	47.92	49.21
	S	(10.76)	(8.24)	(7.54)	(8.91)	(9.10)	(6.55)
	G.E.	4.51	5.66	5.68	4.84	5.47	5.70
	S	(0.94)	(1.29)	(1.18)	(1.15)	(1.61)	(1.16)
	N		162			162	
CONTROL	T	45.91	44.15	46.49	45.39	45.51	48.93
	S	(9.41)	(9.29)	(7.97)	(9.27)	(8.48)	(6.01)
	G.E.	4.16	4.94	5.30	4.48	5.05	5.65
	S	(0.82)	(1.45)	(1.25)	(1.20)	(1.50)	(1.07)
	N		106			106	
TOTAL	T	50.00	50.00	50.00	50.00	50.00	50.00
	S	(10.00)	(10.00)	(8.32)	(10.00)	(10.00)	(6.92)
	G.E.	4.52	5.85	5.85	5.07	5.84	5.84
	S	(0.87)	(1.56)	(1.30)	(1.29)	(1.77)	(1.23)
	N		478			480	

Note: Table entries are T scores (Mean=50,S.D.=10) computed separately for each grade level. Adjusted scores are CTBS (Post) scores adjusted f[illegible] CAT (Pre) scores separately for each grade level.

Table 6

Results of Nested Analyses of Variance, Adjusted Mathematics Achievement, Experiment 2

Source of Variation	Computations d.f.	M.S.	F	p<	Concepts and Applications d.f.	M.S.	F	p<
Treatment	3	10.22	2.71	.076	3	4.52	1.94	n.s.
Class/Teacher Within Treatment	18	3.78	7.43	.001	18	2.33	6.16	.001
Error (within cells)	456	0.51			458	0.38		

Differences between Adjusted Means in Effect Sizes and (Grade Equivalents), Computations

	TAI	AGAT	MMP	CONTROL
TAI	---	-.20 (-.26)	+.35** (+.46)	+ .64*** (+ .84)
AGAT		---	+.55*** (+.72)	+ .84*** (+1.10)
MMP			---	+ .29* (+ .38)
CONTROL				---

*p<.05 using Modified Bonferroni Procedure
**p<.01 using Modified Bonferroni Procedure
***p<.001 using Modified Bonferroni Procedure

Table 7

Means and Standard Deviations,
Attitude Scales, Experiment 2

		Liking of Math Class			Self-Concept in Math		
		Pre	Post	Adjusted	Pre	Post	Adjusted
TAI	$\bar{X}$	24.61	26.46	26.90	24.71	24.50	24.24
	S	5.91	4.93	4.51	4.60	4.67	4.04
	N		95			96	
AGAT	$\bar{X}$	26.07	24.76	24.30	25.07	24.83	24.36
	S	4.77	5.34	4.45	4.58	4.52	3.55
	N		90			89	
MMP	$\bar{X}$	25.46	25.33	25.24	23.84	24.56	24.81
	S	4.85	5.44	4.49	4.64	4.80	3.81
	N		153			157	
Control	$\bar{X}$	25.07	22.81	22.97	23.78	24.88	25.17
	S	5.99	6.63	4.54	4.57	4.29	3.40
	N		86			93	
TOTAL	$\bar{X}$	25.32	24.95	24.95	24.27	24.67	24.67
	S	5.33	5.69	4.67	4.62	4.60	3.73
	N		424			435	

Note: Table entries are raw scores from 8-item attitude scales. Range of possible scores: 8-32. Adjusted scores are posttests adjusted for pretests.

Table 8

Results of Nested Analyses of Variance, Adjusted Attitude Scales, Experiment 2

Source of Variation	Liking of Math Class d.f.	M.S.	F	$p <$	Self-Concept in Math d.f.	M.S.	F	$p <$
Treatment	3	249.70	5.41	.008	3	17.38	0.52	n.s.
Class/Teacher Within Treatment	18	46.17	2.42	.001	18	33.38	2.56	.001
Error (within cells)	402	19.04			413	13.03		

Differences between Adjusted Means in Effect Sizes, Liking of Math Class

	TAI	AGAT	MMP	CONTROL
AI	---	+.56***	+.36**	+.84***
GAT		---	-.20	+.28
MP			---	+.49***
ONTROL				---

** $p < .01$ Using Modified Bonferroni Procedure
** $p < .001$ Using Modified Bonferroni Procedure

about their methods. At the end of the experiments, teachers were allowed to choose any method other than the one they had used before in which to receive training and materials. Every eligible teacher chose TAI, and every TAI teacher has continued to use the program during the next (1983-84) school year. None of the AGAT teachers continued to use the program. One possible reason for the relative unpopularity of the AGAT treatment is the amount of teacher work required; 88% of the AGAT teachers felt that AGAT required more work from them than their usual methods, while only 40% of TAI teachers and 25% of MMP teachers responded this way.

The outcomes of the present studies suggest many directions for further research. First, the effects themselves (particularly the effects of AGAT) should be replicated in field experiments involving larger numbers of teachers and classes at each grade level--combining across grade levels, necessitated by the small number of classes involved, may have obscured important developmental trends. Also, the unfortunate pretest differences on some measures (despite random assignment) could account for a portion of the effects observed. Adjustment of posttests for pretest scores only partially solves this problem; further replication is clearly needed. Second, more detailed and extensive observations of treatment implementation and collection of data on such variables as student motivation and perceptions would allow for better understanding of how and why the various methods produce their effects. Finally, component analyses of the complex TAI and AGAT programs are needed to establish which elements of these programs affect student achievement.

One major drawback of recent research on teaching (e.g., Brophy, 1979) is that it largely consists of correlational (process-product) studies relating teaching practices to student achievement outcomes in traditionally taught classrooms. Although this line of research has added much to our understanding of effective teaching practices, its correlational nature makes it subject to errors in understanding direction of causality, and its restriction to traditionally taught classrooms limits its prescriptions to the range of current widespread practice. Slavin (in press a) has called for a movement in research on teaching toward experimental studies evaluating alterable components of instruction. The present research represents one step in this direction, focusing on means of accomodating heterogeneity in mathematics instruction. Much work of this kind remains to be done before we will have a true scientific basis for instructional practice.

References

Anderson, L.M. (1981). _Student response to seatwork: Implications for the study of students' cognitive processing_. Paper presented at the Annual Convention of the American Educational Research Association, Los Angeles.

Anderson, L.M., Evertson, C., & Brophy, J. (1979). An experimental study of effective teaching in first-grade reading groups. _Elementary School Journal_, _79_, 193-223.

Begle, E. (1975). _Ability grouping for mathematics instruction: A review of the empirical literature_. Stanford University, Stanford Mathematics Education Study Group.

Brophy, J.E. (1979). Teacher behavior and its effects. _Journal of Educational Psychology_, _71_, 733-750.

Clements, B., & Evertson, C. (1982). _Orchestrating small group instruction in elementary school classrooms_. (Technical Report No. 6053). University of Texas, Research & Development Center for Teacher Education.

Dewar, J. (1963). Grouping for arithmetic instruction in sixth grade. _Elementary School Journal_, _63_, 266-269.

Esposito, D. (1973). Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. Review of Educational Research, 43, 163-179.

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

Good, T., & Grouws, D. (1979). The Missouri Mathematics Effectiveness Project: An experimental study in fourth grade classrooms. Journal of Educational Psychology, 71, 355-362.

Glass, G., & Stanley, J. (1970). Statistical methods in education and psychology. Englewood Cliffs, NJ: Prentice-Hall.

Good, T., Grouws, D., & Ebmeier, H. (1983). Active mathematics teaching. New York: Longman.

Good, T., & Marshall, S. (in press). Do students learn more in heterogeneous or homogeneous groups? In P.E. Peterson, L.C. Wilkinson, & M.T. Hallinan (Eds.), Student diversity and the organization, process, and use of instructional groups in the classroom. New York: Academic Press.

Heathers, G. (1969). Grouping. In R. Ebel (Ed.) Encyclopedia of Educational Research (4th ed.). New York: Macmillan.

Hopkins, K. (1982). The unit of analysis: Group means versus individual observations. American Educational Research Journal, 19, 5-18.

Jones, D. M. (1948). An experiment in adaptation to individual differences. Journal of Educational Psychology, 39, 257-273.

Karweit, N., & Slavin, R.E. (1981). Measurement and modeling choices in studies of time and learning. American Educational Research Journal, 18, 157-171.

Keppel, G. (1982). Design and analysis. Englewood Cliffs, NJ: Prentice-Hall.

Kulik, C. L., & Kulik, J. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. American Educational Research Journal, 19, 415-428.

Lord, F. (1960). Large-sample covariance analysis when the control variable is fallible. Journal of the American Statistical Association, 55, 307-321.

Madden, N.A., & Slavin, R.E. (1983). Mainstreaming students with mild academic handicaps: Academic and social outcomes. Review of Educational Research, 53, 519-569.

Miller, R. L. (1976). Individualized instruction in mathematics: A review of research. Mathematics Teacher, 69, 345-351.

Rosenbaum, J. (1980). Social implications of educational grouping. Review of Research in Education, 8, 361-401.

Rowan, B., & Miracle, A. (1983). Systems of ability grouping and the stratification of achievement in elementary schools. Sociology of Education, 56, 133-144.

Schoen, H.L. (1976). Self-paced mathematics instruction: How effective has it been? Arithmetic Teacher, 23, 90-96.

Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. Review of Educational Research, 50, 241-271.

Slavin, R.E. (1980). Cooperative learning. New York: Longman.

Slavin, R. E. (1983a). When does cooperative learning increase student achievement? Psychological Bulletin, 94, 429-445.

Slavin, R.E. (in press a). Component building: A strategy for research-based instructional improvement. Elementary School Journal.

Slavin, R.E. (in press b). Meta-analysis in education: How has it been used? Educational Researcher.

Slavin, R.E., & Karweit, N. (1983). Ability grouped active teaching (AGAT): Teacher's manual. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Slavin, R.E., & Karweit, N. (1984). Within-class ability grouping and student achievement: Two field experiments. Paper presented at the Annual Convention of the American Educational Research Association, New Orleans.

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., Madden, N.A., & Leavey, M. (in press). Effects of Team Assisted Individualization on the mathematics achievement of academically handicapped and non-handicapped students. Journal of Educational Psychology.

Talmage, H. (Ed.). (1975). Systems of individualized education. Berkeley, CA: McCutchan.

Table 1

Means and Standard Deviations in T-Scores
and Grade Equivalents, Mathematics Achievement Measures, Experiment 1

		Computations			Concepts and Applications		
		CAT(Pre)	CTBS(Post)	Adjusted	CAT(Pre)	CTBS(Post)	Adjusted
TAI	T	52.03	53.50	52.66	54.23	53.00	49.84
	S	(10.73)	(9.83)	(8.86)	(10.78)	(10.84)	(6.87)
	G.E.	5.51	7.16	7.01	6.07	7.12	6.57
	S	(1.12)	(1.74)	(1.56)	(1.53)	(1.87)	(1.19)
	N		122			123	
AGAT	T	49.77	52.48	52.67	49.11	49.77	50.41
	S	(10.21)	(9.60)	(8.55)	(8.65)	(8.77)	(7.11)
	G.E.	5.28	6.98	7.01	5.34	6.56	6.67
	S	(1.07)	(1.69)	(1.51)	(1.23)	(1.51)	(1.23)
	N		89			89	
MMP	T	48.40	45.44	46.04	46.90	47.55	49.88
	S	(8.85)	(8.51)	(7.60)	(8.71)	(9.24)	(6.51)
	G.E.	5.13	5.74	5.84	5.03	6.18	6.58
	S	(0.92)	(1.50)	(1.34)	(1.24)	(1.59)	(1.13)
	N		142			142	
TOTAL	T	50.00	50.00	50.00	50.00	50.00	50.00
	S	(10.00)	(10.00)	(8.89)	(10.00)	(10.00)	(6.77)
	G.E.	5.30	6.54	6.54	5.47	6.60	6.60
	S	(1.04)	(1.76)	(1.57)	(1.42)	(1.72)	(1.17)
	N		353			354	

Note: Table entries are T scores (Mean=50, S.D.=10) computed separately fo[r] each grade level. Adjusted scores are CTBS (Post) scores adjusted [for] CAT (Pre) scores separately for each grade level.

Table 2

Results of Nested Analyses of Variance, Adjusted Mathematics Achievement, Experiment 1

	Computations				Concepts and Applications			
urce of iriation	d.f.	M.S.	F	p <	d.f.	M.S.	F	p <
eatment	2	18.61	6.27	.012	2	0.10	0.13	n.s.
.ass/Teacher .thin Treatment	13	2.97	4.95	.001	13	0.77	1.73	.054
:ror (within :lls)	337	0.60			338	0.45		

:fferences between Adjusted Means in 'fect Sizes and (Grade Equivalents), Computations

	TAI	AGAT	MMP
ɪI	---	.00 (.00)	+ .75*** (+1.17)
;AT		---	+ .75*** (+1.17)
ɪP			---

< .001 Using Modified Bonferroni Procedure

Table 3

Means and Standard Deviations,
Attitude Scales, Experiment 1

		Liking of Math Class			Self-Concept in Math		
		Pre	Post	Adjusted	Pre	Post	Adjusted
TAI	$\overline{X}$	25.15	25.16	24.67	24.10	25.01	24.70
	S	5.35	4.78	4.06	4.65	4.11	3.35
	N		120			117	
AGAT	$\overline{X}$	24.99	24.48	24.06	24.15	23.31	22.98
	S	5.04	4.90	4.41	4.12	4.44	3.57
	N		86			80	
MMP	$\overline{X}$	22.71	21.65	22.35	22.86	22.68	23.13
	S	5.96	6.01	5.42	4.66	5.10	4.09
	N		135			139	
TOTAL	$\overline{X}$	24.14	23.60	23.60	23.60	23.64	23.64
	S	5.63	5.55	4.83	4.56	4.72	3.80
	N		341			336	

Note: Table entries are raw scores from 8-item attitude scales. Range of possible scores: 8-32. Adjusted scores are posttests adjusted for pretests.

Table 4

Results of Nested Analyses of Variance, Adjusted Attitude Scales, Experiment 1

Source of Variation	Liking of Math Class d.f.	M.S.	F	p <	Self-Concept in Math d.f.	M.S.	F
Treatment	2	183.74	4.06	.043	2	101.86	4.15
Class/Teacher Within Treatment	13	45.28	2.11	.013	13	24.52	1.82
Error (within cells)	325	21.43			320	13.45	

Differences between Adjusted Means in Effect Sizes, Liking of Math Class

	TAI	AGAT	MMP
TAI	---	+.13	+.48***
AGAT		---	+.35***
MMP			---

Differences between Adjust in Effect Sizes, Self-Con

	TAI	AGAT
TAI	---	+.45***
AGAT		---
MMP		

$p < .01$ Using Modified Bonferroni Procedure

$p < .001$ Using Modified Bonferroni Procedure

Table 5

Means and Standard Deviations in T-Scores
and Grade Equivalents, Mathematics Achievement Measures, Experiment 2

		Computations			Concepts and Applications		
		CAT(Pre)	CTBS(Post)	Adjusted	CAT(Pre)	CTBS(Post)	Adjusted
TAI	T	51.35	52.92	51.83	52.21	51.07	49.27
	S	(9.05)	(9.64)	(7.94)	(8.96)	(9.15)	(7.13)
	G.E.	4.64	6.31	6.14	5.36	6.03	5.71
	S	(0.79)	(1.51)	(1.24)	(1.16)	(1.62)	(1.27)
	N		112			114	
AGAT	T	53.16	54.99	53.49	55.40	57.05	53.31
	S	(8.81)	(10.12)	(8.58)	(10.46)	(9.90)	(7.26)
	G.E.	4.79	6.63	6.40	5.77	7.09	6.43
	S	(0.77)	(1.58)	(1.34)	(1.35)	(1.76)	(1.29)
	N		98			98	
MMP	T	49.84	48.80	48.92	48.19	47.92	49.21
	S	(10.76)	(8.24)	(7.54)	(8.91)	(9.10)	(6.55)
	G.E.	4.51	5.66	5.68	4.84	5.47	5.70
	S	(0.94)	(1.29)	(1.18)	(1.15)	(1.61)	(1.16)
	N		162			162	
CONTROL	T	45.91	44.15	46.49	45.39	45.51	48.93
	S	(9.41)	(9.29)	(7.97)	(9.27)	(8.48)	(6.01)
	G.E.	4.16	4.94	5.30	4.48	5.05	5.65
	S	(0.82)	(1.45)	(1.25)	(1.20)	(1.50)	(1.07)
	N		106			106	
TOTAL	T	50.00	50.00	50.00	50.00	50.00	50.00
	S	(10.00)	(10.00)	(8.32)	(10.00)	(10.00)	(6.92)
	G.E.	4.52	5.85	5.85	5.07	5.84	5.84
	S	(0.87)	(1.56)	(1.30)	(1.29)	(1.77)	(1.23)
	N		478			480	

Note: Table entries are T scores (Mean=50,S.D.=10) computed separately f each grade level. Adjusted scores are CTBS (Post) scores adjusted CAT (Pre) scores separately for each grade level.

Table 6

Results of Nested Analyses of Variance, Adjusted Mathematics Achievement, Experiment 2

Source of Variation	Computations d.f.	M.S.	F	p <	Concepts and Applications d.f.	M.S.	F	p <
Treatment	3	10.22	2.71	.076	3	4.52	1.94	n.s.
Class/Teacher Within Treatment	18	3.78	7.43	.001	18	2.33	6.16	.001
Error (within cells)	456	0.51			458	0.38		

Differences between Adjusted Means in Effect Sizes and (Grade Equivalents), Computations

	TAI	AGAT	MMP	CONTROL
TAI	---	-.20 (-.26)	+.35** (+.46)	+ .64*** (+ .84)
AGAT		---	+.55*** (+.72)	+ .84*** (+1.10)
MMP			---	+ .29* (+ .38)
CONTROL				---

* $p < .05$ using Modified Bonferroni Procedure

** $p < .01$ using Modified Bonferroni Procedure

*** $p < .001$ using Modified Bonferroni Procedure

Table 7

Means and Standard Deviations, Attitude Scales, Experiment 2

		Liking of Math Class			Self-Concept in Math		
		Pre	Post	Adjusted	Pre	Post	Adjusted
TAI	$\bar{X}$	24.61	26.46	26.90	24.71	24.50	24.24
	S	5.91	4.93	4.51	4.60	4.67	4.04
	N		95			96	
AGAT	$\bar{X}$	26.07	24.76	24.30	25.07	24.83	24.36
	S	4.77	5.34	4.45	4.58	4.52	3.55
	N		90			89	
MMP	$\bar{X}$	25.46	25.33	25.24	23.84	24.56	24.81
	S	4.85	5.44	4.49	4.64	4.80	3.81
	N		153			157	
Control	$\bar{X}$	25.07	22.81	22.97	23.78	24.88	25.17
	S	5.99	6.63	4.54	4.57	4.29	3.40
	N		86			93	
TOTAL	$\bar{X}$	25.32	24.95	24.95	24.27	24.67	24.67
	S	5.33	5.69	4.67	4.62	4.60	3.73
	N		424			435	

Note: Table entries are raw scores from 8-item attitude scales. Range of possible scores: 8-32. Adjusted scores are posttests adjusted for pretests.

Table 8

Results of Nested Analyses of Variance,
Adjusted Attitude Scales, Experiment 2

Source of Variation	Liking of Math Class				Self-Concept in Math		
	d.f.	M.S.	F	p <	d.f.	M.S.	F
Treatment	3	249.70	5.41	.008	3	17.38	0.52
Class/Teacher Within Treatment	18	46.17	2.42	.001	18	33.38	2.56
Error (within cells)	402	19.04			413	13.03	

Differences between Adjusted Means
in Effect Sizes, Liking of Math Class

	TAI	AGAT	MMP	CONTROL
TAI	---	+.56***	+.36**	+.84***
AGAT		---	-.20	+.28
MMP			---	+.49***
CONTROL				---

** p < .01 Using Modified Bonferroni Procedure

** p < .001 Using Modified Bonferroni Procedure