DOCUMENT RESUME

ED 378 260 UD 030 233

AUTHOR	Reuman, David A.; Mac Iver, Douglas J.
TITLE	Effects of Instructional Grouping on Seventh Graders' Academic Motivation and Achievement.
INSTITUTION	Center for Research on Effective Schooling for Disadvantaged Students, Baltimore, MD.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CDS-R-50
PUB DATE	Nov 94
CONTRACT	R117R90002
NOTE	31p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Ability; *Academic Achievement; Achievement Tests; *Cooperative Learning; Grade 7; *Grouping (Instructional Purposes); Junior High Schools; *Junior High School Students; Language Arts; Mathematics Instruction; Science Instruction; Social Studies; Standardized Tests; *Student Motivation
IDENTIFIERS	*Student Team Learning

ABSTRACT

This study tested the hypothesis that cooperative learning practices in heterogeneous classes would enhance students' academic motivation and achievement. Half the seventh graders entering a junior high school were randomly assigned to hetergenous classrooms in which their teachers implemented Student Team Learning (STL) techniques (Slavin, 1986). The remaining students were assigned to homogeneous ability-grouped classrooms. The use of STL in heterogeneous classrooms produced achievement benefits, with effects stronger for course grades and end-of-year grades than for results on a nationally-standardized achievement test or a state criterion-referenced test. The effects of STL in heterogeneous classes varied by subject area, with consistent positive effects in language arts, math, and science, and inconsistent effects in mathematics and social studies. The effects of STL on motivation also varied by subject, with STL raising motivational variables in language arts, math, and science, but not in social studies. Such variations can be attributed to variations in teachers' implementation of STL. Ten tables present study findings, and an appendix lists indicators of motivational constructs. (Contains 32 references.) (SLD)

UD

CDS

THE JOHNS HOPKINS UNIVERSITY

Effects of Instructional Grouping on Seventh Graders' Academic Motivation and Achievement

David A. Reuman

Douglas J. Mac Iver

Report No. 50

November 1994

CENTER FOR RESEARCH ON EFFECTIVE SCHOOLING FOR DISADVANTAGED STUDENTS

National Advisory Panel

Beatriz Arias, College of Education, Arizona State University

Mary Frances Berry, Department of History, University of Pennsylvania

Anthony S. Bryk, Department of Education, University of Chicago

Edmund Gordon (Chair), Department of Psychology, Yale University

Ronald D. Henderson, National Education Association

Vinetta Jones, Executive Director, EQUITY 2000, College Board

Arturo Madrid, Director, Tomas Rivera Center

Hernan LaFontaine, Administration and Supervision,
Southern Connecticut State University

William Julius Wilson, Department of Sociology, University of Chicago

Center Liaison

Harold Himmelfarb, Office of Educational Research and Improvement

Effects of Instructional Grouping on Seventh Graders' Academic Motivation and Achievement

David A. Reuman
Trinity College

Douglas J. Mac Iver
Johns Hopkins University

Report No. 50

November 1994

Published by the Center for Research on Effective Schooling for Disadvantaged Students, supported as a national research and development center by funds from the Office of Educational Research and Improvement (OERI), United States Department of Education (R117R90002). The opinions expressed in this publication do not reflect the position or policy of OERI, and no official endorsement should be inferred.

Center for Research on Effective Schooling for Disadvantaged Students
The Johns Hopkins University
3505 North Charles Street
Baltimore, Maryland 21218

The Center

The mission of the Center for Research on Effective Schooling for Disadvantaged Students (CDS) is to significantly improve the education of disadvantaged students at each level of schooling through new knowledge and practices produced by thorough scientific study and evaluation. The Center conducts its research in four program areas: The Early and Elementary Education Program, The Middle Grades and High Schools Program, the Language Minority Program, and the School, Family, and Community Connections Program.

The Early and Elementary Education Program

This program is working to develop, evaluate, and disseminate instructional programs capable of bringing disadvantaged students to high levels of achievement, particularly in the fundamental areas of reading, writing, and mathematics. The goal is to expand the range of effective alternatives which schools may use under Chapter 1 and other compensatory education funding and to study issues of direct relevance to federal, state, and local policy on education of disadvantaged students.

The Middle Grades and High Schools Program

This program is conducting research syntheses, survey analyses, and field studies in middle and high schools. The three types of projects move from basic research to useful practice. Syntheses compile and analyze existing knowledge about effective education of disadvantaged students. Survey analyses identify and describe current programs, practices, and trends in middle and high schools, and allow studies of their effects. Field studies are conducted in collaboration with school staffs to develop and evaluate effective programs and practices.

The Language Minority Program

This program represents a collaborative effort. The University of California at Santa Barbara and the University of Texas at El Paso are focusing on the education of Mexican-American students in California and Texas; studies of dropout among children of recent immigrants have been conducted in San Diego and Miami by Johns Hopkins, and evaluations of learning strategies in schools serving Navajo Indians have been conducted by the University of Northern Arizona. The goal of the program is to identify, develop, and evaluate effective programs for disadvantaged Hispanic, American Indian, Southeast Asian, and other language minority children.

The School, Family, and Community Connections Program

This program is focusing on the key connections between schools and families and between schools and communities to build better educational programs for disadvantaged children and youth. Initial work is seeking to provide a research base concerning the most effective ways for schools to interact with and assist parents of disadvantaged students and interact with the community to produce effective community involvement.

Abstract

This study tested the hypothesis that cooperative learning practices in heterogeneous classes would enhance students' academic motivation and achievement. Half of the seventh graders entering a junior high school were randomly assigned to heterogeneous classrooms where their teachers implemented Student Team Learning (STL) techniques (Slavin, 1986). The remaining students were assigned to homogeneous, ability-grouped classrooms. The use of STL in heterogeneous classes produced achievement benefits, although the effects were stronger for course grades (median Δ = .62) and for grades on end-of-year departmental exams and research projects (median Δ = .48) than for a nationally-standardized achievement test (median Δ = .12) or for a state-standardized, criterion-referenced test of learning objectives for the seventh grade (median Δ = .06). The effects of STL in heterogeneous classes on achievement varied by subject area, with consistent positive effects in language arts and science, and inconsistent effects in math and social studies. The effects of STL in heterogeneous classes on motivational variables also varied by subject area. In language arts, math, and science, STL used in heterogeneous classes raised students' self-concept of academic ability (median Δ = .26), self-reported academic effort (median Δ = .20), and academic task value (median Δ = .13), and lowered students' evaluation anxiety (median Δ = -.27). However, in social studies the motivational effects were all reversed. Variation of effects across subject areas can be attributed to variation in teachers' implementation of STL. By demonstrating that the achievement benefits of STL in heterogeneous classrooms accrue in conjunction with internal changes in students' academic expectancies, values, and motives, we may expect the benefits of these instructional grouping practices to generalize beyond (and to outlast) the classrooms in which they were first generated.

Acknowledgments

This research would not have been possible without the leadership of M. Leon McKinley, principal, Bloomfield (Connecticut) Junior High School; Dr. Paul Copes, superintendent, Bloomfield Public Schools; and the Bloomfield Board of Education. Special credit is due the faculty of Bloomfield Junior High School for their dedication to this project and to the students of Bloomfield Junior High School for their cooperation.

This research was supported by the Center for Research on Effective Schooling for Disadvantaged Students at Johns Hopkins University, with funds from the Office of Educational Research and Improvement, United States Department of Education (Grant R117R90002). The first author was supported by a faculty research leave from Trinity College during the fall of 1990, when the field work was initiated.

The authors thank John Hollifield and Robert Slavin for thoughtful comments on this manuscript.

Introduction

Research on between-classroom ability grouping has shown that the practice does not enhance student achievement at the elementary school level (Slavin, 1987) or at the secondary level (Slavin, 1990a). Research on cooperative learning, by comparison, has demonstrated that this instructional grouping practice can enhance student achievement (Slavin, 1989, 1990b). In response to recent reports questioning the widespread use of between-classroom ability grouping (e.g., Braddock & Mc Partland, 1990; Oakes, 1985, 1986a, 1986b), many schools have begun to restructure their instructional grouping practices (Oakes & Lipton, 1992; Slavin, Braddock, Hall, & Petza, 1989; Wheelock, 1992). Developmentally-appropriate instructional grouping practices have been a special concern for systemic reform in the middle grades (Carnegie Task Force on the Education of Young Adolescents, 1989; Connecticut Task Force on the Education of Early Adolescents, 1991; Eccles & Midgley, 1989; Epstein & Mac Iver, 1990; Maryland Task Force on the Middle Learning Years, 1989).

The present research represents an experimental evaluation of one junior high school's effort to restructure an ability-grouped curriculum. Half of the seventh graders entering the school were assigned to homogeneous, ability-grouped classrooms, as had been the school's practice for many years. The other entering students were assigned to heterogeneous classrooms in which Student Team Learning (STL) techniques (Slavin, 1986, 1990b) were implemented. STL involves the formation of small, mixed-ability teams of students who study together, rather than individually, after the teacher presents a new lesson. STL emphasizes improvement over one's past performance, so all team members have an equal opportunity to contribute to the success of the team. STL requires testing students individually after a period of team study, so individual accountability is maintained. Finally, STL provides team rewards and recognition, so the incentive to work together is built.

This experimental evaluation of a junior high school's restructuring effort differs in an important respect from previous experimental tests of between-classroom grouping and cooperative learning practices. Here, effects of between-classroom grouping and STL are evaluated simultaneously. Typically the effects of between-classroom grouping practices are evaluated by comparing student achievement in heterogeneously- and homogeneously-grouped classrooms, where whole-class instruction and individual (student) seatwork are common to both instructional grouping conditions. Typically the effects of cooperative learning practices are evaluated by comparing student achievement in classrooms that do or do not expose students to cooperative learning practices, where heterogeneous class assignments are a common denominator of both cooperative learning conditions. This junior high school was interested in directly evaluating its traditional instructional grouping practice (between-classroom grouping with whole-class instruction and individual student seatwork) against a comprehensive alternative (heterogeneous class assignments with whole-class instruction and STL). Because the junior high school was already organized into two seventh-grade "houses" (i.e., two teams of academic teachers that each taught approximately 90 seventh graders), it was feasible to arrange an experimental comparison of the two instructional grouping programs of interest by locating each grouping program in its own "house." Because of the small size of the junior high school, between-classroom grouping and STL could not be manipulated as completely crossed factors in the evaluation design.[1]

Although substantial research efforts have focussed on effects of instructional grouping practices (including cooperative learning practices) on student achievement, studies of effects of these practices on motivational variables have usually been limited to global measures of self-esteem or academic self-concept (Kulik & Kulik, 1982;

Slavin, 1990b). The present investigation addresses a range of motivational variables, including students' achievement expectancies, values, and motives (Atkinson & Feather, 1966; Atkinson & Raynor, 1974; Eccles (Parsons), 1983). Each motivational variable was assessed separately for each academic subject area to allow for individual differences among students in motivation for achievement in different academic subject areas and to allow for possible subject-area differences in motivational outcomes due to differences in teachers' implementation of STL. We hypothesize that the inter-dependent task structure of STL models will enhance the value of academic tasks for early adolescent students. We hypothesize that the reward structure of STL models (e.g., using improvement criteria rather than norm-referenced criteria for grading individual students; using team recognition and rewards based on the average improvement points earned by learning teams) will raise students' expectancies for success and lower students' evaluation anxieties with respect to their academic work.

It is important to document the effects of instructional grouping practices on internal motivational variables. If the achievement benefits of cooperative learning practices accrue exclusively because of situational (peer) influences, then we would expect the benefits to be as situation-specific (and as transient) as the implementation of the practice. If the achievement benefits of cooperative learning practices accrue in conjunction with internal changes in students' achievement expectancies, values, and motives, then we would expect the benefits to generalize beyond (and to outlast) the environment in which they were initiated. Benefits of STL for students' academic motivation in the middle grades, for instance, could be expected to enhance students' persistence in optional academic subjects later in secondary school (Eccles, 1984).

Method

Manipulation of instructional grouping conditions

Throughout the 1990-91 school year, all seventh-grade students and teachers of core academic subjects (i.e., language arts, math, science, and geography) at a junior high school participated in an experiment designed to compare two forms of instructional grouping. Half of the students were randomly assigned to homogeneous, ability-grouped classrooms; these classes were called "Blue Team" classes. The other half of the students were randomly assigned to heterogeneous classrooms[2] in which Student Team Learning techniques were implemented; these classes were called "Green Team" classes. In language arts, science, and geography, the odds of Blue Team students being assigned to above-grade, on-grade, or below-grade level classrooms were 2:5:1. In math, the odds of both Blue and Green Team students being assigned to above-grade or on-grade level classrooms were 2:3. The instructional grouping on the Blue Team was effectively what Slavin (1987) has called "ability-grouped class assignment" (rather than "regrouping for specific subjects") insofar as 95 percent of all Blue Team students took at least three of four core academic subjects at the same level. Regardless of the organization of their classrooms, all seventh-grade students were exposed to the same core curriculum during the school year.

Teacher preparation

Teachers were not randomly assigned to Teams;[3] Green Team teachers volunteered to teach in the new instructional grouping program. On-site staff development in the use of Student Team Learning was provided by specialists from Johns Hopkins University in August, October, November, February, and May of the 1990-91 school year (a total of nine in-service days). In addition, teachers using Student Team Learning attended a 3-day intensive workshop in Baltimore either in October, 1990, or March, 1991. Throughout the school year, teachers on both Teams had regular common planning periods. On the Green Team, teachers dedicated at least one common planning period each week to issues connected with the new instructional grouping program.

Sample

The junior high school was located in a suburban school district in Connecticut. It was the only public school serving seventh graders in the district. Students in the seventh grade were 70 percent African American, 25 percent Caucasian, and 5 percent Hispanic; they were 55 percent female. Parents of these seventh graders tended to have achieved high levels of education: 44 and 47 percent of these students' mothers and fathers, respectively, had earned a degree from a 4-year college. Students who received special education services were excluded from all analyses because they were not randomly assigned to a Team; in fact, all students with identified, mild learning disabilities were assigned to the Green Team. The final analysis sample included between 134 and 154 seventh graders, depending on the dependent variable under consideration.[4]

Measures

To establish comparability of the two groups of students at the outset of the school year, scores from sixth-grade administrations of a nationally standardized, norm-referenced achievement test (the California Achievement Test; CTB/McGraw-Hill, 1986) and of a state-standardized, criterion-referenced test (the Connecticut Mastery Test; State of Connecticut Department of Education, 1990a, 1990b) were collected, as well as information from sixth-grade report cards on achievement in core academic subjects. To assess effects on achievement of the two forms of instructional grouping in the seventh grade, we collected scores from seventh-grade or early eighth-grade administrations of the same standardized tests, as well as information from seventh-grade report cards. To assess motivational variables, we used questionnaires filled out by students at school in November, February, and May of the seventh grade school year. The questionnaires included multiple indicators of students' self-concepts of ability in academic subjects (Reuman, 1989), perceptions of the intrinsic and utility value of academic subjects (Eccles, 1984; Mac Iver & Reuman, 1988), and measures of evaluation anxiety in the context of each academic subject (Feld & Lewis, 1969; Reuman, 1991; Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960). Internal consistency reliabilities ranged from .62 to .91 (see Table 1). A single indicator was used to assess students' academic effort. All indicators are documented in Appendix A.

Table 1 about here

Results

Pre-treatment comparability of groups

Several tests were performed to demonstrate the initial comparability of Blue and Green Team students. As would be expected from the random assignment procedure, no significant differences between Teams were observed with respect to the distribution of student gender, race/ethnicity, mother's educational attainment, father's educational attainment, grade level at which the student entered the public school system, any battery of the California Achievement Test administered in the sixth grade, or any learning objective assessed by the sixth-grade Connecticut Mastery Test. Contrary to expectation, Green Team students had earned significantly higher achievement marks in some academic areas in the first marking period of the sixth grade; however, no achievement differences between Blue and Green Team students were significant in the second or third marking periods of the sixth grade. Overall, we conclude that the random assignment procedure produced initially equivalent groups.

By design, Blue Team students were homogeneously grouped for all core academic subjects, whereas Green Team students were only homogeneously grouped for mathematics instruction. Manipulation checks indicated that Green Team classrooms were more heterogeneous, as planned. Classroom heterogeneity was assessed by computing th within-classroom variances of batteries of the California Achievement Test, administered when students had been sixth graders. As can be seen in Table 2, Green Team classes were more heterogeneous in language arts, $F\ (1, 6) = 6.62, p = .04$; in science, $F\ (1, 8) = 9.97, p = .01$; and in geography, $F\ (1, 7) = 5.29, p = .05$. As expected, there was not a significant difference between teams with respect to heterogeneity in mathematics, $F\ (1, 8) = .01, p = .91$.

Table 2 about here

Effects of instructional grouping on achievement

Course grades. Table 3 displays grades earned in core academic subjects at all marking periods during the 1990-91 school year. Effect size statistics show consistent and substantial achievement benefits in language arts, mathematics, and science of membership on the Green Team, where students were assigned to heterogeneous classes and exposed to Student Team Learning. No benefits for Green Team students were evident in geography (the core social studies course). Repeated-measures MANOVAs confirm these effects of team membership on grades. Pooling across all four seventh-grade marking periods, Green Team students earned significantly higher grades in language arts, $F\ (1, 149) = 21.01, p < .0001$; in mathematics, $F\ (1, 148) = 13.27, p = .0004$; and in science, $F\ (1,147) = 30.36, p < .0001$. There was not a significant difference between teams with respect to grades earned in geography, $F\ (1, 148) = .09, p = .76$.

Table 3 about here

The finding that Green Team students earned higher grades may mean that they have learned more, but it may also mean that Green Team teachers were using different criteria for assigning grades, such as improvement criteria. Comparisons with grades earned in Grade 6 revealed that Green Team students maintained the grades they had been earning in Grade 6, whereas Blue Team students showed a decline in grades earned after they entered the junior high school. The finding that Green Team students tended to earn comparable grades in the sixth and seventh grades argues against the inference that Green Team teachers are simply inflating the grades of their students, unless one is willing to argue that all Grade 6 teachers inflated students' grades.

Repeated-measures MANOVAs also indicated substantial effects of marking period in all subject areas and team-by-marking period interactions in mathematics and geography. In language arts, grades declined linearly over the school year, $F\ (3. 147) = 21.88, p < .0001$. Because the team-by-period interaction was not significant, $F\ (3, 147) = 0.61, p = .61$, we may infer that the decline in grades during the school year (about two letter-grade units) was the same for Blue and Green Team students. In math, grades declined abruptly after the first marking period, and then stabilized, $F\ (3, 146) = 26.52, p < .0001$. The drop in math grades was larger for Blue Team students (about 1.7 letter-grade units) than for Green Team students (about 1.1 letter-grade units), yielding a significant team-by-period interaction, $F\ (3, 146) = 5.11, p = .0022$. In science, grades increased slightly after the first marking period, then dropped and stabilized at the third marking period. $F\ (3, 145) = 18.13, p < .0001$. The team-by-period interaction was not significant in the case of science grades, F (3, 145) = 2.24, p = .09. In geography, significant effects were observed for marking period, F (3, 146) = 6.92, p = .0002, and for the team-by-period interaction, F (3, 146) = 7.25, p < .0001. Geography grades were level throughout the school year for Green Team students, whereas grades rose at the third marking period and then declined substantially at the fourth marking period for Blue Team students.

Several factors may help account for the changes in grades seen over the four marking periods of seventh grade. In many junior high schools, particularly those that matriculate students from several different elementary schools, the seventh grade begins with a great deal of review. It is possible that students would earn higher grades on material that is being reviewed, and that grades would drop as new material is introduced. Because this (7-8) junior high school matriculated seventh graders from the same (5-6) middle school, it may be that there was less review than would be common in other junior high schools, but this is still one likely factor underlying the patterns of change observed in students' grades. Another factor underlying declines in grades during the seventh grade may be progressive difficulty of the curriculum. Mathematics, in particular, tends to be a subject where the curriculum builds on itself and later learning is dependent on concepts learned earlier in the school year. As the difficulty of the curriculum increases, grades may decline, unless students compensate with more studying time. Finally, there may be changes in students' academic motivation during the school year. For some early adolescents, social relations may become increasingly important and time-consuming, at the expense of school work.

Special grades. Year-end departmental exams in core academic subjects and research projects in language arts and science were evaluated with standardized criteria for students on both teams. Table 4 displays grades earned on these achievement measures and effect size statistics for team comparisons. Because the year-end departmental exams and research projects were evaluated with standardized criteria, these performance measures were not vulnerable, to the same extent that course grades were, to the criticism that they may be biased by factors such as the expectations held by the person making the evaluation. With respect to year-end research projects, Green Team students earned higher grades than Blue Team students in both disciplines where projects were assessed [in language arts, F (1, 145) = 14.73, p = .0002; and in science, F (1, 143) = 18.24, p < .0001]. Green Team students outperformed Blue Team students on the year-end departmental exam in language arts, F (1, 147) = 12.62, p = .0005. Blue Team students outperformed Green Team students on the departmental exam in geography, F (1, 149) = 8.99, p = .003. There were no differences in performance on the departmental exams in mathematics, F (1, 151) = 0.42, p = .52, or science, F (1, 144) = 2.29, p = .13. Overall, the predicted achievement benefits of STL in heterogeneous classes are confirmed in language arts and science, contraindicated in geography, and equivocal in mathematics.

Table 4 about here

California Achievement Test. Analyses of batteries of the California Achievement Test (CTB/McGraw-Hill, 1986) administered in April (1991) of the seventh grade generally showed no Team differences. Table 5 displays mean battery scores for each team and effect size statistics for team comparisons. Green Team students outperformed Blue Team students on the Total Language Arts Battery of the seventh-grade California Achievement Test, F (1, 151) = 3.84, p = .05, but no significant Team differences were observed for the Total Reading, F (1, 151) = 1.68, p = .20; Total Math, F (1, 152) = 0.00, p = .95; Total Battery, F (1, 151) = 1.65, p = .20; Science, F (1, 150) = 0.15, p = .70; or Social Studies Batteries, F (1, 150) = 0.64, p = .42.

Table 5 about here

Connecticut Mastery Test. Three measures from a released form of the Grade 8 Connecticut Mastery Test (CMT) were administered in the last week of the seventh-grade school year (1990-91).[5] The top portion of Table 6 displays the mean performance for each team at the end of Grade 7 and effect-size statistics for team comparisons. MANOVAs indicated that Blue and Green Team students showed comparable mastery of learning objectives on the Mathematics - Conceptual Understandings battery, F (1, 132) = 1.78, p = .18; and on the Language Arts - Reading Comprehension battery, F (1, 134) = 0.58, p = .45. We observed a marginally significant Green Team advantage on the Degrees of Reading Power battery, F (1, 152) = 3.31, p = .07. The bottom portion of Table 6 displays the mean CMT performance for each team at the beginning of Grade 8 and effect-size statistics for team comparisons. None of the team comparisons was statistically significant (F - statistics ranged from 0.00 to 1.10). The magnitude of effect sizes for the three repeated CMT measures (i.e., Math: Conceptual Understandings; Language Arts: Reading Comprehension; and Degrees of Reading Power) were all attenuated during the interval between the Grade 7 (June) and Grade 8 (October) administrations, although mean performance on each of the measures was higher at the Grade 8 administration. The attenuation of effect sizes in Grade 8 may be due to the fact that all Grade 8 classrooms (except mathematics) were heterogeneously grouped in 1991-92, and all Grade 8 teachers had been given staff development in STL and were encouraged to implement STL. Even though students had only been in Grade 8 for several weeks when they took the CMT, the fact that they had all been in "treatment" classrooms would have reduced program effects from the previous school year.

Table 6 about here

Effects of instructional grouping on motivational variables

Self-concept of academic ability. Table 7 displays students' self-concepts of ability in academic subject areas at three points during the seventh-grade school year. Effect-size statistics vary substantially across subject areas. Pooling across all three waves of observations, repeated-measures MANOVAs show marginally significant benefits of membership on the Green Team for self-concept of ability in language arts, F (1, 138) = 3.06, p = .08; and in science, F (1, 135) = 2.65, p = .10; no team effects in math, F (1, 134) = 0.00, p = .96; and a disadvantage of membership on the Green Team for self-concept of ability in geography, F (1, 135) = 18.43, p < .0001.

Table 7 about here

Repeated-measures MANOVAs also indicate effects of wave in all subject areas and team-by-wave interactions in science and geography. In language arts, self-concept tended to decline over the school year, F (2, 137) = 2.53, p = .08, with no team-by-wave interaction, F (2, 137) = 1.49, p = .23. In math, self-concept declined after November and then remained level, F (2, 133) = 7.93, p = .0006, with no team-by-wave interaction, F (2, 133) = 0.22, p = .80. In science, self-concept declined over the school year for Green Team students but remained relatively level for Blue Team students, yielding both a wave effect, F (2, 134) = 5.86, p = .004, and a team-by-wave interaction effect, F (2, 134) = 4.33, p = .02. In geography, self-concept declined and then returned to its original level for Green Team students but increased steadily for Blue Team students, yielding both a wave effect, F (2, 134) = 3.58, p = .03, and a team-by-wave interaction effect, F (2, 134) = 3.53, p = .03.

The large effects of instructional grouping conditions on course grades did not translate into comparable effects on self-concept of academic ability. The median effect size of grouping conditions on grades in language arts was .71, compared to a median effect size of grouping conditions on self-concept in language arts of .27. In math, the median effect size of grouping conditions on grades was .51 but the median effect size for self-concept was -.02. In science, the median effect size of instructional grouping on grades was .89 but the median effect size for self-concept was only .26. Finally, in geography, the median effect size on grades was -.05 whereas the median effect size on self-concept was -.83. This pattern suggests that students were not systematically comparing their own grades with those of students on the other team (i.e., house). If social comparison of course grades was used in academic self-evaluations, it was most likely based on within-team comparisons.

The observed levels of academic self-concept correspond more closely to grade differences among subject areas and, to a lesser extent, to changes in course grades over the school year. For instance, relative to their grades in language arts and science, Green Team students were earning substantially lower grades in geography, and these students' self-concepts of ability show a similar profile. Among Blue Team students, grades were typically lowest in science and highest in geography, and these students' self-concepts show a corresponding pattern. The decline in language arts grades during the school year corresponds to the wave-related decline in students' self-concept of ability in language arts. The drop in math grades after the first marking period, followed by level math grades, is strikingly similar to the wave-related pattern of change in self-concept of math ability. Overall, these similarities suggest that students' self-evaluations of academic ability may be more strongly linked to grade comparisons they make across subject areas and over time.

Self-reported effort. Table 8 displays students' self-reported effort in academic subject areas at three waves during the seventh-grade school year. Effect-size statistics indicate that effects of instructional grouping on effort vary according to subject area. Pooling across all three waves, repeated-measures MANOVAs show a marginally significant beneficial effect of Green Team membership on effort in language arts, F (1, 138) = 3.33, p = .07; no overall team effects in math, F (1, 132) = 2.51, p = .12, or in science, F (1, 135) = 1.70, p = .19; and a significant detrimental effect of Green Team membership on effort in geography, F (1, 135) = 8.22, p = .005.

Table 8 about here

Repeated-measures MANOVAs also indicated effects of wave and team-by-wave interactions. In language arts, effort increased slightly in February for Blue Team students and then declined substantially, whereas effort was level throughout the year for Green Team students, contributing to a significant wave effect, F (2, 137) = 3.93, p = .02, and team-by-wave interaction, F (2, 137) = 5.37, p = .006. In math, effort declined steadily for students on both teams, leading to a significant wave effect, F (2, 131) = 9.02, p = .0002, but no team-by-wave interaction, F (2, 131) = 0.20, p = .82. In science, Blue Team students reported low effort throughout the year, whereas Green Team students' effort declined during the year until they were at a comparably low level, resulting in a wave effect, F (2, 134) = 5.13, p = .007, and a team-by-wave interaction effect, F (2, 134) = 3.81, p = .02. In geography, there was no wave effect, F (2, 134) = 1.37, p = .26, or team-by-wave interaction, F (2, 134) = 0.90, p = .41. To summarize, self-reported effort was generally level during the school year, except for declines on both teams in math effort, a decline on the Blue Team in language arts effort, and a decline on the Green Team in science effort.

Academic task value. As was the case with other motivational variables, effects of instructional grouping on students' academic task values varied for different subject areas. Table 9 displays students' task values near the beginning and end of seventh grade and effect-size statistics for team comparisons. Green Team students valued science in November more highly than did Blue Team students, but the team difference was not significant by the end of the school year, producing a team-by-wave interaction, $F(1, 137) = 23.23$, $p < .0001$, and a wave effect, $F(1, 137) = 23.23$, $p < .0001$, but no overall effect of team $F(1, 137) = 0.63$, $p = .43$. Contrary to predictions, Blue Team students valued geography more highly than Green Team students, $F(1, 138) = 9.17$, $p = .003$, with no effect of wave, $F(1, 138) = 1.36$, $p = .25$, and no team-by-wave interaction, $F(1, 138) = 2.03$, $p = .16$. No team, wave, or team-by-wave interaction effects were observed in students' valuing of language arts or mathematics. Comparisons across subject areas suggest that Blue Team students placed distinctively low value on their work in science, whereas Green Team students placed distinctively low value on their work in geography.

Table 9 about here

Evaluation anxiety. Analyses of evaluation anxiety (assessed in February) showed effects of instructional grouping conditions that varied according to the academic subject under consideration. Table 10 displays descriptive statistics on evaluation anxiety for each team and effect-size statistics for team comparisons. As predicted, Green Team students reported lower evaluation anxiety in science, $F(1, 149) = 6.21$, $p = .01$. Contrary to predictions, Blue Team students reported lower anxiety in geography, $F(1, 149) = 4.57$, $p = .03$. No team differences were observed in evaluation anxieties in language arts, $F(1, 145) = 0.20$, $p = .65$, or mathematics, $F(1, 151) = 2.94$, $p = .09$.

Table 10 about here

Discussion

Overall, assignment to heterogeneous classes and exposure to Student Team Learning resulted in higher student achievement. These achievement benefits were strongest for course grades (median $\Delta = .62$), moderately strong for grades assigned on end-of-year departmental exams and research projects (median $\Delta = .48$), and weak but positive for the Grade 7 California Achievement Test (median $\Delta = .12$), "Grade 7" Connecticut Mastery Test (median $\Delta = .13$), and Grade 8 Connecticut Mastery Test (median $\Delta = .06$). This pattern is similar to Slavin's (1989) summary finding that, among cooperative learning methods that use group goals and individual accountability, the effect size on achievement was stronger for measures overall (median $\Delta = .30$) than for standardized measures (median $\Delta = .21$).

End-of-year departmental exams and research projects represent the most meaningful achievement outcome measures in this study. The departmental exams used uniform evaluation criteria for students in all classrooms, rather than the idiosyncratic criteria underlying teachers' course grades. The departmental exams and research projects were "authentic," and directly related to what was taught on both teams, whereas the California Achievement Test and even the Connecticut Mastery Test were more distantly related to the curriculum of the school.

The intervention in this junior high school was ambitious. Teachers were asked to learn how to use Student Team Learning and to adapt their curriculum to it continuously during the school year. Almost all of the teachers reported that they did not have enough planning time to carry out this curriculum adaptation to their own satisfaction. The geography teacher on the Green Team was teaching the seventh grade curriculum for the first time, having just been reassigned from the high school in the district, and she reported particular problems with insufficient planning time. This individual circumstance of the Green Team geography teacher may help explain the markedly different student outcomes in that subject area. Some of the teachers implemented certain components of Student Team Learning models frequently (such as team study) but neglected implementing other components (such as team recognition based on the average improvement points earned by members of each study team), even after the omissions were brought to their attention. Program evaluation was carried out in the first year of implementation, and it is possible that more robust effects on achievement must await more development of teachers' in-class practices.

The effects of STL in heterogeneous classes on motivational variables varied by subject area. In language arts, math, and science, STL used in heterogeneous classes raised students' self-concept of academic ability (median $\Delta = .26$), self-reported academic effort (median $\Delta = .20$), and academic task value (median $\Delta = .13$), and lowered students' evaluation anxiety (median $\Delta = -.27$), as predicted. However, in social studies the motivational effects were all reversed. As noted above, this was the subject area taught by a teacher who had just been reassigned from the high school. Frequent, informal classroom observations revealed substantial variation in teachers' implementation of STL, despite the fact that all Green Team teachers had received comparable staff development. Implementation was most frequent and most thorough in language arts and science. Variation of STL effects on motivational variables across subject areas is most likely attributable to this variation in teacher's implementation of STL. Individual differences among teachers were completely confounded with subject area.

Ultimately, by demonstrating that achievement benefits of STL in heterogeneous classrooms generally accrued in conjunction with internal changes in students' academic expectancies, values, and motives, we may expect that the benefits of these instructional grouping practices will generalize beyond the classrooms in which they were first generated.

Footnotes

[1] A completely-crossed design would have required four "houses," where one house would incorporate homogeneous grouping with STL, another would incorporate homogeneous grouping without STL, a third would incorporate heterogeneous grouping with STL, and the last house would incorporate heterogeneous grouping without STL. Even if the size of the junior high school had not precluded this design, securing permission from the Board of Education for randomly assigning students to *four* conditions very likely would have exceeded the investigators' powers of persuasion.

[2] In mathematics, students were assigned either to "Math 7" (an on-grade level course) or "Pre-Algebra" (an above-grade level course). The two levels of math instruction were used on both the Blue and the Green Teams. In this respect, classroom heterogeneity did not differ across Teams in math. However, the number of levels used for math instruction was reduced from three levels used in previous years.

[3] Insofar as there was only one teacher in each subject area on each team (except for language arts), random assignment of teachers to the instructional grouping conditions would not have produced teacher equivalence within subject areas any more than non-random assignment would have. In language arts, one teacher taught entirely with the Blue Team, one entirely with the Green Team, and one taught half her sections with the Blue Team and the other half with the Green Team.

[4] Due to the small size of the junior high school, the design had (unavoidably) low statistical power (Kraemer & Thiemann, 1987). Consequently, both effect-size coefficients and statistical tests are reported.

[5] Normally, the Grade 8 Connecticut Mastery Test is administered at the end of the first month of Grade 8. Because it is designed to assess learning objectives that are taught in Grade 7 mathematics and language arts curriculums, the Grade 8 Connecticut Mastery Test should have comparable content validity whether administered near the end of seventh grade or the beginning of eighth grade. A released form of the Grade 8 Connecticut Mastery Test was used in this special administration at the end of seventh grade so that students would not re-take the same test as part of Connecticut's program of statewide student assessment at the beginning of eighth grade. The released form of the test had not been used for statewide student assessment since 1988; otherwise, it was a parallel form of the non-released Grade 8 Connecticut Mastery Test.

References

Atkinson, J. W., & Feather, N. T. (1966). *A theory of achievement motivation.* New York: Wiley.

Atkinson, J. W., & Raynor, J. O. (1974). *Motivation and achievement.* Washington, D.C.: Halsted Press.

Braddock, J. H., II, & Mc Partland, J. M. (1990). Alternatives to tracking. *Educational Leadership, 47,* 76-79.

Carnegie Task Force on the Education of Young Adolescents. (1989). *Turning points: Preparing American youth for the 21st century.* New York: Carnegie Council on Adolescent Development, Carnegie Corporation of New York.

Connecticut Task Force on the Education of Early Adolescents. (1991). *Right in the middle.* Middletown, CT: Connecticut State Department of Education.

CTB/McGraw-Hill. (1986). *California Achievement Tests: Forms E and F.* Monterey, CA: Author.

Eccles (Parsons), J. (1983). Expectancies, values, and academic behaviors. In Janet T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-146). San Francisco, CA: Freeman.

Eccles, J. (1984). Sex differences in achievement patterns. In T. B. Sonderegger (Ed.), *Nebraska Symposium on Motivation: Psychology and Gender* (pp. 97-132). Lincoln, NE: University of Nebraska Press.

Eccles, J.S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames and R. Ames (Eds.), *Research on motivation in education (Volume 3): Goals and cognitions* (pp. 139-186). San Diego, CA: Academic Press.

Epstein, J.L., & Mac Iver, D. J. (1990). *Education in the middle grades: National practices and trends.* Columbus, OH: National Middle School Association.

Feld, S.C., & Lewis, J. (1969). The assessment of achievement anxieties in children. In C. P. Smith (Ed.), *Achievement-related motives in children.* (pp. 151-199). New York: Russell Sage Foundation.

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

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research.* Newbury Park, CA: Sage Publications.

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

Mac Iver, D., & Reuman, D. A. (1988). Decision-making in the classroom and early adolescents' valuing of mathematics. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Maryland Task Force on the Middle Learning Years. (1989). *What matters in the middle grades?* Baltimore, MD: Maryland State Department of Education.

Oakes, J. (1985). *Keeping track: How schools structure inequality.* New Haven: Yale University Press.

Oakes, J. (1986a). Keeping track, Part 1: The policy and practice of curriculum inequality. *Phi Delta Kappan, 68*, 12-17.

Oakes, J. (1986b). Keeping track, Part 2: Curriculum inequality and school reform. *Phi Delta Kappan, 68,* 148-153.

Oakes, J., & Lipton, M. (1992). Detracking schools: Early lessons from the field. *Phi Delta Kappan, 73,* 448-454.

Reuman, D. A. (1989). How social comparison mediates the relation between ability grouping practices and students' achievement expectancies in mathematics. *Journal of Educational Psychology, 81,* 178-189.

Reuman, D. A. (1991). Change in ability grouping and evaluation anxiety in mathematics at the transition to junior high school. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Sarason, S.B., Davidson, K.S., Lighthall, F.F., Waite, R.R., & Ruebush, B.K. (1960). *Anxiety in elementary school children: A report of research.* New York: Wiley.

Slavin, R. E. (1986). *Using Student Team Learning (3rd ed.).* Baltimore: Johns Hopkins University, Center for Research on Elementary and Middle Schools.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research, 57,* 293-336.

Slavin, R. E. (1989). Cooperative learning and student achievement. In Robert E. Slavin (Ed.), *School and classroom organization* (pp. 129-156). Hillsdale, NJ: Lawrence Erlbaum.

Slavin, R. E. (1990a). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research, 60,* 471-499.

Slavin, R. E. (1990b). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice Hall.

Slavin, R. E., Braddock, J. H., II, Hall, C., & Petza, R. J. (1989). Alternatives to ability grouping. Baltimore: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

State of Connecticut Department of Education. (1990a). *Grade 6 Mastery Test Results: Summary and interpretations, 1989-90.* Hartford, CT: Author.

State of Connecticut Department of Education. (1990b). *Grade 8 Mastery Test Results: Summary and interpretations, 1989-90.* Hartford, CT: Author.

Wheelock, A. (1992). *Crossing the tracks: How "Untracking" can save America's schools.* New York: The New Press.

Table 1

Internal Consistency Reliabilities (Cronbach's alpha) for Measures of Motivational Constructs

	Subject Area			
Motivational Construct	Language Arts	Math	Science	Geography
Self-Concept of Ability				
November	.78	.62	.82	.85
February	.80	.63	.83	.84
May	.81	.74	.84	.87
Academic Task Value				
November	.84	.71	.85	.86
May	.89	.70	.87	.88
Evaluation Anxiety				
February	.88	.85	.91	.91

Note. Self-concept of ability, academic task value, and evaluation anxiety were composite measures, computed by summing four, four, and three indicators, respectively.

Table 2

Effects of Instructional Grouping Conditions on Within-Classroom Heterogeneity

	Team						
	Blue			Green			
Subject Area	n	Mean	SD	n	Mean	SD	Δ
Language Arts	4	328	313	4	1117	527	2.52
Mathematics	5	520	749	5	477	421	-.06
Science	5	649	327	5	1935	851	3.93
Geography	5	401	192	4	754	269	1.84
Overall	19	482	434	18	1086	788	1.39

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes (except in mathematics) with STL. Heterogeneity was assessed by computing within-classroom variances of batteries of the California Achievement Test (i.e., classrooms are the unit of analysis). Positive effect-size coefficients (Δ) indicate greater heterogeneity in Green Team classes.

Table 3

Effects of Instructional Grouping Conditions on Course Grades During the 1990-91 School Year

	Team				
	Blue		Green		
Achievement Measure	Mean	SD	Mean	SD	Δ
Language Arts	n = 76		n = 75		
Period 1	6.58	2.96	8.55	3.35	.66
Period 2	5.70	2.84	7.56	3.47	.65
Period 3	5.28	2.64	7.32	3.30	.77
Period 4	4.54	2.66	6.89	3.43	.88
Mathematics	n = 75		n = 75		
Period 1	6.31	3.06	7.59	2.57	.42
Period 2	4.59	3.13	6.47	3.25	.60
Period 3	5.16	3.21	6.19	3.15	.32
Period 4	4.33	3.05	6.51	2.89	.71
Science	n = 74		n = 75		
Period 1	5.27	2.86	8.05	3.35	.97
Period 2	5.90	3.52	8.75	2.61	.81
Period 3	4.92	3.57	6.91	3.42	.56
Period 4	4.40	2.78	7.23	3.78	1.02
Geography	n = 76		n = 74		
Period 1	6.45	1.92	6.39	2.43	-.03
Period 2	6.37	2.04	6.20	2.50	-.08
Period 3	6.91	1.84	6.20	2.62	-.38
Period 4	5.70	2.61	6.22	2.36	.20

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate an achievement advantage for students in Green Team classes. Course grades are coded <13>A+, <12>A, <11>A-, <10>B+, <9>B, <8>B-, <7>C+, <6>C, <5>C-, <4>D+, <3>D, <2>D-, and <1>F.

Table 4

Effects of Instructional Grouping Conditions on Grades For Year-End Departmental Exams and Research Projects

	Team						
	Blue			Green			
Achievement Measure	n	Mean	SD	n	Mean	SD	Δ
Language Arts							
Exam	72	3.19	2.61	77	5.05	3.64	.71
Research Project	71	4.58	3.34	76	6.92	4.01	.70
Mathematics							
Exam	77	5.82	3.21	76	5.47	3.36	-.11
Science							
Exam	70	5.17	3.45	76	6.09	3.87	.27
Research Project	69	5.97	3.01	76	8.24	3.35	.75
Geography							
Exam	75	5.41	3.87	76	3.78	2.75	-.42

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate an achievement advantage for students in Green Team classes. Course grades are coded <13>A+, <12>A, <11>A-, <10>B+, <9>B, <8>B-, <7>C+, <6>C, <5>C-, <4>D+, <3>D, <2>D-, and <1>F.

Table 5

Effects of Instructional Grouping Conditions on Batteries of the California Achievement Test (CAT) Administered in Grade 7 (April, 1991)

	Team						
	Blue			Green			
CAT Battery	n	Mean	SD	n	Mean	SD	Δ
Total Reading	76	751.8	29.8	77	758.0	28.4	.21
Total Language Arts	76	722.1	34.8	77	733.6	37.7	.33
Total Math	77	769.1	25.3	77	769.4	25.6	.01
Total Battery	76	747.6	27.7	77	753.4	28.4	.21
Science	76	717.5	45.5	76	720.3	43.0	.06
Social Studies	76	722.8	37.2	76	727.4	32.5	.12

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Battery scores are scale scores (i.e., units of a single, equal-interval scale that is applied across all levels of the CAT regardless of grade or time of year of testing; these scores are expressed as numbers that may range from 0 through 999 (CTB/McGraw-Hill, 1986). Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate an achievement advantage for students in Green Team classes.

Table 6

Effects of Instructional Grouping Conditions on Connecticut Mastery Test (CMT) Performance

	Team						
	Blue			Green			
CMT Measure	n	Mean	SD	n	Mean	SD	Δ
Grade 7							
Math: Conc Und	66	8.17	2.38	68	7.56	2.87	-.26
L.A.: Read Comp	66	1.27	1.22	70	1.43	1.17	.13
Degrees Reading Power	79	57.30	13.96	75	61.27	13.01	.28
Grade 8							
Math: Conc Und	68	9.04	2.08	69	8.65	2.29	-.19
Comp Skills	68	6.88	2.15	69	6.84	2.51	-.02
Prob Solv	68	7.22	2.37	69	6.93	2.30	-.12
Meas / Geom	69	2.62	1.52	69	2.77	1.54	.10
Overall	68	25.79	6.92	69	25.19	7.34	-.09
L.A.: Writ Mech	69	3.22	1.04	68	3.37	0.79	.14
Study Skills	69	1.55	0.70	68	1.60	0.63	.07
List Comp	69	1.14	0.88	69	1.14	0.86	.00
Read Comp	68	1.94	1.13	69	2.01	1.12	.06
Overall	68	7.93	2.78	68	8.10	2.53	.06
Degrees Reading Power	69	64.23	14.41	69	65.59	13.03	.09
Writing Sample	69	5.26	1.41	68	5.50	1.42	.17

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Components of the Connecticut Mastery Test (CMT; State of Connecticut Department of Education, 1990) are abbreviated as follows: Math: Conc Und = Conceptual Understandings; Comp Skills = Computational Skills; Prob Solv = Problem Solving / Applications; Meas / Geom = Measurement / Geometry; L.A.: Writ Mech = Language Arts: Writing Mechanics; List Comp = Listening Comprehension; Read Comp = Reading Comprehension. Descriptive statistics were calculated on the number of objectives mastered on each component of the CMT. Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate an achievement advantage for students in Green Team classes.

Table 7

Effects of Instructional Grouping Conditions on Self-Concept of Academic Ability

	Team						
	Blue			Green			
Subject Area	n	Mean	SD	n	Mean	SD	Δ
Language Arts							
November	69	20.42	4.22	71	22.01	4.92	.38
February	69	20.33	4.77	71	20.92	4.84	.12
May	69	19.55	5.37	71	20.98	4.60	.27
Math							
November	68	20.46	4.28	68	20.78	3.94	.07
February	68	19.37	4.27	68	19.24	3.40	-.03
May	68	19.31	5.24	68	19.22	4.28	-.02
Science							
November	68	18.68	5.45	69	21.16	4.04	.46
February	68	17.87	5.22	69	19.23	5.33	.26
May	68	18.81	5.89	69	18.72	5.62	-.02
Geography							
November	69	20.12	5.39	68	17.87	5.41	-.42
February	69	20.52	3.78	68	16.32	5.75	-1.11
May	69	21.00	4.05	68	17.65	6.26	-.83

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Indicators for self-concept of academic abilities were summed to form a composite, with scores ranging from 4 (low self-concept) to 28 (high). Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate a self-concept advantage for students in Green Team classes.

Table 8

Effects of Instructional Grouping Conditions on Self-Reported Effort

	Team						
	Blue			Green			
Subject Area	n	Mean	SD	n	Mean	SD	Δ
Language Arts							
November	72	76.2	16.8	68	81.2	18.5	.30
February	72	79.2	15.7	68	78.4	19.5	-.05
May	72	69.3	22.5	68	78.7	20.6	.42
Math							
November	69	76.1	14.9	65	79.1	17.1	.20
February	69	73.2	18.6	65	76.6	13.7	.18
May	69	67.0	23.2	65	72.2	18.2	.22
Science							
November	71	70.9	19.2	66	80.6	14.3	.50
February	71	71.0	21.0	66	74.1	22.9	.15
May	71	69.8	23.6	66	68.3	25.5	-.06
Geography							
November	70	77.6	18.5	67	70.3	22.7	-.39
February	70	79.8	16.0	67	69.2	23.5	-.66
May	70	75.7	15.3	67	68.6	26.4	-.46

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. The indicator for effort ranged from 0 (low effort) to 100 (high). Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate an effort advantage for students in Green Team classes.

Table 9

Effects of Instructional Grouping Conditions on Academic Task Value

	Team						
	Blue			Green			
Subject Area	n	Mean	SD	n	Mean	SD	Δ
Language Arts							
November	71	21.18	5.05	69	22.31	4.78	.22
May	71	20.66	5.90	69	21.71	5.50	.18
Math							
November	69	22.26	4.32	69	22.00	4.29	-.06
May	69	22.23	4.72	69	21.00	4.71	-.26
Science							
November	72	19.11	6.07	67	22.19	4.68	.51
May	72	19.11	6.56	67	17.46	6.63	-.25
Geography							
November	70	21.03	5.70	70	19.08	5.49	-.34
May	70	21.16	5.69	70	17.80	6.84	-.59

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Indicators for academic task value were summed to form a composite, with scores ranging from 4 (low value) to 28 (high). Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); positive coefficients indicate a value advantage for students in Green Team classes.

Table 10

Effects of Instructional Grouping Conditions on Evaluation Anxiety in Academic Subjects

	Team						
	Blue			Green			
Subject Area	n	Mean	SD	n	Mean	SD	Δ
Language Arts	76	8.50	5.14	71	8.90	5.70	.08
Math	79	9.35	4.86	74	8.04	4.59	-.27
Science	77	10.71	5.87	74	8.42	5.43	-.39
Geography	79	10.75	5.25	72	12.74	6.17	.38

Note. Students on the Blue Team were assigned to ability-grouped classes in core academic subjects, with no exposure to Student Team Learning (STL); Green Team students were assigned to heterogeneous classes with STL. Evaluation anxiety was assessed by a 3-item composite, with scores ranging from 3 (low anxiety) to 21 (high). Δ is an effect-size statistic (Glass, McGaw, & Smith, 1981, p.29); negative coefficients indicate an anxiety advantage (i.e., lower evaluation anxiety) for students in Green Team classes.

Appendix A

Indicators of Motivational Constructs

Self-Concept of Academic Ability

How good at [subject] are you?

Not at all good						Very good
1	2	3	4	5	6	7

If you were to rank all the seventh graders in your school from the worst to the best in [subject], where would you put yourself?

The worst						The best
1	2	3	4	5	6	7

How much have your skills in [subject] improved during the past several weeks of school?

My skills have not improved at all						My skills have improved a lot
1	2	3	4	5	6	7

How much natural ability do you have in [subject]?

No ability at all						A lot of ability
1	2	3	4	5	6	7

Academic Task Value

How useful do you think the [subject] you are learning this year will be for you after you finish school?

Not very useful						Very useful
1	2	3	4	5	6	7

How important is it to you to be good at [subject]?

Not at all important						Very important
1	2	3	4	5	6	7

How worthwhile is the effort it will take to do well this year in [subject]?

Not very worthwhile						Very worthwhile
1	2	3	4	5	6	7

How useful do you think high school [subject] will be for what you want to do after you finish school?

Not very useful						Very useful
1	2	3	4	5	6	7

Appendix A (continued)

Indicators of Motivational Constructs

Evaluation Anxiety

Before you take a test in [subject], how nervous do you get?

I'm not nervous at all						I'm very nervous
1	2	3	4	5	6	7

While you take a test in [subject], how nervous do you get?

I'm not nervous at all						I'm very nervous
1	2	3	4	5	6	7

Do [subject] tests scare you?

Not at all						Very much
1	2	3	4	5	6	7

Effort

If a student works to his or her highest potential in a class, then we would say that he or she is putting forth 100% effort to learn the subject matter. How much effort do you usually put forth in [subject] class?

I am not trying at all										I am working to my highest potential
0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%