Teachers' Grouping Practices in Elementary Science Classrooms

CSE Technical Report 456

Noreen M. Webb
University of California, Los Angeles

Gail P. Baxter
University of Michigan, Ann Arbor

Laurie Thompson
Pasadena Unified School District

July 1996

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported in part by grants from the National Science Foundation (Grant Number ES1-90-55443); the Academic Senate on Research, Los Angeles Division, University of California; and the Office of Educational Research and Improvement of the U. S. Department of Education Grant to the Center for Research on Evaluation, Standards, and Student Testing (Grant Number R117G10027). However, all opinions in this paper are those of the authors.

TEACHERS' GROUPING PRACTICES IN ELEMENTARY SCIENCE CLASSROOMS

Noreen M. Webb
University of California, Los Angeles

Gail P. Baxter
University of Michigan, Ann Arbor

Laurie Thompson
Pasadena Unified School District

Abstract

Although small-group learning is an essential part of classroom instruction in nearly all of today's elementary schools little is known about how teachers actually group students in their classrooms. In the study reported here, multiple observations of 30 fifth-grade teachers in a large urban school district revealed that group compositions for hands-on science instruction varied greatly, both between teachers and within classrooms. Generally, teachers tended to form heterogeneous groups given the composition of their classrooms whereas students tended to form homogeneous groups with respect to gender and ethnicity and to a lesser extent achievement level. However, classroom composition (i.e., percentage minority or low achievers) severely constrained teachers' grouping options and in many situations teachers formed groups that contradicted recommended practices. Teachers' stated grouping practices on the basis of student achievement level did not correspond to classroom observations. Further factors not explicitly mentioned by teachers, such as ethnicity, predominated many group compositions. Implications of these findings for practice and further research are discussed.

Introduction

Small-group learning is an essential part of classroom instruction in nearly all of today's elementary schools. Mandates and recommendations for cooperative and collaborative methods of instruction come from school districts, state departments of education, and national research and teacher organizations (e.g., California Department of Education, 1990; California State Department of Education, 1985, 1992; National Research Council, 1989, 1995; National Center for Improving Science Education, 1989; National Council of Teachers of Mathematics, 1989). Coinciding with this burgeoning interest in small-group learning as a pedagogical strategy is research demonstrating its effectiveness for improving achievement, social skills, self-esteem, and attitudes toward others (Bossert, 1988; Slavin, 1990).

In this context, curricular reform efforts in mathematics, science, and other subjects recommend the increased use of student groups for instruction and a variety of published guides and workshops have been developed to advise teachers on how to implement cooperative and collaborative learning in the classroom (e.g., Aronson, Blaney, Sikes, & Snapp, 1978; Burns, 1981; Cohen, 1986; Dishon & O'Leary, 1984; Johnson & Johnson, 1986; Johnson, Johnson, Holubec, & Roy, 1984; Kagan, 1989; Sharan & Sharan, 1992; Slavin, 1990). Although the advice given in these contexts and others is usually to group students heterogeneously on the basis of ability, social, and demographic characteristics (e.g., Kagan, 1992; Slavin, 1990), little systematic research has been done on the nature of group compositions that optimize learning and other outcomes. Even less is known about how teachers actually group students in their classrooms given the large variation in student composition between and within schools. Nor is it known when, why, or if teachers grouping practices deviate from recommended practices.

The purpose of this paper is to document the grouping practices implemented among teachers who heavily emphasize group work in their classrooms. The study reported here was conducted in fifth-grade classrooms in an urban school district in which science instruction at the elementary school level is conducted primarily in groups and in which group work is an explicit component of class work. Classrooms were observed on three different occasions to determine whether groups were heterogeneous or homogeneous on three student characteristics—achievement level, gender, and ethnic/cultural background, to determine the variation of kinds of groups formed in each classroom, and to determine the stability of group

assignments over time. Further, this study examined the correspondence between these classroom observations and teachers' reports about how they grouped students in their classrooms.

In the following section, a review of literature summarizes empirical research on three grouping dimensions—student achievement, gender, and ethnicity and discusses the impact of these student characteristics on the achievement of individuals within a group. This literature review provides a basis for considering the extent to which teachers' grouping practices observed in this study correspond to suggested practice stemming from these empirical studies. Because classrooms are more or less homogeneous in terms of student characteristics, teachers may be constrained in how they can form groups. Given that the extant literature may offer little specific guidance in these situations, how then do teachers form groups for instruction? In this study, teachers' grouping practices are considered within the context of their classroom defined by several student characteristics such as percentage of minority students or females in the class. Understanding the correspondence between teacher grouping practices and empirical research that addresses the relationship between group composition and achievement (1) lays the groundwork for research that can inform instructional practice and (2) provides concrete feedback to teacher leaders and others concerned with staff development about the ways in which teachers adapt recommended grouping practices for use in their classrooms. Results of this study may also draw attention to the implications of various district and/or school policies for assigning students to schools or classrooms.

Achievement and ability level.

Most empirical research on group compositions has focused on the mixture of ability or achievement levels in the group. It is commonly believed, and the research generally concurs, that heterogeneous groups will benefit lower-achieving students by giving them access to the intellectual resources of higher-achieving students. Studies in a number of subject matters and using a variety of tasks have shown that low-ability or low-achieving students learn more in heterogeneous groups than in homogeneous groups (e.g., developing the concept of conservation, model-building tasks, mathematical balance beam tasks, mathematical problem solving, computer learning (Bell, Grossen, & Perret-Clermont, 1985; Doise & Mugny, 1984; Hooper & Hannafin, 1988; Mugny & Doise, 1978; Murray, 1982; Perret-Clermont, 1980; Tudge,

1989; Webb, 1980). An examination of the verbal interaction among students in heterogeneous groups showed that low-achieving or low-ability students in the group benefit from receiving explanations and other kinds of help from high-achieving or high-ability students in the group (e.g., Azmitia, 1988; Tudge, 1989; Webb, 1980).

Not all research, however, has come to the same conclusion about the benefits of heterogeneous grouping for low-ability students. Some studies have found heterogeneous grouping to be no more effective than homogeneous groups for low-ability or low-achieving students (e.g., Ames & Murray, 1981; Glachan & Light, 1982; Hooper, Ward, Hannafin, & Clark, 1989). The observations made by Mugny and Doise (1978) of pairs of students learning how to construct models of a village with different spatial orientations offer one explanation for why low-ability students may not always benefit from heterogeneous groups. Mugny and Doise found that low-ability students learned more when paired with medium-ability students than when paired with high-ability students. They observed that medium-ability students involved low-ability students in the collaboration and gave them explanations, whereas high-ability students directed group work without explaining how to solve the problems. Mugny and Doise also suggested that low-ability students may not have understood explanations by high-ability students, even when they were offered, due to the wide gap in understanding between the high and low students.

While using heterogeneous groups for low-ability or low-achieving students is generally not controversial, using heterogeneous groups for high-ability or high-achieving students is. Educators, and even more, parents, worry that heterogeneous grouping will hold back high-ability students (Oakes, 1990). Empirical research does not support this belief, however. High-ability and high-achieving students show equally strong learning outcomes in heterogeneous and homogeneous groups (e.g., Azmitia, 1988; Hooper & Hannafin, 1988; Hooper et al., 1989; Skon, Johnson, & Johnson, 1981). There is even evidence that high-ability or high-achieving students benefit more from working in heterogeneous groups than from working in homogeneous groups or individually due to the teacher roles they adopt in heterogeneous groups (e.g., Bereiter & Scardamalia, 1989; Webb, 1991). Giving explanations to less-able students helps more-able students reorganize and reconceptualize material, develop new perspectives, and solidify their understanding (Bargh & Schul, 1980; Benware & Deci, 1984; Yackel, Cobb, & Wood, 1991).

The benefits of different group compositions for medium-ability or medium-achieving students are less clear cut. Some studies have found that medium-ability students may be excluded from the helping relationships that develop between high and low members of heterogeneous groups, with consequent detrimental effects on their learning (see review by Webb, 1991). These same studies show that medium-ability students may participate more, and learn more, in groups that are homogeneous or that mix fewer ability or achievement levels together in a group (e.g., lows and mediums, or mediums and highs, instead of highs, mediums, and lows).

Gender.

Not only does the relative ability or achievement of group members influence group processes and achievement outcomes, but there is also evidence that group composition based on social characteristics, such as gender and ethnic background, may influence group dynamics and learning outcomes. Gender has been shown to influence group interaction and achievement, with boys tending to dominate interaction in mixed-gender groups (Hazelwood, Roth, Hasbach, Hoekwater, Ligett, Lindquist, Peasley, & Rosaen, 1992; Lockheed & Harris, 1984). Not all mixed-gender groups function in the same way, however. In classes with predominantly high-achieving white students, Webb (1984a) found that girls participated less and learned less in mixed-gender groups with an unequal number of girls and boys (either girls outnumbering boys or boys outnumbering girls) than in groups with an equal number of girls and boys. Only in groups with an equal number of girls and boys did girls learn as much as boys. These effects were not related to academic competence; in all group compositions, girls and boys showed equal ability and achievement at the outset of the study. A similar study in classes with predominantly low-achieving African-American students, however, showed no effects of gender or gender composition on either group interaction or achievement (Webb & Kenderski, 1985; see also Grant, 1986).

Ethnic/cultural background.

One of the original goals of cooperative learning theory and research was to use heterogeneous groups to facilitate improved cross-race relations in mixed classrooms and schools (Sharan, Kussell, Hertz-Lazarowitz, Bejarano, Raviv, & Sharan, 1985). However, heterogeneous groups may highlight differences in members' status and produce differences in influence, and thus undermine the

desired goals. According to Expectations States Theory (Berger, Rosenholz, & Zelditch, 1980), when members of a group do not know each other and do not have a way to judge each others' competence on the task, they will often rely on social characteristics to determine status and influence in the group. Laboratory studies carried out by Cohen (1982) offer some support for this theory. In these laboratory studies white students dominated interaction in multiracial groups by talking more, being more assertive, and by giving more ideas and information.

To circumvent differences in status and influence in the group, Miller and Harrington (1990) suggest forming heterogeneous groups so that social characteristics such as race are less salient. They suggest forming groups with as much diversity as possible (e.g., four ethnic backgrounds instead of two), or varying several social characteristics simultaneously (e.g., a black male, a white male, a black female, and a white female; instead of two black females and two white males). Miller and Harrington also advise against "solo status" groups, such as groups with a single minority student, because it makes differences in social characteristics especially obvious.

Cohen (1982, 1994) emphasized that social characteristics such as race and ethnic background are most likely to define status and influence behavior in the group when group members do not know each other or each others' levels of competence. When group members know each other, as would be the case in many classrooms, Cohen (1982, 1994) argues that perceived academic competence (academic status) and popularity (peer status) are more likely than social characteristics to determine status and influence. The power of perceived academic competence was clearly demonstrated in a study by Dembo and McAuliffe (1987) who "determined" competence randomly by assigning students fictitious scores on a pretest. Students who were "higher" on this artificial ability variable initiated more activity and gave more help in heterogeneous groups than did students who were "lower" on this artificial ability. As Cohen (1982, 1994) describes, it is difficult to separate the effects of academic and peer status from those of social characteristics such as race and ethnic background because race and ethnic background are often correlated with academic competence.

The limited research on the effects of group composition on group dynamics and learning offers few guidelines for creating or forming classroom groups to optimize learning. There is general consensus that some kinds of groupings should be avoided, such as homogeneous low-ability groups, placing medium-ability

students in wide-range heterogeneous groups, and forming groups with a "solo" low-status student. But the existing research is silent about the relative effectiveness of other group compositions. The present study, then, investigated the options teachers had for forming groups and the kinds of groups they formed in their classrooms. While researchers have investigated a variety of group compositions, teachers must work within the constraints of their own classrooms and may have the option to form some kinds of groups but not others. This study documents the grouping practices of teachers in an urban school district where grouping is mandatory for science instruction because of limited materials and a philosophy that group work is an effective strategy for helping students learn and understand science.

Method

Sample

A large, urban school district in Southern California, characterized by its diverse ethnic, linguistic, and socio-economic makeup, served as the site for this study. The district serves about 22,500 students of whom 44% are Latino, 35% are African-American, 17% are Anglo and 4% are Asian and other. One quarter of the students are classified as Limited English Proficient (LEP) and the majority of students (55%) qualify for federally supported free or reduced-cost lunch programs.

Thirty fifth-grade classrooms (845 students) from 16 elementary schools participated in this study. All participating teachers had attended twelve days of inservice instructional activities over a two-year period designed to develop teachers' understanding of science content, the nature of inquiry, and how to use effective pedagogical strategies (e.g., questioning, cooperative groups, science journal writing) to help students learn the key concept and process skills that were specific to each science unit.

Class size ranged from 23 to 42, with an average of 31. Due to student absence, the number of students per class used in the analyses was slightly lower, ranging from 23 to 37. The ethnic makeup of the sample was 38% Latino, 32% African-American, 20% Anglo, 3% Asian-American, 3% from other ethnic/racial backgrounds, and 5% whose ethnicity was unknown.

Science Instruction

Elementary science instruction in this district centers around four inquiry-based units at each grade level. In the fifth grade, students work in small groups to carry out investigations of the behavior of crayfish; the nature and function of electric circuits; the physical and chemical properties of common substances; and the astronomical relationship between the sun and the earth. Teachers guide students in using relevant equipment and materials, such as batteries, bulbs, wires, chemical indicators, or shadow boards to carry out a sequence of activities designed to build understanding of a particular topic.

As part of teachers' formal preparation on how to implement these investigations in their classrooms, they receive general guidance about how to group their students. Teachers are told that students should work in groups of four (science materials are prepared for groups of four) and that it is desirable to form groups that are heterogeneous in terms of academic level, ethnic background, and gender. The rationale is that groups of this sort contribute a variety of life experiences, prior knowledge, and perspectives to each activity thereby maximizing learning opportunities for all students. Despite this rule of thumb, teachers are told to use their judgment and form homogeneous groups for students who may not otherwise participate actively, such as forming all-girl groups for girls who do not participate when grouped with boys, or pairing students of the same background to make them more comfortable when the activity involves a sensitive issue for a particular racial or ethnic group. Further, teachers are told that the science groups do not have to be the same as the reading groups and that they can decide how frequently to change group membership.

Data Collection Procedures

Standardized achievement test scores, survey, and observational data were collected in each classroom. The district supplied the most recent reading, mathematics, and language percentile scores on the Comprehensive Test of Basic Skills (CTBS) for students in the sample.[1] Prior to classroom observations in the study, a survey was mailed to each teacher asking them to rate each of their students on: (a) competence in reading, writing, mathematics, and science (high, medium, or

[1] Normal curve equivalents (NCEs) would be preferable to percentiles in the analyses because equal changes in students' raw scores correspond to equal changes in NCE scores (but not in percentile scores; see Linn & Gronlund, 1995; Oosterhof, 1994), but they were not available from the district.

low), (b) motivation, responsibility, and behavioral problems (yes or no), and (c) special needs (e.g., limited English proficient). Classroom observations of grouping practices were made on three occasions approximately one month apart. A science resource teacher from the school district visited each classroom and noted the grouping arrangements, including the name of each student in a group, the number of students in each group, and the number of groups in the class. During the visit, the classroom teacher was asked by the science resource teacher to indicate how groups were formed for science instruction. In addition to asking this general question about the teacher's grouping practices, the resource teacher pointed to a group and asked how that particular group was formed. The intent was to gain some clarification of the teacher's grouping rule with minimal disruption to classroom schedules.

Determining Achievement Variables for Analysis

Preliminary analyses of CTBS scores and teachers' ratings of student achievement level showed that: (a) the three CTBS scores (reading, mathematics, language) were highly intercorrelated (r's ranged from .76 to .81), (b) the four teacher ratings (reading, writing, mathematics, and science) were highly intercorrelated (r's ranged from .58 to .79), and (c) the CTBS scores were highly correlated with teachers' ratings (r's ranged from .61 to .66). The high correlations among scores showed that teachers' ratings agreed substantially with standardized test scores, and that scores or ratings in different subject-matter areas were highly related. Consequently, a single composite measure based on teachers' ratings was used as the measure of student achievement in this study.

For purposes of analysis, the composite teacher rating—the mean of the four teacher ratings—was categorized as low (1.00 and 1.50), medium (1.75 to 2.50), or high (2.75 to 3.00). To be classified at a particular level, a student had to receive the majority of ratings at that level or above. Students classified as low had two, three, or four ratings at the low level and the rest at the medium level. Students classified as medium typically had the majority of their ratings at the medium level. Of the students classified as medium, 73% had three or four ratings at the medium level, 19% had two medium ratings, and 8% had one medium rating and the remainder high and low ratings. Students classified as high had three out of four ratings that were at the high level and the remainder at the medium level. The percentage of the sample categorized at each level was 27% low, 50% medium, and 24% high. This

distribution corresponds closely to the 25%-50%-25% decision rule often used in previous research for classifying students into ability levels (e.g., Webb, 1989, 1991; Peterson, Wilkinson, Spinelli, & Swing, 1984).

Some students (2% of the sample) had neither CTBS scores nor teacher ratings. Analyses were conducted to determine whether other data available on these students such as gender, ethnic/cultural background, teacher, or school could be used to form a regression estimate of their composite achievement rating. However, in the overall sample, students' CTBS scores and teacher ratings were only weakly related to student background characteristics, so this information could not be used to estimate achievement level. For students with neither CTBS scores or teacher ratings, no determination of achievement score was possible.

Results and Discussion

In this study, three student characteristics (i.e., achievement level, gender, and ethnicity) were considered as possible factors in teachers' grouping strategies. The classrooms in this study differed markedly in terms of the variation in student achievement level and ethnicity and to a lesser extent gender. Some classrooms were heavily represented by one ethnic background (Latino, African-American, or Anglo). The student populations in other classrooms exhibited a mixture of ethnic backgrounds, with no ethnic background in the majority. Regardless of ethnic distribution, some classrooms had students with a wide range of achievement level, and others had students who were primarily at or below grade level. Because the composition of the class places constraints on how groups can be formed in them, the results for group composition in this section are presented by ethnic composition of the class. In subsequent sections, results are examined with respect to the gender and achievement distributions within a class.

Teachers' Rules for Forming Groups

Classes were categorized as one of two types. "Majority" classes had more than 65% of students with the same ethnicity. "Heterogeneous" classes had less than 65% of students of the same ethnic background and each ethnic background is represented by at least 10% of the students. There were three types of majority classrooms: Latino, African-American, and Anglo. And there were two types of

heterogeneous classes: those with Latino, African-American, and Anglo students; and those with predominantly Latino and African-American students.

Table 1 gives the number of teachers by category of classroom who used each grouping rule. The term "grouping rule" here and throughout this paper refers to teachers' stated decision-making strategies for forming groups. All teachers of majority ethnic classrooms assigned students to groups. Some teachers of

Table 1
Teachers' Rules for Forming Groups

Grouping Rule	Majority Ethnic Classrooms[a]			Heterogeneous Ethnic Classrooms[b]	
	Latino	African-American	Anglo	Latino, African-American, Anglo	Latino, African-American
Teachers choose	4[c]	5	2	8	4
Homogeneous gender	1	0	0	0	0
Homogeneous achievement	1	0	0	0	0
Homogeneous gender, mix achievement	0	0	0	1	0
Mix gender	0	0	0	1	0
Mix achievement	1	1	0	2	1
Mix gender, mix achievement	1	1	0	1	0
Proximity	0	0	1	2	2
Random	0	0	0	0	1
Unknown	0	3	1	1	0
Students choose	0	0	0	4	3

[a]Greater than 65% of students in the class with the same ethnicity.
[b]Less than 65% of students in the class with the same ethnicity; at least 10% with each ethnicity.
[c]Number of teachers.

heterogeneous ethnic classrooms assigned students to groups; others let students choose their own groups. When teachers formed groups, they used a variety of rules based on gender and achievement levels. A few teachers stated that they formed groups based on where students sat in the classroom (proximity) whereas others did not report enough information to determine their grouping rule (unknown). No teacher mentioned "ethnic background" as a factor in assigning students to groups. The results described in the following sections examine the extent to which the groups observed in the classrooms conformed to the teachers' stated grouping rules and whether teachers and students differed in the types of groups they formed.

Group Size[2]

Group sizes ranged from 1 to 8 in the following distribution: 1 (6% of the groups), 2 (20%), 3 (19%), 4 (44%), 5 (9%), 6 (2%) and 8 (less than 1%). The most popular group size was four, with two-person and three-person groups occurring frequently. Kits of science equipment were packaged for groups of four students, which accounts for the large proportion of 4-person groups. The group sizes other than four (especially pairs and groups with six and eight students) occurred mostly in classrooms studying a unit on crayfish. Rather than one set of materials per group, as is typical for science instruction, crayfish are kept in large tubs on a central table in each classroom to facilitate behavioral observations by students. The teacher can then decide how large the observation groups should be. Some teachers created large observation groups of 6 or 8 students and other teachers created small observation groups of 2 students.

About two percent of the students (n = 15) worked alone. Table 2 gives the characteristics of students who worked alone and students who worked in groups. Students who worked alone had lower academic achievement than students who worked in groups, both in terms of standardized test scores ($t(762) = 2.87$, $p = .004$) and in terms of teachers' ratings of achievement ($t(812) = 2.24$, $p = .025$)[3]. Whether students worked alone was also significantly related to their being rated by the

[2] The results reported in this and the following sections are based on observations from the first of three visits to the classrooms. As described more fully in a later section, because the results across the three visits were not statistically different, the results for the second and third visits are not presented here. Tables of results are available from the author upon request.

[3] Differences in degrees of freedom across statistical tests is due to missing data. Some students did not have standardized test scores or some teacher ratings.

teacher as having behavior problems ($chi^2(1) = 6.84$, $p < .01$). Working alone versus working in groups was not significantly related to teachers' ratings of motivation ($chi^2(1) = 3.44$, $p = .06$), teachers' ratings of responsibility ($chi^2(1) = 0.43$, $p = .51$), gender ($chi^2(1) = 0.16$, $p = .69$), or ethnic background ($chi^2(1) = 2.44$, $p = .12$). It appears that teachers' concerns for classroom management may have outweighed the potential benefits of grouping the students for whom working in groups may be difficult: relatively low-achieving students with behavioral problems.

Table 2
Characteristics of Students Who Worked Alone and in Groups

	CTBS Scores[a]		Teacher Ratings[b]		Behavior Problems	Low Motivation	Not Responsible	Female	Ethnic Minority[c]
	M	SD	M	SD	%	%	%	%	%
Worked Alone	23	16	1.68	73	47	33	29	53	93
Worked in Groups	44	28	2.05	.63	19	16	21	48	73

[a]Percentile scores. CTBS = Comprehensive Test of Basic Skills.
[b]Three-point scale: 3 = high; 1 = low.
[c]Latino and African-American vs. Anglo and Asian and other ethnic backgrounds.

Group Composition on Ethnic/Cultural Background

Groups within a classroom were classified as homogeneous if all students in the group had the same ethnicity, majority if more than 50% of the students in the group had the same ethnicity, and heterogeneous if 50% or less of the students in the group had the same ethnicity. No teacher mentioned using ethnic background as a factor in assigning students to groups, so we could not examine whether group compositions on ethnicity corresponded to teachers' grouping rules. The only comparison possible was between classes where students chose their own groups and classes where teachers assigned students to groups. Table 3, then, shows the

distribution of students by group composition for classes with student-chosen groups and for classes with teacher-assigned groups.

In majority-ethnic classrooms, all teachers assigned students to groups. Because by definition more than 65% of the students in these classes were of the same ethnic background, not surprisingly most groups were homogeneous or had a majority of students with the same ethnicity. This is particularly apparent in majority-Latino classrooms where a prevalence of homogeneous groups stemmed from a high proportion (92%) of Latino students in these classrooms. Majority-African-American and majority-Anglo classrooms, with approximately a 70-30 distribution of majority to other ethnic groups, showed a prevalence of majority groups. Although teachers did not indicate that ethnicity was a factor in determining groups, the fairly low percentage of homogeneous groups in these classrooms suggests that teachers tried to mix ethnic backgrounds in each group to the extent that classroom composition (percentage of various ethnic backgrounds) permitted.

Heterogeneous classes had students with a mixture of ethnic backgrounds thereby giving teachers (or students) opportunities to create all of the kinds of groups represented in Table 3. When teachers formed groups, they tended to form groups that were heterogeneous and rarely formed groups that were homogeneous. The pattern for student-chosen groups depended on the composition of the class. For heterogeneous Latino, African-American, Anglo classes, students were equally likely to form homogeneous, heterogeneous or majority groups. Group composition was not significantly related to whether groups were teacher-chosen or student-chosen in these classrooms ($chi^2(3) = 5.92$, $p = .11$). Both teachers and students formed a mix of groups with no one type of group predominating. In heterogeneous Latino, African-American classes, students tended to put themselves in homogeneous groups whereas teachers tended to put students in heterogeneous groups. The relationship between group composition and whether groups were teacher-chosen or student-chosen in these classrooms was statistically significant ($chi^2(3) = 10.74$, $p = .01$).

The results given in Table 3 for student-chosen groups mask important and statistically significant variation between classes. Students formed homogeneous groups much more frequently in some classes than in others. In heterogeneous-Latino, African-American, Anglo classes, the differences in percentages of homogeneous groups varied markedly between classes: two classes had low percentages of homogeneous groups (0% and 13%), whereas two classes had much

higher percentages (44% and 57%). Similarly, in heterogeneous-Latino, African-American classes, one class had no homogeneous groups and the other two classes had very high percentages of homogeneous groups (69% and 71%). The relationship between classroom and group ethnic composition was statistically significant for

Table 3
Distribution of Ethnic Background and Ethnic Group Compositions

	Ethnic Backgrounds (Percentage of Students)				Ethnic Group Compositions (Percentage of Groups)			
Classroom Ethnic Composition	Latino	African-American	Anglo	Asian	Hom.[a]	Major.[b]	Heter.[c]	Unknown
Majority[d]								
Latino	92	2	1	0	78	19	0	3
African-American	19	73	6	1	21	62	14	3
Anglo	10	10	71	9	20	47	33	0
Heterogeneous (Latino, African-American, Anglo)								
Teachers chose groups	31	28	27	4	16	24	55	5
Students chose groups	22	26	39	2	33	28	33	7
Heterogeneous (Latino, African-American)								
Teachers chose groups	40	42	6	1	11	32	43	14
Students chose groups	51	33	5	4	50	21	18	11

Note: Percentages do not always sum to 100 because students with unknown ethnic background are not included here.
[a] All students in the group have same ethnicity.
[b] More than 50% of the students in the group have same ethnicity.
[c] 50% or less of the students in the group have same ethnicity.
[d] Teachers chose groups in all majority ethnic classrooms.

both the heterogeneous-Latino, African-American, Anglo classes (four classrooms: $chi^2(9) = 17.92$, $p = .04$) and the heterogeneous-Latino, African-American classes (three classrooms: $chi^2(6) = 22.51$, $p = .001$). These results show that students in some classes, but not all, tended to segregate themselves according to ethnic background.

The classes in which students segregated themselves by ethnic background had significantly lower CTBS scores than the classes in which students did not segregate themselves by ethnic background (segregated classes: $M = 39$, $SD = 27$; nonsegregated classes: $M = 60$, $SD = 28$; $F(1,195) = 26.83$, $p < .0001$). Classes in which students segregated themselves by ethnic background also had lower teacherratings of achievement (segregated classes: $M = 2.12$, $SD = .69$; nonsegregated classes: $M = 2,33$, $SD = .58$; $F(1,198) = 5.22$, $p = .02$). Slavin (1983, 1990; see also Gerard & Miller, 1975) described how black, white, and Hispanic students in mixed-ethnic schools are much more likely to have friends of their own ethnic background than to form cross-ethnic friendships, and how they often compete with each other for teacher approval, grades, and positions of power in school. The findings reported here, however, suggest that this mechanism may operate in low-achieving schools but not necessarily in higher-achieving schools. It is possible that in higher-achieving schools students of different ethnic backgrounds may be more likely to work together because they recognize that all students have some academic competence and skills and expertise to contribute. Or perhaps students in higher-achieving classes perceive ethnic background as a less salient distinguishing characteristic than do students in lower-achieving classes. Interviewing students about their reasons for forming groups would help clarify this issue.

Gender Group Composition

Among majority-Latino, majority-African-American, and majority-Anglo classrooms, the type of classroom was independent of gender group composition ($chi^2(6) = 5.59$, $p = .48$). Consequently, these three types of classrooms were combined as majority ethnic in Table 4. Similarly, among heterogeneous-Latino, African-American, Anglo and heterogeneous-Latino, African-American classrooms, because type of classroom was not related to gender group composition ($chi^2(3) = 0.30$, $p = .96$), these two types of classrooms were combined as heterogeneous ethnic in Table 4.

As shown in Table 4, teachers who reported forming homogeneous or mixed groups on gender did follow their reported decision-making rule. Teachers who reported forming same-gender groups formed mostly homogeneous groups. Similarly, the teachers who reported forming mixed-gender groups did form mostly mixed-gender groups, although the teachers' definition of mixed is not clear from the data. Teachers may have used mixed to mean a variety of gender groups, including some with more males or more females as well as some groups with equal numbers of males and females.

What is apparent is that teacher-assigned groups in heterogeneous-ethnic classrooms were more likely to be homogeneous with respect to gender and less likely to have disproportionate numbers of males or females in a group (i.e., more boys or more girls) than teacher-assigned groups in majority-ethnic classrooms. Among teachers who did not mention using gender to form groups, the groups were mostly mixed on gender.

Statistical tests showed, first, that group gender composition was significantly related to whether or not teachers said they formed same-gender groups ($chi^2(6) = 28.9$, $p < .0001$ for majority classrooms; $chi^2(6) = 25.54$, $p < .0001$ for heterogeneous-ethnic classrooms). Second, group gender composition was not significantly related to whether teachers said they formed mixed-gender groups or did not mention gender as a grouping rule ($chi^2(3) = 1.20$, $p = .76$ in majority-ethnic classes; $chi^2(3) = 3.82$, $p = .29$ in heterogeneous-ethnic classes). In other words, even when teachers did not explicitly state that they considered student gender when forming groups, classroom observations suggest that indeed gender was a determining factor in assigning students to groups.

In a substantial portion of the teacher-assigned groups in this study (21%), girls were the "solo girl" in a group with several boys, a position that has been shown in some previous research to be detrimental for learning and performance (see Webb, 1984a). A number of these groups were found in one classroom. The other "solo girl" groups were spread across the remaining classrooms, with at most one or two appearing in any one classroom. The one class where a large number of solo groups were observed had more boys than girls (68% vs. 32%). Given this gender imbalance, the teacher's options would be to concentrate girls in a few groups and have a number of groups consisting only of boys or to spread the girls across the groups to create mixed-gender groups. This teacher followed the latter strategy: 90% of the groups in this class were mixed, although the majority had only one girl. This

teacher did not report using gender as a factor when grouping students, but these results suggest a perhaps conscious (but unstated) strategy of mixing girls and boys in groups even with a relatively small number of girls in the class.

As can be seen in Table 4, when students chose their own groups, they usually formed same-gender groups. In the seven classes in which students formed their own groups, only one class had a low proportion (11%) of homogeneous groups with respect to gender. This class had a variety of mixed-gender groups: equal male-equal female (33%), solo-female (44%), and solo-male (11%). It is interesting that even in the face of an imbalance of boys and girls in this class (62% of the students were boys), students most often chose mixed-gender groups instead of all-boy groups. The remaining six classes showed very high percentages of same-gender groups (ranging from 77% to 100%). The findings for teacher-selected vs. student-selected groups suggest that teachers' efforts to form mixed-gender groups were at odds with students' preferences for same-gender peers as grouping partners, a factor that has not been considered in previous research.

Table 4
Distribution of Gender and Gender Group Compositions

	Gender (Percentage of Students)		Gender Group Compositions (Percentage of Groups)			
Classroom Composition and Gender Grouping Rule	Female	Male	Hmong Gender	More Girls	Equal Gender	More Boys
Majority Ethnic[a]						
Homogeneous	38	62	78	22	0	0
Mixed	54	46	9	27	36	27
No gender rule	51	49	9	32	21	38
Heterogeneous[b]						
Homogeneous	32	68	100	0	0	0
Mixed	48	52	30	15	45	10
No gender rule	57	43	19	16	34	31
Students choose	53	47	79	6	7	7

[a]Includes majority Latino, majority African-American, and majority Anglo classrooms.
[b]Includes heterogeneous Latino, African-American, Anglo; and heterogeneous Latino, African-American classrooms.

Achievement Group Composition

Table 5 gives the percentage of students at each achievement level (as rated by the teacher), mean and standard deviation in CTBS scores, and the percentage of groups in each achievement group composition. The following group compositions were defined according to achievement: (1) homogeneous groups had students at one achievement level, (2) medium-range groups had high- and medium-achieving students, or medium- and low-achieving students, (3) one type of wide-range group had high-, medium-, and low-achieving students, and (4) the other type of wide-range group had high-achieving and low-achieving students.

Because the three types of majority ethnic classrooms differed significantly on CTBS scores ($F(2,245) = 93.15$, $p < .00001$) and on the distribution of achievement level as rated by the teacher ($chi^2(4) = 27.86$, $p < .0001$), they are presented separately in Table 5. Similarly, because the two types of heterogeneous ethnic classes different significantly on CTBS scores ($F(1, 514) = 56.73$, $p < .00001$) and on the distribution of achievement level as rated by the teacher ($chi^2(2) = 17.08$, $p = .0002$), they are also presented separately in Table 5.

As can be seen in Table 5, the teacher who reported forming groups that were homogeneous in achievement level did not do so. Only one group in that class was homogeneous; the others were mixed, either medium-range groups or wide-range groups. This teacher said that she usually tried to keep "gifted" students (those who were identified as part of the Gifted and Talented Education program) together. However, the homogeneous group in her classroom was medium, not high. High-achieving students were distributed across heterogeneous groups, sometimes in groups with low-achieving students and sometimes in groups with medium-achieving students. One reason for the discrepancy may be the teacher's definition of homogeneous. In two out of three of the medium-range groups, all but one student came from the same achievement level (for example, one medium-achieving student and four high-achieving students). The teacher may have considered these groups to be homogeneous in achievement because the majority of the students in the group were of the same achievement level. Even though the majority of students in a group may have had the same achievement rating as judged by the teacher (i.e., homogeneous) the range of CTBS scores in these groups was very large (i.e., heterogeneous). In 57% of the groups in this class, the range of CTBS scores in each

group exceeded 48 points out of a possible 100. So these groups were not homogeneous in any absolute sense.

Teachers who reported forming groups that were mixed in achievement generally did form mixed-groups. Most of these groups were medium-range groups, not wide-range groups. Among teachers who did not report using achievement level to group students, the percentage of heterogeneous groups varied. Most of the groups in those classes were also medium-range groups.

Teachers and students tended to form similar kinds of groups. In heterogeneous-Latino, African-American, Anglo classes, group composition was not significantly related to achievement grouping rule ($chi^2(8) = 11.20$, $p = .19$). In heterogeneous-Latino, African-American classes, a significant relationship between group composition and grouping rule appeared ($chi^2(8) = 19.81$, $p = .01$); it was due to the large proportion of groups of unknown achievement composition in the three classes in which teachers did not explicitly use achievement to group students (see right hand column of Table 5). In these three classes, 10% of students had neither standardized test scores nor teacher ratings of achievement. These students were distributed across groups, resulting in the relatively high proportion of groups with unknown achievement composition. When groups with unknown achievement composition were omitted from the analysis, the relationship between group composition and grouping rule was not statistically significant ($chi^2(6) = 8.89$, $p=.18$).

Students tended to form different kinds of groups in the two types of heterogeneous ethnic classrooms. When students formed their own groups in heterogeneous-Latino, African-American, Anglo classes, they formed mostly mixed-achievement groups and few homogeneous groups (13%). In heterogeneous-Latino, African-American classes, students tended to form more homogeneous groups (46%). Among classes in which students chose groups, the distributions of group compositions were significantly different in the two types of heterogeneous ethnic classes: group composition was significantly related to the type of class ($chi^2(4) = 12.83$, $p = .01$).

Five percent of the groups in this study consisted of low-achieving students. Few of them appeared in classes in which teachers chose groups. The majority of them (64%) appeared in classes in which students chose groups. One possibility

Table 5
Distribution of Achievement Levels and Achievement Group Compositions

Classroom Composition and Achievement Grouping Rule	Teacher Ratings of Achievement (Percentage of Students)			CTBS Scores		Achievement Group Compositions Based on Teacher Ratings (Percentage of Groups)				
	Low	Med	High	*M*	*SD*	Hom.[a]	Med.[b]	Wide (HML)	Wide (HL)	Unknown
Majority Latino										
Homogeneous	27	43	30	32	26	14	43	43	0	0
Mixed	44	48	8	12	5	19	50	19	0	12
No ach. rule	31	62	7	27	13	44	44	0	11	0
Majority African-American										
Mixed	25	66	9	20	16	0	80	10	0	10
No ach. rule	12	57	31	43	22	16	53	21	5	5
Majority Anglo										
No rule	12	48	40	72	23	0	67	27	7	0
Heterogeneous (Latino, African-American, Anglo)										
Mixed	34	47	19	50	24	15	52	27	0	6
No ach. rule	30	45	25	44	27	31	52	10	3	3
Students choose	13	43	44	58	29	13	63	10	8	8
Heterogeneous (Latino, African-American)										
Mixed	13	69	19	44	21	33	44	22	0	0
No ach. rule	32	63	5	30	18	16	47	5	5	26
Students choose	41	41	18	31	21	46	46	0	7	0

[a] Homogeneous: All students in group have same ability level.
[b] High-medium or medium-low groups.

suggested by this result is that teachers tended to avoid the potentially harmful practice of placing low-achieving students together, whereas students were not reluctant to group themselves this way. On the other hand, teachers may have considered classroom management—high achievers monitoring or helping low

achievers—as the most important concern in forming groups, a concern not necessarily shared by students.

The results just described for group formation on the basis of achievement level used teacher ratings and not CTBS scores to define achievement level. Nearly all teachers used the full range of ratings for the students in their classrooms, rating them from low to high, regardless of the range of achievement level in the class as indicated by students' CTBS scores. Classes varied considerably, however, in the distribution of CTBS scores. As such, some teachers made low-medium-high distinctions among primarily low-achieving students and other teachers made low-medium-high-distinctions among students who varied from low-achieving to high-achieving as indicated by their scores on the standardized test. As a consequence, "mixed" groups in some classrooms were more homogeneous than "mixed" groups in other classes; some heterogeneous groups were in effect homogeneous and vice versa.

As an example, two of the classes whose teachers formed mixed groups had similar distributions of teacher ratings (27% low, 60% medium, 13% high, and 25% low, 56% medium, 19% high, respectively) but had much different mean CTBS scores: the first was extremely low (16.69) and the second was about average on national norms (53.21). CTBS scores in the first class ranged from 1 to 63, but only two students had scores above 39: 93% of the students in this class had scores *below* 39! In the second class, CTBS scores ranged from 8 to 93, with 78% of the scores *above* 39. Nearly all of the students in the first class had CTBS scores comparable to the "low" students in the second class. Although the groups in the first class were all mixed according to the teacher's ratings, most or all of them (especially the low-medium groups) would be considered as homogeneous-low according to the second teacher's ratings. For example, the maximum CTBS score in any of the low-medium groups in the first class was 36. These groups would be homogeneous-low according to the classification rules of the second class. Even the low-medium-high group in the first class, with a range of CTBS scores of 1 to 39, might be considered a homogeneous-low group according to the classification rules of the second class. This example shows that, especially in predominantly low-achieving classrooms, many groups that were classified as mixed were more homogeneous than the teacher rating data suggest. This example shows the tendency for many teachers to view achievement level as relative to the distribution of achievement in their class, not relative to the fifth-grade population in the district.

Multiple Characteristics Considered Simultaneously

Discussion of teachers grouping practices to this point has largely considered group compositions on one student characteristic at a time, first ethnic background, then gender, and then achievement level. Nevertheless, groups may vary simultaneously on one or more of these characteristics. It is important to consider how groups varied on multiple characteristics, especially whether groups that were homogeneous on one characteristic were also homogeneous on others. It is possible, for example, that teachers (or students) prefer homogeneous groups for some types of students (e.g., Latino females) but not for others. Comparisons are made here between teacher-assigned and student-chosen groups with respect to the overall homogeneity of the groups.

Recall from previous discussion that teachers attempted to form heterogeneous groups on any given characteristic, with the exception of two teachers who explicitly formed same-gender groups. One of these teachers who formed same-gender groups also tended to form homogeneous groups on the basis of ethnicity and achievement as well. Students in this class were predominantly Latino (31 out of 32 students), so all groups were constrained to be homogeneous in ethnicity. This classroom also had a fairly high percentage of homogeneous-achievement groups (44%), all at the medium-achievement level according to the teachers' ratings. Again this is a function of the achievement distribution in the class where most of the students were rated medium (63%), some were rated low (28%) and only a few were rated high (10%). The preponderance of medium-achieving students placed constraints on the kinds of groups that could be formed, making it almost impossible to avoid forming some homogeneous-medium groups. In the other classroom in which the teacher formed same-gender groups, all groups were mixed on ethnic background and 88% were mixed on achievement level. Unlike the first class described above, this class is best described as multiethnic with a full range of achievement levels represented. Consequently, this teacher could consider a wider variety of grouping options than the teacher whose class was relatively homogeneous in ethnicity and achievement.

From these results it is not apparent on what basis teachers make the grouping choices they do. Some teachers are severely constrained by the class composition and in effect don't have a choice. It is unknown what decisions they would make with a more heterogeneous group of students. Other teachers appear to make a

conscious choice which may or may not depend on the particular class of students and/or the teachers' perceptions of these students.

When students chose groups, they were much more likely than teachers to form groups that were homogeneous on several characteristics. As described in previous sections students tended to form mostly same-gender groups and often formed groups that were homogeneous in ethnicity. Further analysis of these groups showed that students often formed groups that were homogeneous on both gender and ethnicity. In classes in which students chose groups, 37% of the groups were homogeneous on both gender and ethnic background. The types of groups students formed included: Latino female (9%), Latino male (7%), African-American female (4%), African-American male (10%), Anglo female (3%), and Anglo male (3%). When students chose their own groups, they often chose to work with students who were like themselves in terms of both ethnic background and gender.

Students were also more likely than teachers to choose groups that were homogeneous on both ethnic background and achievement (12% of groups) or that were homogeneous on both gender and achievement (16% of groups). Latino, African-American, and Anglo students all formed some groups that were homogeneous on achievement. Similarly, both girls and boys did so. Although these percentages are smaller than the percentage of groups that were homogeneous on ethnic background and gender (37%, discussed above), they still show some tendency among students to segregate themselves by achievement level as well as by gender and ethnic background. It was not the case, however, that students of any one achievement level (such as high-achieving students) were especially likely to separate themselves from other students. Of the groups that were homogeneous on both ethnic background and achievement, 50% were low-achieving, 17% were medium-achieving, and 33% were high-achieving. Of the groups that were homogeneous on both gender and achievement, 40% were low-achieving, 40% were medium-achieving, and 20% were high-achieving.

Stability of Group Compositions Across Visits

Classrooms were observed on three separate visits taking place at approximately one-month intervals. The previous sections presented the distributions of group size and group composition by student ethnic background, gender, and achievement level for the first visit. However, comparable data were compiled for all three visits and the results were compared across visits. To examine

variation across visits, repeated measures analyses of variance were carried out for group size and for each group composition variable with classroom as the unit of analysis. For example, the percentages of groups (by class) that were homogeneous on gender were compared across the three visits and the difference between means was not significant. For the 19 classes that had complete data from all three visits, the means (and standard deviations) for the percentage of groups that were homogeneous on gender were 29 (34) for visit 1, 35 (29) for visit 2, and 31 (26) for visit 3. This result showed that the percentage of groups that were homogeneous on gender remained similar from one visit to the next ($F(2, 36) = 0.29$, $p = .75$). Similar analyses were carried out for all of the group size and group composition variables. The probability level of every statistical test exceeded .05, showing that the distributions of group sizes and group compositions were not significantly different across the three visits to the classroom. Consequently, those results are not presented here.

Stability of Groups Over Time

The final issue examined in this paper is the extent to which students stayed in the exact same groups over time. Although the previous section showed that the group sizes and the compositions of the groups remained stable over time, the particular students who worked together were not necessarily the same. For example, a classroom may have had one group out of eight (12%) that was homogeneous on gender (e.g., four boys) on each occasion. However, the particular boys who appeared in that group may have changed from one visit to the next. This section examines the extent to which students who worked in each group stayed the same from one visit to the next.

The percentage of students who worked together in the same groups on consecutive visits was fairly high: 65% for Visits 1 and 2, and 63% for Visits 2 and 3, averaging over all classes. The percentage of students who were in the same groups in Visits 1 and 3 was lower, 37%. These results suggest that, on average, about a third of the students in the class changed groups from one visit to the next. Despite this apparent stability in group membership, there was great variation from classroom to classroom. The percentage of students who remained in the same groups from visit to visit ranged from 5% to 100%.

That is, in one classroom nearly all students changed groups from one visit to the next, whereas in another classroom all students remained in the same groups across all three visits.

Table 6 shows the percentage of students who stayed in the same groups across visits to the classroom by ethnic distribution of the class and by whether students or the teacher chose groups. Among the 19 classrooms with complete data from all three visits, the percentages of students who stayed in the same groups across visits were similar for majority ethnic and heterogeneous ethnic classes (Visit 1 to 2: $F(1,17) = 0.12$, $p = .73$; Visit 1 to 3: $F(1,17) = 0.47$, $p = .50$). Nor did results for stability differ significantly according to whether students chose groups or whether teachers assigned students to groups (Visit 1 to 2: $F(1,17) = 0.86$, $p=.37$; Visit 1 to 3: $F(1,17) = 0.25$, $p = .63$). Surprisingly, students did not always keep the same groups over time. It is possible that teachers told students to change groups from one science unit to another, rather than as the data implies, students chose to do so.

Table 6
Stability of Groups Over Time[a]

	Percentage of Students Who Stayed in the Same Group Over Time			
	Visit 1 to Visit 2		Visit 1 to Visit 3	
	M	*SD*	*M*	*SD*
Majority ethnic[b]	66	38	39	29
Heterogeneous ethnic[c]	59	41	29	28
Students chose groups	74	39	35	35
Teachers chose groups	56	39	27	19

[a] $n = 19$ classes with complete data from all three visits.
[b] Includes majority Latino, majority African-American, and majority Anglo classrooms.
[c] Includes heterogeneous Latino, African-American, Anglo; and heterogeneous Latino, African-American classrooms.

Because all classrooms worked on the same science unit (mystery powders, batteries and bulbs, crayfish, or daytime astronomy) on two of the three consecutive observations, analyses were performed to determine whether the stability of groups

was higher in classrooms that were studying the same unit on consecutive observations than in classrooms that changed units between observations. The results showed that stability of group membership did correspond to whether the same unit was used on consecutive visits, although the effect was more pronounced for some observations than for others. The percentage of students who were in the same groups in Visits 1 and 2 was slightly higher in classrooms using the same unit on both occasions (mean of 69% across 12 classrooms) than in classrooms studying different units on the two occasions (mean of 55% across 9 classrooms). The percentage of students who were in the same groups in Visits 2 and 3 was dramatically higher in the classrooms studying the same unit on both occasions (mean of 89% across 8 classrooms) than in the classrooms studying different science units (mean of 39% across 10 classrooms). Some teachers were excluded from these analyses because not enough information was available to determine which science unit they were studying on each occasion. All teachers changed units between Visits 1 and 3. In summary, although teachers were more likely to keep students together during a unit and change student groups when they changed instructional units, the actual group composition mix within and across classes was fairly constant across visits. The reasons for these changes and the corresponding impact on group performance and learning are unknown.

Conclusions

This study documented the kinds of groups that teachers formed for elementary school science instruction, examined the correspondence between teachers' stated grouping rules and classroom observations of actual groups, and determined the stability of the groups within and between science units. The results show considerable variability among teachers, both in terms of their reported grouping practices and in the group compositions that were observed in different classrooms. In short, there is no "one" way that teachers group students because classroom composition set limits on the nature and variety of groups that can be formed. Indeed, it was virtually impossible for some teachers to form groups consistent with recommended practice because of the homogeneous nature of their students with respect to ethnicity and achievement level. In the end, it appears that teachers' efforts to group heterogeneously were at odds with district and/or school policies that permitted the formation of homogeneous classrooms of students in a school district characterized by its diverse student population. Research should

consider the situations, if any, in which the homogeneity of group members facilitates learning, increases motivation, or enhances self-efficacy more than heterogeneous or other possible groupings. With empirical evidence to support decision-making, policy makers and teachers can align their practices in ways that maximally benefit the student populations they serve.

Generally, teachers' reported grouping practices with respect to gender agreed with the groups observed in their classrooms. For the most part, this was also true for teachers' reported grouping practices with respect to achievement level. Most teachers who used achievement level to group students stated that they formed mixed groups, and the majority of groups in these classes were, indeed, mixed. Because teachers judged students' achievement relative to other students in the class and not with respect to the larger population of 5th-grade students, groups were mixed relative to the range of achievement levels in the class; they were not necessarily mixed relative to the population in the school district. In classes of relatively low- achieving students as measured by CTBS, for example, teachers made distinctions among students who differed only minimally on an absolute basis, rating them low, medium, and high, and forming mixed groups accordingly. These teachers grouped students heterogeneously relative to the constrained range of achievement level in their classes, but the groups were homogeneous relative to the range of achievement level in the larger population.

Whether heterogeneity should be defined relative to the range of achievement in the classroom or relative to the range of achievement in the larger population is still an open question. Some previous studies have shown that relative standing in the group is a stronger predictor of a student's behavior in the group, and sometimes the student's learning, than is absolute ability (Dembo & McAuliffe, 1987; Webb, 1984b; Webb & Kenderski, 1984). Further, the same student may behave, and learn, differently depending on whether he or she is high, medium, or low relative to the other students in the group. Whether relative standing would function in the same way in classrooms and groups with constrained ranges of achievement level, especially uniformly low-achieving classrooms, has yet to be investigated. The impact (on group composition and learning) of absolute and relative achievement level in constrained classrooms needs to be studied systematically so that concrete strategies for defining heterogeneity with respect to student achievement level can be established.

Perhaps more important than the correspondence between stated and observed practices is the extent to which unstated practices were often seen to influence group composition. For example, no teacher reported considering ethnic background when forming groups. Nevertheless, classroom observations suggest that teachers did attempt to mix ethnicity when possible in their classrooms. In classrooms where students were predominantly of one ethnicity, opportunities to mix students were not present and students were grouped homogeneously.

As another example of the distinction between implicit and explicit grouping strategies, teachers stated that they put students who sat near each other (i. e., proximity) in groups together. What these teachers didn't say was that they had assigned students to seats to facilitate the kinds of groups they wanted to form. These classrooms, like other classrooms with teacher-assigned groups, had low percentages of homogeneous groups (12% homogeneous on ethnicity, 24% homogeneous on gender, 18% homogeneous on achievement level). Clearly, probing questions or extensive interviews are needed to understand the process by which teachers form groups. Having teachers articulate their decision-making process makes it available for questioning, debate, and, if necessary, revision, not only for this particular teacher but for others who are faced with a similar situation in their own classrooms. Indeed some teachers in this study were not aware of how students formed groups in their class until the visiting resource teacher pointed out the homogeneous nature of them. Here is a situation in which teacher guides and workshops can play an important role in effectively guiding teachers through the various options they might consider in forming groups in their classrooms. Of necessity is a strong empirical basis on which to suggest the relative merits of various strategies for particular groups of students or particular student outcomes.

In contrast to the implicit or explicit efforts of teachers to form heterogeneous groups for science instruction, students tended to group themselves homogeneously by gender, ethnicity, and to a lesser extent achievement. This finding is consistent with other empirical work indicating the tendencies for students to form homogeneous peer groups (Slavin, 1990). While there is no systematic research on the effects of the ethnic composition of small groups within the classroom on students' academic performance or social or emotional development, research on cooperative learning shows that working closely and collaboratively with others in mixed groups promotes cross-racial friendships and improves students' attitudes toward others. Slavin (1977, 1990) summarized the results of a number of studies

showing positive effects of cooperative reward structures on interracial friendships, interracial attitudes, interracial helping on school tasks, and reductions in interracial tension. While teachers may have felt allowing students to self-select their groups would be effective in motivating students or reinforcing their sense of self-efficacy through being in charge of their own learning, the groups that students ultimately formed may in the end have defeated or at least ignored other goals of grouping such as interracial relations and achievement. The literature is moot with respect to the relationship among the various goals of grouping (e.g., improving cross-race relations, developing teamwork and communication skills, increasing attitudes toward school and the subject matters learned, increasing learning) and the goals of science education (e.g., motivation, interest, epistemological beliefs, achievement) or other curricular reform efforts which promote grouping as an essential instructional practice. Are some groups more optimal for some goals than for others or does one group composition serve cognitive, social, and motivational goals equally well?

In summary, the mismatch between intended and actual grouping practices, the extent to which implicit rules versus explicit rules influenced group assignment, and the differences between teacher-assigned and student-formed groups merit further study. The main unanswered questions concern grouping according to student gender and ethnic background, the need, if any, to maintain stability of groups over time, and the extent to which grouping practices observed in this study are district, curriculum, or grade-level specific. Future research needs to explore in some detail the mechanisms by which groups were formed, the reasoning that teachers and students used to form particular groups, and the effects of varying group compositions on students' experiences in the classroom and on their learning, motivation, and social-emotional development. Particular attention should be given to when and why groups change and the impact of these changes on group performance. Positive group interaction does not just happen. Students need to be taught how to work effectively in a group and then given adequate time to practice and develop group-interaction skills (Webb & Farivar, 1994; Farivar & Webb, 1994). As such, it is necessary to understand more clearly the evolution of group dynamics over time and whether changing groups on some regular basis either between disciplines (reading and science) or within a discipline (science units) impacts this process in any way.

Finally, the generalizability of the findings reported here to other districts that are more or less diverse or to different grade levels, different curricula (e.g., reading,

mathematics), or some combination of these are open to challenge. Is it the case that students approaching adolescence seek homogeneity in their relationships both within and outside of school but younger students do not? Do teachers give primacy to different factors when grouping students for some subjects than others? What components of teachers' grouping strategies are context-specific and which aspects or decisions are independent of a particular context? With answers to these questions, it will be possible to give recommendations to teachers about the optimal group compositions to use in small-group instruction in their classrooms: groups that facilitate learning and social integration or other goals. Further, it will be possible to give recommendations to policy makers that support or challenge existing policies for assigning students to schools and/or classrooms. In summary, this study, carried out in one urban school district, raises issues that need to be investigated to help define and clarify optimal grouping strategies for maximizing achievement, supporting social and affective development, and facilitating positive cross-ethnic relations. Responding to these issues will help inform educational policy and classroom practice in specific rather than general ways.

References

Ames, G. J., & Murray, F. B. (1981). When two wrongs make a right: Promoting cognitive change by social conflict. *Developmental Psychology*, *18*, 894-897.

Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The Jigsaw Classroom.* Beverly Hills, CA: Sage.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59,* 87-96.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefit of teaching. *Journal of Educational Psychology, 72,* 593-604.

Bell, N., Grossen, M., & Perret-Clermont, A. N. (1985). Sociocognitive conflict and intellectual growth. In M. W. Berkowitz (Ed.), *Peer conflict and psychological growth* (pp. 41-54). San Francisco, CA: Jossey-Bass.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal, 21,* 755-765.

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 361-392). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Berger, J., Rosenholtz, S. J. , & Zelditch, M. (1980). Status organizing processes. *Annual Review of Sociology, 6,* 479-508.

Bossert, S. T. (1988). Cooperative activities in the classroom. *Review of Research in Education, 15*, 225-252.

Burns, M. (1981, September). Groups of four: Solving the management problem. *Learning*, 46-51.

California Department of Education. (1990). *Science Framework for California Public Schools, Kindergarten Through Grade Twelve*. Sacramento, CA: Author.

California Department of Education. (1985). *Mathematics Framework for California Public Schools, Kindergarten Through Grade Twelve.* Sacramento, CA: Author.

California Department of Education. (1992). *Mathematics Framework for California Public Schools, Kindergarten Through Grade Twelve.* Sacramento, CA: Author.

Cohen, E. (1986). *Designing groupwork: Strategies for the heterogeneous classroom.* New York: Teachers College Press.

Cohen, E. G. (1982). Expectation states and interracial interaction in school settings. *American Review of Sociology, 8*, 209-235.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*, 1-36.

Dembo, M. H., & McAuliffe, T. J. (1987). Effects of perceived ability and grade status on social interaction and influence in cooperative groups. *Journal of Educational Psychology, 79,* 415-423.

Dishon, D., & O'Leary, P. W. (1984). *A guidebook for cooperative learning*. Cooperation Unlimited, P. O. Box 68, Portage, MI 49081.

Doise, W. & Mugny, G. (1984). *The social development of the intellect.* Oxford: Pergamon Press.

Farivar, S., & Webb, N. M. (1994). Helping and getting help: Essential skills for effective group problem solving. *Arithmetic Teacher, 41*, 521-525.

Gerard, H. B., & Miller, N. (1975). *School desegregation: A long-range study.* New York: Plenum.

Glachan, M., & Light, P. H. (1982). Peer interaction and learning: Can two wrongs make a right? In G. Butterworth & P. Light (Eds.), *Social cognition: Studies of the development of understanding* (pp. 238-262). Chicago, IL: University of Chicago Press.

Grant, L. (1986, April). *Classroom Peer Relationships of Minority and Nonminority Students.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Hazelwood, C. C., Roth, K. J. Hasbach, C., Hoekwater, E., Ligett, C., Lindquist, B., Peasley, K., & Rosaen, C. (1992). *Gender and discourse: The unfolding "living text" of a science lesson.* Technical Report No. 60, The Center for the Learning and Teaching of Elementary Subjects, Michigan State University, East Lansing.

Hooper, S., & Hannafin, M. J. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research, 4,* 413-424.

Hooper, S., Ward, T. J., Hannafin, M. J., & Clark, H. T. (1989). The effects of aptitude composition on achievement during small group learning. *Journal of Computer-Based Instruction, 16*, 102-109.

Johnson, D. W., & Johnson, R. T. (1986). *Learning together and alone* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Johnson, D. W., Johnson, R. T., Holubec, E. J., & Roy, P. (1984). *Circles of learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Kagan, S. (1989). *Cooperative Learning Resources for Teachers*. San Juan Capistrano, CA: Resources for Teachers.

Kagan, S. (1992). *Cooperative Learning*. San Juan Capistrano, CA: Resources for Teachers (27128 Paseo Espada, Suite 622).

Linn, R. L., & Gronlund, N. E. (1995). *Measurement and Assessment in Teaching* (6th edition). Englewood Cliffs, NJ: Prentice Hall.

Lockheed, M. E., & Harris, A. M. (1984). Cross-sex collaborative learning in elementary classrooms. *American Educational Research Journal, 21,* 275-294.

Miller, N., & Harrington, H. J. (1990). A situational identity perspective on cultural diversity and teamwork in the classroom. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 39-76). New York, NY: Praeger.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8,* 181-192.

Murray, F. B. (1982). Teaching through social conflict. *Contemporary Educational Psychology, 7,* 257-271.

National Center for Improving Science Education. (1989). *Science and Technology Education for the Elementary Years: Frameworks for Curriculum.* Washington, DC: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Research Council (Mathematical Sciences Education Board). (1989). *Everybody counts: A report to the nation of the future of mathematics education.* Washington, DC: National Academy Press.

National Research Council. (1995). *National Science Education Standards*. Washington, DC: National Academy Press.

Oakes, J. (1990). *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science* (Report R-3928-NSF). Santa Monica, CA: The Rand Corporation.

Oosterhof, A. (1994). *Classroom applications of educational measurement* (2nd. edition). New York: Macmillan Publishing Company.

Perret-Clermont, A-N. (1980). *Social interaction and cognitive development in children.* (European Monographs in Social Psychology, 19, Series Editor: H. Tajfel). New York: Academic Press.

Peterson, P. L., Wilkinson, L. C., Spinelli, F., & Swing, S. R. (1984). Merging the process-product and the sociolinguistic paradigms: Research on small-group processes (pp. 125-152). In P. L. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes*. Orlando, FL: Academic Press, Inc.

Sharan, S., Kussell, P., Hertz-Lazarowitz, R., Bejarano, Y., Raviv, S., & Sharan, Y. (1985). Cooperative learning effects on ethnic relations and achievement in Israeli junior-high-school classrooms. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 313-344). New York: Plenum Press.

Sharan, Y., & Sharan, S. (1992). *Expanding cooperative learning through group investigation*. New York: Teachers College Press.

Skon, L., Johnson, D. W., & Johnson, R. T. (1981). Cooperative peer interaction versus individual competition and individualistic efforts: Effects on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology, 73,* 83-92.

Slavin, R. E. (1977). Classroom reward structure: An analytic and practical review. *Review of Educational Research, 47,* 633-650.

Slavin, R. E. (1983). *Cooperative Learning.* New York: Longman.

Slavin, R. E. (1990*). Cooperative Learning: Theory, Research, and Practice.* Englewood Cliffs, NJ: Prentice-Hall

Tudge, J. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19,* 123-138.

Webb, N. M. (1980). A process-outcome analysis of learning in group and individual settings. *Educational Psychologist, 15,* 69-83.

Webb, N. M. (1984a). Sex differences in interaction and achievement in cooperative small groups. *Journal of Educational Psychology, 76,* 33-34.

Webb, N. M. (1984b). Stability of small group interaction and achievement over time. *Journal of Educational Psychology, 76,* 211-224.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13,* 21-40.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22,* 366-389.

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31,* 369-395.

Webb, N. M., & Kenderski, C. M. (1984). Student interaction and learning in small group and whole class settings. In P. L. Peterson , L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes* (pp. 153-170). New York, NY: Academic Press.

Webb, N. M., & Kenderski, C. M. (1985). Gender differences in small group interaction and achievement in high-achieving and low-achieving classrooms. In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 209-226). Orlando, FL: Academic Press.

Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22,* 390-408.

Author Notes

This study was supported in part by grants from the National Science Foundation (Grant Number ES1-90-55443); the Academic Senate on Research, Los Angeles Division, University of California; and the Office of Educational Research and Improvement of the U. S. Department of Education Grant to the Center for Research on Evaluation, Standards, and Student Testing (Grant Number R117G10027). However, all opinions in this paper are those of the authors.

We would like to thank Jeff Shih for helping to create the computer data files and for conducting preliminary analyses, and Becky Smerdon for her help with the data analysis. We are especially grateful to the teachers who participated in this project. Requests for reprints should be sent to Noreen M. Webb, Graduate School of Education and Information Sciences, UCLA, Los Angeles, California 90095.