Tracking Middle School Students

for Instruction:

A Study of Homogeneous and Heterogeneous Grouping

A Thesis Presented to

the Faculty of

California State University, San Marcos

by

Jonathan A. Flores

Spring 1999

Approved by:

Janet E. McDaniel, Ph.D.  
5/10/99  
Date
TRACKING MIDDLE SCHOOL STUDENTS FOR INSTRUCTION:
A STUDY OF HOMOGENEOUS AND HETEROGENEOUS GROUPING

A Paper

Presented to the Faculty of

California State University San Marcos

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in Education

by

Jonathan Andrew Flores

Spring 1999
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. Introduction and Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>II. Review of the Literature</td>
<td>5</td>
</tr>
<tr>
<td>III. Method</td>
<td>28</td>
</tr>
<tr>
<td>IV. Findings</td>
<td>36</td>
</tr>
<tr>
<td>V. Discussion</td>
<td>45</td>
</tr>
<tr>
<td>VI. Conclusion and Recommendations</td>
<td>51</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td></td>
</tr>
</tbody>
</table>
Abstract

Tracking Middle School Children for Instruction:
A Study of Homogeneous and Heterogeneous Grouping

by Jonathan A. Flores

This study examined the effects of tracking on the academic
achievement and attitudes of 56 seventh grade students. The study took
place in a public middle school of 600 students located in a middle class
neighborhood of a large suburban California school district. The
socioeconomic makeup of the school was predominantly middle-class,
with White students comprising about 75% of the enrollment.

The 56 students were divided into two homogeneously grouped
classes for instruction during the first quarter, then regrouped into two
heterogeneous classes for instruction during the second quarter. Both
classes were taught the same curriculum and received all instruction from
the same two classroom teachers. Academic achievement of the two
groups were compared using the students' grade point averages and their
mean scores on teacher-created math tests. Student attitudes regarding
the classroom learning environments were compared using written
reflections collected from the students after each quarter of instruction.
T-tests for statistical comparisons of the students’ GPAs and math test scores showed no significant differences. Thus, the results of the study showed negligible changes in academic achievement when tracking students for instruction. However, a chi-square analysis of the written reflections showed a significant difference in the students’ positive attitudes towards their classes, which favored the heterogeneous groups. The teacher observed that the students, overall, exhibited better cooperation and behavior when grouped heterogeneously.

Recommendations for further studies call for research which examines the effects of two treatments on heterogeneously grouped students. First, improved training for teaching classrooms with all types of learners. Second, the provision of high-quality teaching materials and programs of enrichment for classroom use. Finally, there is a general recommendation for increased administrative support for heterogeneous grouping in our schools.
CHAPTER 1

INTRODUCTION

The reform of public education is an issue of critical debate in the United States. Public schools are the fundamental base of American democracy. They are a reflection of their communities, and to some extent the nation as a whole. Therefore, their success or failure may be the crux upon which our entire society succeeds or is relegated to the pages of history. Among the myriad of issues surrounding our schools and education in general is the controversial practice called “tracking”, which divides students into separate classes, groups, or sections based upon some measure of achievement or ability. Public schools have long employed this method of sorting students and it continues to be pervasive in U.S. schools today.

Tracking is an issue that merits much scrutiny because it raises many concerns in regards to equality, class privilege, social justice, and sound educational practice. Our public schools are faced with the incredible challenge of educating a very diverse population of students which reflect the ever-increasing multicultural makeup of modern-day America. These students come from many different ethnic and economic backgrounds, and bring many different learning needs into the classroom. The question of how to meet the needs of such a diverse population
necessitates the examination of tracking students and to what degree it has its place in American education.

**Background**

In my five years of teaching middle school math and science, I have always taught classes made up of students of varied ability. A typical class usually has a few high-achieving students, many average-achievers, and a few low-achievers. Not surprisingly, I have observed that the high-achievers often complete the lessons and assignments more quickly than the other students, while the low-achievers struggle to keep up with the pace of the class. The high-achievers seem capable of learning more than I could plan for the entire class, while the low-achievers need more time and support to grasp the presented material.

I have attempted to remedy the situation by providing homework that is more challenging for the high-achievers, with the intent being to provide them with a deeper understanding of the topic of study. I have also offered extra credit research assignments with the same intent, which the high-achievers usually complete for the grade. In retrospect, this really amounts to assigning them more work rather than teaching them more concepts.

Low-achieving students who are identified as learning disabled can receive extra help from a resource teacher assigned to their case. This is helpful, but it does not rectify their ability to keep up with the pace of my
class, which must press forward in order to cover the required grade level curriculum. Other low-achievers do not qualify for special education, and they usually do not take advantage of help that is available. My classroom is open during lunch for extra help, reviewing of tests, and for completing homework. The school also has a homework club, an after-school program which offers tutoring and a quiet place to work on assignments.

Lost in this entire discussion is the average student. Those students who plug along getting C's and B's and do not stick out are easy to overlook. My immediate concerns focus primarily on those students who are failing, then on those high-achievers who need a little extra push. The urgency to address the needs of those students who are doing satisfactory work and not causing trouble are not as pressing as the other students. Thus, I find myself not having the time to meet the instructional needs of all my students.

Statement of the Problem

As a classroom teacher, I am deeply concerned with the academic achievement of all my students, and the associated implications for my instructional practice. From the many topics in educational research that deal with the issue of having students of varied abilities in the same classroom, I chose to study the regrouping students by perceived ability, also known as “tracking”. It is thought by grouping the students into classes of similar ability levels, the problems of pacing and delivery of
instruction should be greatly reduced, thus making the teacher's job easier. However, the heart of the issue remains the students' attainment of both the academic and social goals set for them. The purpose of this study was to examine the effect of tracking on the classroom learning environment and the academic achievement of my seventh grade students over a limited period of time.
CHAPTER 2
REVIEW OF THE LITERATURE

In this review of the literature, I will first define key terms relevant to the discussion. Next, I will examine tracking practices in the U.S. and the rationale supporting them. Then, I will present the criticism of tracking as an educational practice based on the issues of equality and negative social consequences. Finally, I will examine the research conducted in this area, and the efforts to reform the educational practice of grouping students, known as “detracking”.

Definition of Terms

The following list of terms will be mentioned throughout the literature review and are key to the understanding and clarification of the discussions presented.

Tracking—The practice of dividing students into separate classes, groups, sections or schools for high, average, and low achievers.

Ability Grouping—A school selection system that groups students on the basis of perceived ability and differentiates instruction by quantity and intensity of work. Typically, this system segregates students all or most of the day and bases placement on general ability or achievement rather than skill in a specific subject. Examples are special education and gifted and talented programs.
Heterogeneous Grouping—The absence of a structured grouping plan based on students’ perceived ability. Students are assigned to a classroom irrespective of such factors as intelligence, social maturity or achievement.

Cooperative Learning—Instructional methods in which students work in small learning groups toward a group goal. Cooperative learning differs from within-class ability grouping which are usually heterogeneous and they engage in a task-focused activity, such as studying together or completing a group assignment.

U.S. Tracking Practices

Since the 1920's a majority of schools have offered a “tracked” curriculum, a sequence of academic classes that range from slow-paced remedial courses to rigorous academic ones, in which students are sorted into homogeneous groups according to their past academic achievement and presumed educational needs (Oakes & Lipton 1999). Most schools group students homogeneously, not only by age but also according to academic ability, educational disadvantage, learning and behavioral disabilities, language proficiency, educational aspirations, and college potential (Smith-Maddock & Wheelock 1995). Nearly all middle and senior high schools ability group some or all academic subjects based on students’ past grades, test scores, and teacher recommendations. About twenty states now provide special funding for gifted and talented students,
and nearly all schools have highly visible gifted programs where the highest-achieving students are grouped together for enrichment or accelerated instruction, either in separate classes, in pull-out programs which remove students from the regular classroom for special instruction, or in gifted clusters in which students are grouped for instruction according to ability and/or achievement within regular classrooms (Oakes & Lipton 1999).

At the high school level, tracking is most prevalent in mathematics, where students are placed in vocational, general, or college-preparatory mathematics courses. It also occurs at the middle or junior high school level in those schools that offer algebra at the eighth grade (Slavin 1990). Because some math courses follow a sequence, students’ assignments in earlier grades determine how far they can progress by the time they graduate. The percentage of young adolescents grouped homogeneously for math increases every year, growing from 57 percent in fifth grade to 94 percent in ninth grade. By seventh grade, students are more likely to be grouped by whole class in math than for any other subject. Typically, students who will be in the top math classes all through high school are identified by the sixth grade or before (Wheelock 1992).

In addition to grouping students by their academic ability or prior achievement in various subjects, high schools also proscribe different sequences of classes, or tracks, for students with different futures. Most
senior high schools offer some advanced placement (AP) classes for exceptionally high-achieving eleventh and twelfth graders bound for the most competitive universities (Oakes & Lipton 1999).

Not all ability grouping results in distinct and separate class assignments. Students are often homogeneously grouped into small groups within a classroom, where clusters are based on ability or achievement within that particular classroom. This has been an established practice for reading instruction at the elementary school level for years (Oakes, 1990; Slavin 1987). Teachers frequently will organize their classrooms in a similar format for mathematics instruction.

In summary, there is a pattern of sorting students in American schools according to perceived measures of achievement, ability, behavior, and potential. This practice is prevalent in our schools and often begins at an early age and increases as students approach graduation from high school.

The Rationale for Tracking

The persistence of tracking in the public school system is based on several assumptions: (1) students learn better when they are grouped with other students who are considered to be academically similar; (2) students develop more positive attitudes towards themselves and school when they are not placed in groups with others who are more capable; (3) placement processes used to separate students into grouping both fairly and
accurately reflect past achievement and native abilities; and (4) it is perceived as easier for teachers to accommodate individual differences in homogeneous groups (Oakes 1985).

Closely tied to the practice of tracking is the central question of how to bring sound instruction to gifted students in regular classroom settings. This is a question many teachers struggle to answer as they attempt to create a meaningful learning environment for students with wide-ranging learning styles and abilities. Teachers identify three primary characteristics of gifted and talented students. First, they tend to get their work done quickly and may seek further assignments or direction. Second, they ask probing questions which tend to differ from their peers in depth of understanding. Finally, gifted learners may have interest in areas more like the interest of older students, and they often know more about the subject being taught than their instructors (Tomlinson 1995).

These traits lead many educators to conclude the best solution for meeting the needs of gifted learners is to separate them into classroom settings where their unique talents can be stimulated and challenged. Many schools have Gifted And Talented Programs (GATE) which exist to serve this purpose. Gifted education exists to foster high-end excellence. It stresses practices most likely to promote "expertise" in learners with advanced potential. Educators of the gifted value the benefits of ability grouping for advanced learners. They are concerned about a lack of
emphasis on differentiated instruction for academic diversity in heterogeneous classrooms, and reject a one-size-fits-all approach to educating students. They find the overuse of some cooperative learning strategies result in a lack of challenge for advanced learners, as well as the inappropriate use of these learners as "junior teachers" (Tomlinson 1995). Ability grouping is seen as a way to narrow the range of performance and motivation in a group of students, thereby making teaching easier and preventing less able students from "holding back" those with greater academic talent (Brewer, Rees & Argys 1995).

Hence, many educators are concerned the efforts to raise the achievement of students who traditionally have been classified as having average or below-average abilities are at the expense of the academically gifted. Friedler, et al., in a paper published in 1993, argue our school systems are actually giving tacit approval to creating underachievement in one ability group so that the needs of the other ability groups can be served (Thornton 1995). Thus, the conventional wisdom for homogeneous grouping is that once the school identifies educationally relevant differences, teachers can teach groups of students with meaningful similarities, and students will benefit from instruction in these groups (Oakes & Lipton 1999).
Criticisms of Tracking

Criticism of tracking usually falls into one of two categories, inequality or social-psychological theory. I will first examine the issues of inequality many claim to be inherent in tracking practices, then I will look at effects of tracking on students’ attitudes and behavior.

The pressure to eliminate tracking emanates from many studies that have shown the practice has negative consequences for the future educational opportunities and schooling outcomes of many children. These negative consequences disproportionately effect low-income, African-American, and Latino children. Ability grouping and the differential distribution of expectations that accompanies the practice are key elements in locking students out of meaningful opportunities for future success (Smith-Maddock & Wheelock 1995). Critics suggest ability grouping often limits the instructional experience of lower-track students to little more than rote drill on basic skills. Further, because mobility between tracks is rare, students placed in low tracks at a young age may never be transferred to the upper tracks where higher-order skills are typically taught (Burnett 1995).

Brewer, Rees, and Argys (1995) examined the effect of 10th grade tracking on end-of-year 10th grade achievement in mathematics for a sample of more than 3,900 public school students. Their data confirmed the proposition that high-track classes received more educational
resources than lower-track classes. For example, students in above-average math courses tended to be taught by more experienced teachers than students in either below-average or heterogeneous classes and were more likely to be taught by a teacher with a master’s degree. Thus, their evidence suggests lower-track classes tend to be assigned to the least qualified teachers and, in general, tend to receive less than their share of educational resources. Similarly, upper-level classes seem to receive a disproportionately large share of resources.

In schools that group by ability, the best students are more likely to have access to teachers who specialize in a subject, to newer materials and facilities, and perhaps most importantly, to higher expectations from teachers and a more challenging curriculum. “All the things we think matter in terms of a high-quality education, we disproportionately give to high-achievers,” says Jeannie Oakes, a professor of education at UCLA (Ratnesar)

A 1990 NEA study examined the relationship between tracking and equality of opportunity. Below are the study’s findings:

- Tracking is very common and becomes progressively more rigid with each subsequent school grade.
- Lower socioeconomic, Black and Hispanic students tend to be over-represented in lesser academic classes, and female students are still underrepresented in higher science and mathematics classes.
• Early grouping and tracking deprive students of access to educational classes that qualify them for advanced placement.

• Educationally effective grouping procedures keep students in heterogeneous classes except for a particular subject, are constantly reevaluated and accommodate all achievement levels (Minnesota Education 1998).

According to the data compiled by the National Educational Longitudinal Study (NELS:88), African-American, Latino, Native American, and low-income eighth graders are twice as likely as white or upper-income eighth graders to be in remedial math courses. Not only do students in remedial settings receive less demanding curriculum, but their teachers are also more likely to be less experienced in the classroom. For example, researcher Lorraine McDonnell (1990) and her colleagues found that teachers in 42 percent of the remedial, vocational, and general-mathematics sections have been teaching for five years or less, compared with 19 percent of those in the pre-algebra and algebra-1 sections.

Defenders of homogeneous grouping practices claim group assignments are objective or color blind, and they attribute the disproportionate assignments into college-prep or learning-disabled classes to unfortunate differences in students' backgrounds and abilities. They claim the true policy question is one that goes to the heart of the
excellence versus equity arguments. Equity does not consist of giving the same material to all students, regardless of aptitude or past performance, rather, it means giving each student what he or she needs, and that would require differentiating the curriculum to meet the needs of gifted students (Gallager 1995).

Many of the reform efforts which seek to address the issue of equity has been focused at the middle level. Middle school reforms view education through an equity lens, where all students have an equal opportunity to succeed. Middle school educators emphasize the negative impact of homogeneous grouping on at-risk learners. They argue the tenets of gifted education (e.g., enrichment, high-level thinking, problem solving) are good education for all learners, and should not be reserved for any single group of students (Tomlinson 1995).

Heterogeneous grouping strategies support education as an instrument of social justice and have become a hallmark of good middle schools. However, because homogeneous grouping practices are so deeply entrenched in schools and because these practices closely connect to fundamental cultural values and political interests, changing them proves extraordinarily challenging (Oakes & Lipton 1999).

In summary, much of the criticism of tracking point to the disproportional allocation of resources and opportunities denied to students of color. In their view, in order for the promise of equal
educational opportunity to be realized certain steps must be taken to eliminate practices that divide students into categories of “more able” or “less able” learners and instead provide equal access to knowledge.

Social Criticism of Tracking

Another criticism of tracking is based on a social-psychological theory of human behavior. Proponents of this theory contend an individual's ability to learn depends on the expectations of “significant others” and follows several steps:

1. The social norms and expectations of others define appropriate behavior for persons in various social situations.

2. Each person learns definitions of appropriate behavior through interaction with others who are important and significant to him or her.

3. The individual learns to behave in ways she or he perceives are appropriate and proper for her or him.

4. The individual also acquires conceptions of his/her ability to learn various types of behavior through interaction with others whose evaluations are important to him or her (Brookover, W.B., and E.L. Erickson, 1969, Society, Schools and Learning. Boston: Allyn and Bacon).

Jeannie Oakes examined effects of tracking on several dimensions of classroom practices, including student attitudes toward school and toward themselves. She found that students in higher-track classes had significantly more positive attitudes about themselves and significantly
higher educational aspirations than students in lower-track classes. Low-track students were more likely than other students to view themselves as not as well-liked by other people and as having many things about themselves they would like to change. These feelings, in turn, foster significant differences in classroom climate, and more people are questioning the fairness and morality of the practice (Smith-Maddock & Wheelock 1995).

In short, there is evidence that shows when low-achieving students are isolated from high-achieving peers, they have little opportunity to acquire the high achievement norms and abilities required for future success.

Research on Tracking

Studies examining the effects of homogeneous grouping on achievement tend to take two approaches: (1) comparisons of the achievement of students in heterogeneous classes with comparable students in ability-grouped classes or (2) comparisons of the achievement of students in different ability groups (Davenport 1993). Studies of the grouping practices which follow the first approach have been found to have little or no effect on student achievement. Studies which follow the second approach have shown to make a difference in student outcomes, the benefits of which are the topic of much debate.
Programs that entail only minor adjustment of course content for ability groups usually have little or no effect on student achievement. An example of such a program is one in which school administrators assign students by test scores and school records into high, middle, and low classes, in which they expect all groups to follow the same basic curriculum. The traditional name for this approach is XYZ grouping. Pupils in middle and lower classes in XYZ programs learn the same amount as equivalent pupils do in mixed classes. Students in the top classes in XYZ programs outperform equivalent pupils from mixed classes by about one month on a grade-equivalent scale. In XYZ programs self-esteem of lower aptitude students rises slightly and the self-esteem of higher aptitude students drops slightly (Kulik 1992). Slavin refers to these kind of grouping studies as "track/no-track" studies. He concurs that studies of this kind find the achievement effects of ability grouping are near zero for high, average, and low achievers (Slavin 1995).

A meta-analysis of studies examining the effects of ability grouping on achievement of secondary students (middle, junior high, high school) reported that in comparisons of ability grouping and heterogeneous grouping over periods of from one semester to five years, overall achievement effects were found to be essentially zero at all grade levels (Slavin 1990). There have been at least eight meta-analyses investigating the effects of ability grouping. Although the details of these analyses have
been subject to much critical comment, the overall effect size from the eight meta-analyses is 0.18 (standard error=0.03). Thus, on average, the achievement effect of ability grouping is quite small, equaling less than one-fifth of a standard deviation in student achievement. Since the effect is very small, it is not surprising many studies report positive effects of ability grouping and many other studies report negative effects (Jaeger & Hattie 1995).

Jeannie Oakes, in her report *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and sciences.* (1990), concurs with these findings as well. She writes:

"While most people (including many educators) assume that students will learn better if they are grouped together with those who have similar capabilities, research has shown that putting children into separate classes to accommodate their differences from earliest school years is neither necessary nor very effective. Tracking does not work well for students in the low and middle ability groups, who experience clear and consistent learning disadvantages. Tracking does not necessarily promote achievement for high-ability children either. Many studies show highly capable students do as well in mixed ability classes." (pg.6)

Grouping strategies following the second approach, in which students are grouped according into high, medium, and low tracks based
on ability, have shown to benefit some students. The key element of these programs is they entail more substantial adjustment of curriculum to ability, and appear to have clear positive effects on some children. Cross-grade and within-class programs provide both grouping and curricular adjustment in reading and math for elementary students. Pupils in such programs outperform equivalent control students from mixed-ability classes by two or three months on a grade-equivalent scale (Kulik 1992).

Programs of enrichment and acceleration, which usually involve the greatest amount of curricular adjustment, have the largest effects on student learning. In typical evaluation studies, talented students from accelerated classes outperform non-accelerates of the same age and IQ by almost one full year on achievement tests. Talented students from enriched classes outperform initially equivalent students from conventional classes by 4 to 5 months on grade equivalent scales (Kulik 1992). In light of these findings it appears highly talented children profit greatly from work in accelerated classes, and from programs of enrichment.

Criticism of literature reviews by Slavin and Oakes focus on the results of these kind of studies. The critics point out a balanced review cannot eliminate analyses of studies that deal with gifted students without limiting the conclusions that can be drawn from such a review.

"When a synthesis of data has been conducted that focuses on gifted students and differentiated content, the results are quite different
from those found by Slavin and Oakes. Grouping gifted students for special instruction appears to yield very positive results..." (Gallager 1995, pg.216).

Slavin refers to these second kind of studies as "high track/low track" studies, in which students are assigned to high, middle, and low tracks or ability groups, controlling for pretest scores, I.Q., and other indicators of initial ability or performance. He admits these studies almost always find students gain from being in high-ability groups. Slavin writes:

"The problem with the high track/low track studies is they compare children with substantially different pretest or I.Q. scores and then attempt to control for those differences. The statistical controls just don't work when the groups being compared are far apart on covariates that are strongly related to outcomes...The track/no track studies are far more meaningful. In these, experimenters typically know for sure that grouped and ungrouped students were equivalent at the pretest, and they knew for sure what grouping plan was actually implemented. Among such studies, it is difficult to find even a single instance of important gains for students in the high-ability groups." (Slavin 1995, pg.220).

Brewer, Rees & Argys, who favor certain ability grouping practices, admit the following:

"Since assignments to tracks is made at least in part on the basis
of prior ability, any real attempt to measure the effect of tracking itself must disentangle the influence of tracking from the process of assignment. It is likely that factors that researchers cannot observe (e.g. student motivation that are correlated with student performance) also help determine the track to which a student is assigned." (pg. 210)

The allocation of resources is a factor in which the research on ability grouping practices may be flawed. As previously mentioned, there is strong evidence that those students assigned to the higher ability groups receive a disproportionate share of the resources available to schools. Therefore, it should not be surprising that these students should show an improvement in achievement. Following this same line of reasoning, it is also not surprising that students placed in lower ability groups should show a decrease in achievement. If this pattern of allocation indeed exists, then it is obviously important to control for it in any effort to examine the effect of ability grouping on learning.

Perhaps the most important aspect of the unequal distribution of resources is the classroom teacher. More than any other factor in which schools have some control over, it is the interaction between the teacher and students which has the greatest impact on student achievement.

"The effects of ability grouping are trivial when compared to those
of intra-classroom innovations controlled by teachers. Teachers, their training and background, and the choices they make have a far greater influence than do any innovations having to do with class organization, including ability grouping.” (Jaeger & Hattie 1995, pg.218).

**Detracking Efforts**

The question at the core of the tracking debate is how best to educate large numbers of students whose backgrounds and abilities differ widely. Opponents of the practice argue tracking stigmatizes students who are consigned to low-track classes with less experienced teachers, fewer resources, and lower expectations. They maintain the average and above-average students do not derive substantial academic benefits from being grouped together (Brewer, Rees & Argys 1995). Untracking is desirable because it abandons a strategy which sorts students according to individual weakness, and favors a heterogeneous approach, one which groups them in terms of collective strength. In light of the information age in which we live, our democratic society cannot survive with a continued unequal distribution of knowledge (Wheelock 1992).

Those who favor detracking point out the weaknesses of homogeneous grouping:

"Many educators believe homogeneity is necessary for schools to work well. And yet, this powerful organizing idea and structure is
deeply flawed. Beyond the huge problems inherent in the socially constructed and reified nature of students' classifications, other powerful factors make group assignments no more than rough approximations based on guesswork or bias. These factors include the fallibility of tests, the subjectivity of placement criteria and procedures, and parents' and students' own activism. Moreover, resources, scheduling problems, and other organizational constraints often limit schools' placement options (Oakes & Lipton 1999, pg.294).

Some state-policy documents have served to build momentum for the detracking movement. In Massachusetts, the Board of Education's 1990 policy advisory, *Structuring Schools for Student Success; A Focus on Ability Grouping*, has given a push to untracking and has set the stage for local school reform in its recommendation that "schools consider eliminating the use of widespread ability grouping and tracking practices, and implement alternatives." (Wheelock 1992). In California, *Caught in the Middle*, a report published by the California Department of Education, included recommendations for substantial reduction of tracking and the advice that no student should be tracked in grades six, seven, or eight according to ethnicity, gender, general ability, primary language, or handicap (Wheelock 1992).
Some schools have responded to the call for reform and replaced tracking with heterogeneous grouped classrooms, in which cooperative learning is an integral part of the teaching. Others have abandoned tracking for heterogeneous classes in which students are divided into two or three small ability groups for math or reading instruction (Slavin, 1993). These small groups may or may not use cooperative learning techniques. Although this method sorts students, it may provide advantages over large-scale tracking practices. The small group size is more fluid, so it is more likely students will be able to move into higher tracks as their achievement improves. In addition, using small groups may make it possible to tailor curricula and teaching methods more closely to the needs of individual students.

In order to continue the reform momentum for detracking, it will be necessary to gather research evaluating the effects of detracking on student performance. Not surprisingly, this can be problematic. Student performance is related to a host of observable (and many unobservable and unmeasurable) individual, family, teacher, school, and community characteristics. If statistical models ignore these factors, then biased results are likely (Brewer, Rees & Argys 1995). To resolve the unanswered questions about detracking, Slavin (1993) recommends several directions for such research, including studies of:
• the use of cooperative learning, within-class ability grouping, and other models, specifically in schools that are detracking;
• the effectiveness of other teaching methods, such as mastery learning and mixed-age groupings in heterogeneous settings;
• methods such as individualized studies and supplementary tutoring for helping low-achieving students succeed in high-quality, challenging heterogeneous settings;
• typical characteristics of both successful and unsuccessful tracking efforts.

The conventional wisdom seems to have evolved into a belief that tracking is universally bad for low-ability students and neutral for other students. If this is true, the policy prescription is very clear: detrack. Everyone will gain and no one will lose. However, Brewer, Rees & Argys (1995) found that placement in a below-average math class, as compared to a heterogeneous one, was associated with a decrease in achievement of approximately five percentage points. Placement in an above-average math class was associated with an achievement increase of roughly the same magnitude; placement in an average class was associated with an increase of somewhat less than two percentage points. These results suggest that detracking would create winners and losers. Although students in lower tracks would realize achievement gains by being placed
in a heterogeneous class, this would be at the expense of students placed in higher-level tracks (Brewer, Rees & Argys 1995).

Summary

There is a pattern of sorting students in schools, often called "tracking", which usually begins at the elementary level and becomes more prevalent at the secondary level. The assumption by supporters of tracking practices believe students are easier to teach and they learn better when they are grouped with others of similar ability. Those opposed to tracking cite the patterns of inequality they associate with such practices. These patterns include a disproportionate share of educational resources allocated for the higher tracks, and a large percentage of students of color and poverty assigned to lower tracks. Other critics of tracking point to studies which show negative affects on students' self-esteem.

Research on the effects of tracking are mixed. Some studies show no significant difference in academic achievement when students of similar ability are grouped homogeneously and heterogeneously. Other studies show homogenous grouping does produce significant differences in achievement of gifted students when the curriculum is modified, or differentiated to their ability level.

Reform efforts to eliminate tracking from schools, known as detracking, attempt to address the inequality in schools and the uneven
distribution of educational resources. These reforms call for heterogeneous classrooms, cooperative learning strategies, high expectations and a challenging curriculum for all students. Despite these efforts, many U.S. schools continue to utilize some form of tracking, or ability grouping.
CHAPTER 3

METHOD

The participants for this study are 56 seventh grade students from a public middle school. The students ages range from 11-13 years; 34 are male and 22 are female. The school is located in a middle class neighborhood of a large suburban school district serving approximately 24,000 students. It is the newest and smallest of the district’s four middle schools, housing 600 students in portable classrooms. The socioeconomic makeup of the school is predominantly White middle-class, with White students comprising about 75% of the enrollment, Hispanics about 10%, African-Americans about 6%, and other minorities about 9%. The parents are very active in the school, primarily through the PTA and volunteering in classrooms. The school and the district are very responsive to the standards and accountability movements in education. The district has adopted 8th grade exit standards which students must pass before they can continue on to high school, and is currently in the process of realigning the curriculum with the newly created state standards for mathematics and science.

The school faculty is organized into grade level teaching teams of two teachers, with each team sharing 68 students in class sizes of approximately 34 per teacher. The two teachers team to instruct what the school defines as the core subjects of math, science, language arts, and
social studies. They also share the responsibility of providing instruction in exploration and foreign language. The only subjects they do not teach are physical education and school band. Thus, each teaching team has a student contact of 68 and several preparations for instruction. Because of the teaming approach embraced by the school, the administration allows the teachers to group their students as they deem necessary. Most teams prefer to randomly group their students for instruction, thus creating heterogeneous classes with students of wide-ranging abilities and needs. Some prefer to track their students by ability, creating high and low groups. This is particularly true in eighth grade math, where students are tested and placed in either Beginning Algebra or Integrated Math.

For this study, I was able to persuade my teaching partner to allow me to group our students according to perceived ability for the first quarter of instruction. We have always randomly placed our students into heterogeneous groups, and this was the first time we have utilized this practice in our four years together at the school.

Prior to organizing our incoming seventh graders into homogeneous groups, I consulted with their sixth grade teachers. The sixth grade teachers rated each student's academic ability and performance as either high, medium, or low. I used these recommendations as the primary factor for student placement. All students who received high ratings were placed into the high ability group
and all who received a low rating were placed into the low ability group. For the students who received a medium rating, I looked at their SAT-9 scores from sixth grade. Each student had percentage scores in the four areas tested: Total Reading, Total Math, Total Language, and Spelling. I added these together to arrive at an overall point total for each student. Those receiving the highest point totals filled the remaining spots in the high ability group, and the rest filled the spots in the low ability group. Since some students were high in one area, such as math, and low in other areas, such as reading, there were some students who were misplaced in one or more subject areas. Due to scheduling constraints, this was unavoidable.

Thus, the students were homogeneously grouped into high and low classes for the core classes during the first quarter. For the second quarter of instruction, the students were randomly placed into heterogeneous groups of mixed ability. Again, all students received the same basic curriculum and were held to the same standards for grading and behavior. The intent was to try both grouping strategies to see which was more successful.

Table A shows the daily schedule for the students for the first and second quarters of instruction, which is the time period this study took place. Each quarter is two months long. I taught the math/science component of the core classes during periods 1-2 and 6-7. My partner
taught the language arts/social studies component during these same periods. Together, we shared the responsibility for teaching the explore and foreign language classes during periods 3 and 5, respectively. The students were homogeneously grouped for the core classes only.

The data used in this study were limited to those students who were present for both quarters of instruction and did not have schedule changes which prevented them from full participation in the grouping configurations. Students in both the high and the low groups received the same basic curriculum. Both groups were held to the same

<table>
<thead>
<tr>
<th>TABLE A</th>
<th>DAILY SCHEDULE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Quarter 6-7</strong></td>
<td><strong>Periods 1-2</strong></td>
</tr>
<tr>
<td>Math/Sci. teacher</td>
<td>high core</td>
</tr>
<tr>
<td>L.Arts/S.S. core Teacher</td>
<td>low core</td>
</tr>
</tbody>
</table>

| **Second Quarter 6-7** | **Periods 1-2** | **Period 3** | **Period 4** | **Period 5** | **Periods** |
| Math/Sci. core Teacher | mixed core | foreign lang. | prep | explore | mixed |
| L.Arts/S.S. core Teacher | mixed core | foreign lang. | prep | explore | mixed |
standards for grading and behavior. Every effort was made to make the classroom environments as similar as possible.

My hypothesis states that heterogeneous grouping of 7th grade students will result in increased academic achievement. In order to support my hypothesis, I collected and used extant data to analyze the following variables: students' core GPA's, and students' math test scores. I also used generated data to analyze the affective aspect the grouping strategies had on the students. This was done in the form of quickwrites, in which the students wrote responses of their impressions of the classroom settings and reflected on their own learning. The primary group used for comparison for all of the variables is the entire group of students, that is both the high and the low ability subgroups combined. In addition, I compared the data of the high and low ability subgroups.

The first piece of extant data is the core GPA's for all students in the study. I limited the student GPA comparisons to the core classes of math, science, language arts, and social studies because these were the only classes in which the homogeneous and heterogeneous grouping strategies applied. Physical education, foreign language, and exploration were not part of the study. In addition, these core classes are the most academically rigorous in the students' schedules, and are a better measure of academic progress. Every student's GPA was calculated for the first and second quarter using a point scale of 4 to 0 for the
corresponding letter grades of A to F. The means of these GPA's were used in a t-test in order to determine if their difference was significant. The same analysis was also done for the high ability and low ability subgroups.

The next piece of extant data examined was math test scores. I created the tests, and they were given each Friday. The tests were timed and consisted of 20 problems, most of which were multiple-choice. Calculators were not allowed. The tests aimed to assess student understanding of the material taught during that week, and they also contained selected problems from previous weeks. I recognize that using a non-standardized form of assessment could result in unreliable data. However, these tests are the primary form of assessment I use in my math classes, and they are the format with which all of my students have become accustomed. The means of the students' test scores were calculated for each quarter and compared in a t-test in order to determine if their difference was significant. Analysis was done for the high ability group, low ability group, and both groups combined.

The third and final piece of data collected and analyzed were students' quickwrite responses. At the beginning of the second and third quarters, the students were asked to write their reflections about the previous quarter of instruction. The instructions for the quickwrite
were as follows: “Think about your first (or second) quarter experiences in your L.Arts/S.S. and Math/Sci. classes.”

1. Explain in your own words how you felt about the mood of the classes during first (or second) quarter.

2. Did you feel you learned a lot? Why or why not?

The quickwrite responses used for analysis were limited to those students who were part of the GPA and math test score studies, and who were present for both quickwrites. For the purpose of using a chi-square analysis, these responses were categorized as being either positive (yes), or negative (no). The chi-square tests for significance the difference between the expected and the observed results. Analysis was done for the high ability group, low ability group, and for both groups combined.

However, limitations are in order whenever using student writings. The first is in regard to the authenticity of the students' responses, as reflected by Igoa (1995):

“At times it is difficult to do research with children because they may be apt to say what adults want to hear or they may say 'I don't know' when they are afraid or they don't wish to reveal their feelings or thoughts.” (p 73)
Another limitation is the students may not have understood the writing prompt. I attempted to avoid this problem by reading the prompt to the students, and then leading a class discussion in which they volunteered their interpretations of the meaning of the word "mood". These discussions led to the definition of mood as being related to the students' comfort level (e.g. happy, tense), how they viewed the class learning environment (were they able to learn), and their ability to concentrate (were they distracted?)

This leads to my third and last limitation, which states that in attempting to avoid a misunderstanding of the writing prompt, there exists the possibility I influenced the students' responses. Thus, I recognize the quickwrites could contain unreliable data. However, despite these limitations, feedback from those whom our teaching strategies directly effect, namely the students themselves, can provide meaningful insight. I believe these quickwrites provide valuable information from the students' point of view and they warrant examination.
CHAPTER 4

FINDINGS

Core GPA

The first piece of extant data I examined is the students' core GPA's from the first and second quarter. These data were examined using all of the 56 students as the primary group of study. Next, I looked at the two subgroups, which consisted of students placed into the high and low ability groups for the first quarter of instruction. Table B shows the mean GPA's for all of the students in the study.

<table>
<thead>
<tr>
<th>TABLE B</th>
<th>MEAN CORE GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Quarter</td>
</tr>
<tr>
<td>ALL STUDENTS</td>
<td>2.83</td>
</tr>
</tbody>
</table>

All students were homogeneously grouped for the first quarter of instruction. During this term the mean GPA was 2.83 for all of the students combined. For the second quarter the students were randomly placed into heterogeneous groups. During this term the cumulative GPA was 2.67 for all students. When two groups are being compared, the significance of the difference between the means can be determined by applying a t test, and then consulting a table of t values (Charles 1998). Using the means of the first and second quarter GPA's yielded a t value of
0.81, with a degree of freedom equal to 54. The t value of 0.81 demonstrates that the difference in the mean GPA's for each quarter of instruction was not significant.

For further comparisons, student GPA's were examined according to the high ability and low ability subgroups. This was done because there was the possibility a significant change in the mean GPA for one or both subgroups could exist despite the fact the net effect for the entire group was not significant.

Table C shows that those students who were placed in the high ability group for the first quarter had a combined GPA of 3.32. This number decreased to 3.18 when these students were randomly placed into heterogeneous groups for the second quarter. Again, using these means in a t test yields the t value of 0.64, which is not a significant difference.

Table C also shows that those students who were in the low ability subgroup had a combined GPA of 2.31 for the first quarter of instruction. This number decreased to 2.13 when they were randomly placed into

<table>
<thead>
<tr>
<th>TABLE C</th>
<th>MEAN CORE GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Quarter</td>
</tr>
<tr>
<td>HIGH GROUP</td>
<td>3.32</td>
</tr>
<tr>
<td>LOW GROUP</td>
<td>2.31</td>
</tr>
</tbody>
</table>
heterogeneous groups for the second quarter. Using these mean GPA's in a t test yields the t value of 0.67, which is not significant. Therefore, the mean GPA's of the whole group as well as both subgroups demonstrates no significant changes occurred when comparing both quarters of instruction.

Math Test Scores

The next piece of extant data to be compared is the students' mean test scores for mathematics. I followed the same procedure used for comparing student's GPA's. First, I looked at the entire group, then I looked at the high ability and low ability subgroups. Table D shows the mean test scores for all groups.

<table>
<thead>
<tr>
<th>TABLE D</th>
<th>MEAN MATH TEST SCORES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Quarter</td>
</tr>
<tr>
<td>ALL STUDENTS</td>
<td>79.83%</td>
</tr>
<tr>
<td>HIGH GROUP</td>
<td>85.13%</td>
</tr>
<tr>
<td>LOW GROUP</td>
<td>74.13%</td>
</tr>
</tbody>
</table>

* p < .05

The primary group for comparison is all students, who had a mean math test score of 79.83% for the first quarter of instruction. This average dropped to 75.33% for the second quarter of instruction. The difference between these two means, when applied in a t test yields a t value of 2.00,
which is below the t value of 2.01 needed for significance at the .05 level, and is therefore not significant.

Next, the high ability subgroup had a mean math test score of 85.13 for the first quarter. This average dropped to 82.72 for the second quarter. Using these numbers in a t test yields the t value of 1.19, which is not significant. Those in the low ability subgroup had a mean math test score of 74.13 for the first quarter of instruction. This average also dropped to 67.39 for the second quarter, when all of the students were randomly placed into heterogeneous groups. The difference between these means yields a t value of 2.10, which is significant at the .05 level. Therefore, when comparing the means of the math test scores, there was a significant difference in the low ability subgroup, but this was not sufficient to effect a significant difference in the overall group.

Quickwrites

Research has shown that the grouping practice employed in this study slightly lowers the self-esteem of the higher-aptitude students while slightly raising the self-esteem of lower-aptitude students (Kulik 1992). My third and final data set, the student quickwrites, revealed a pattern similar to these results (Tables E and F).

Chi-square is an inferential statistical procedure that is frequently used in research when data can be placed into categories (Charles 1998). Therefore, for the purposes of conducting a chi-square analysis, I
categorized the student responses as either positive (yes), or negative (no). One expects to observe equality of choice when conducting a random sampling (Charles 1998). Thus, I expected student responses to be 50% positive and 50% negative. By statistically comparing what I expected to observe against what I actually observed, I was able to determine whether the differences between the two were significant (Charles 1998). Again, I considered all of the students as the primary group of study. Then, I looked at the high ability and the low ability subgroups. Table E shows the results of the quickwrites for the first quarter of instruction.

<table>
<thead>
<tr>
<th>TABLE E</th>
<th>QUICKWRITES - FIRST QUARTER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed YES</td>
</tr>
<tr>
<td>2)</td>
<td></td>
</tr>
<tr>
<td>ALL STUDENTS</td>
<td>33</td>
</tr>
<tr>
<td>HIGH GROUP</td>
<td>21</td>
</tr>
<tr>
<td>LOW GROUP</td>
<td>12</td>
</tr>
</tbody>
</table>

** p < .01

The sum total of student quickwrite responses used for this analysis was limited to those students who were part of the GPA and math test score comparisons, and who were present for both quickwrites. This left a total of 48 responses which I was able to use for the chi-square calculations. I would expect one-half, or 24 of the responses to be
positive (yes). The observed number of yes responses was 33. A chi-square calculation for all of the students yielded a value of 3.38 with a degree of freedom equal to 1, which is not significant. However, the high ability subgroup had 21 observed yes responses, which when compared to the expected value of 12 yields a chi-square value of 6.75, a difference significant to the .01 level. The low ability subgroup responded as expected and yielded a chi-square of zero.

Table F shows the quickwrite results for the second quarter. The observed number of yes responses was 34. When compared to the expected result of 24, the chi-square calculation yields a value of 4.17, which is significant to the .05 level. The exponential aspect of the chi-square calculation shows a significant difference for the whole group despite the fact both the high and low subgroups had values of 2.08, which are not significant.

<table>
<thead>
<tr>
<th>TABLE F</th>
<th>QUICKWRITES - SECOND QUARTER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed YES</td>
</tr>
<tr>
<td>ALL STUDENTS</td>
<td>34</td>
</tr>
<tr>
<td>HIGH GROUP</td>
<td>17</td>
</tr>
<tr>
<td>LOW GROUP</td>
<td>17</td>
</tr>
</tbody>
</table>

* P < .05
The quickwrites show a pattern in which the positive responses by the higher-aptitude students dropped slightly, from 21 to 17, when they were heterogeneously grouped. At the same time, the positive responses of the lower-aptitude students rose slightly, from 12 to 17. Overall, there were 34 positive responses to the heterogeneous grouping. A chi-square analysis of these positive responses showed a significant difference to the .05 level. The results reveal a crucial aspect of student learning, that is, the affective environment of the classroom.

The quickwrite responses for the first quarter of instruction show that one group of students, the high core, was very happy with their classroom setting, and another group, the low core, was not. While 88% of the high core students responded positively to the homogeneous groups, only 50% of the low core responded likewise. Therefore, the result was one class (high core) in which the students were motivated to learn and well-behaved, while the other class (low core) experienced far more behavior problems and interruptions. Samples from the students' quickwrites illustrated this point many times.

These are quotes from the high group quickwrites, regarding the homogeneous grouping for first quarter:

"...the mood was cheery. I had a good learning atmosphere."

"I think that everyone was attentive (in my class) and
into the learning process.”;

“There wasn’t a lot of behavior problems in my class so we were able to get some work done.”;

“I love the way that there wasn’t any craziness.”;

“I was able to cooperate.”

Other words used to describe the classroom were “happy”, and “calm”. It was clear the students enjoyed the atmosphere of the class and the fact there were few distractions.

By contrast, the low group quickwrites regarding the first quarter of instruction paint a different picture:

““Our class was bad, we were always messing around.”;

“My classmates would often play around and talk...”;

“I didn’t like my class for first quarter because we would always get in trouble...”;

““I didn’t like the class I was in because we had all the screw offs and it seemed like we were not getting anything done.”;

““I didn’t feel like I learned much last quarter because the rowdy kids would always mess around in class.”

Other words used to describe the class setting were “disruptive”, “loud”, and “noisy”. Clearly, many students were bothered by the class setting.
Summary

An examination of the data when students are grouped homogeneously and heterogeneously shows no significant differences in the core GPA's and the math test scores for the entire group of students. However, the student quickwrite responses favoring heterogeneous grouping practices do show a difference significant to the .05 level. This is contrary to their responses to the homogeneous grouping practices, which showed no significant difference.
CHAPTER 5
DISCUSSION

I have compared the academic achievement of 56 students when they were grouped homogeneously (first quarter) and heterogeneously (second quarter) for instruction in their core classes. In both situations, the students followed the same basic curriculum. Table G reveals a pattern of decline for all students’ core GPA’s and math test scores from first to second quarter.

<table>
<thead>
<tr>
<th>TABLE G</th>
<th>ALL STUDENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Quarter</td>
</tr>
<tr>
<td>CORE GPA</td>
<td>2.83</td>
</tr>
<tr>
<td>MATH TEST SCORES</td>
<td>79.83%</td>
</tr>
</tbody>
</table>

Nevertheless, my findings show no significant difference in the students’ core GPA’s and math test scores associated with the grouping practices. This is consistent with the findings of similar programs, known as XYZ grouping (Kulik 1992). In these types of programs, students are placed into high, medium, and low classes and follow the same basic curriculum. Programs of this type have little or no effect on student achievement (Kulik 1992). Slavin (1995) concurs with these results, which are near zero for high, medium, and low achievers.
A limitation in these findings is the possibility the decline in student achievement may reflect the effect of conducting this study at the beginning of the school year more so than it reflects the grouping strategies employed. The curriculum during first quarter contained considerable review from the sixth grade. As the year progressed, the curriculum became more demanding. Therefore, if I were to replicate this study, one change I would make would be to conduct it during the second and third quarter of the school year. By doing so, the newness of the school year would have worn off and the level of curriculum would be more consistent.

The contrast in classroom environments experienced by my students during the first quarter are illustrative of what can happen when children are tracked. The high-functioning students are placed into a class which is well-behaved and productive for learning. On the other hand, underachieving students, and those with learning disabilities and behavior problems, are lumped together forming an environment in which learning is difficult. Kathleen Cotton (1991) identified practices harmful to students, one of which is academic tracking. She writes:

"Many volumes have been written about the harmful effects of academic tracking on those assigned to low tracks. Ironically, low-track placements, which are supposed to help slower learners...make learning an
unproductive and unpleasant experience.”

Her findings show that students in these lower-ability tracks experience lower levels of student-student cooperation and support, and have lower attitudes about themselves as learners. They also experience more student and teacher interruptions in their classes and more “dead time”. This leads to my next point, which is the relationship between the teacher and the students are crucial to the success of any classroom setting. More than any other factor in which schools have some control over, it is the interaction between the teacher and students which has the greatest impact on student achievement. Jaeger & Hattie (1995) write:

“The effects of ability grouping are trivial when compared to those of intra-classroom innovations controlled by teachers. Teachers, their training and background, and the choices they make have a far greater influence than do any innovations having to do with class organization, including ability grouping.” (p.218).

The homogeneous grouping of the students in this study directly influenced their day to day interactions with one another and with their teachers, that is my partner and me.

I experienced the reality of having two core classes which were very different during the first quarter. The high core was pleasant and wonderful to teach, and the low-core was a constant challenge of my
resolve to maintain a positive learning environment. Both classes received the same basic curriculum, but in many instances the high core was able to receive and follow directions much more quickly and efficiently, especially in math. Thus, they were usually able to complete the basic problem set for that day and move on to the more challenging problems. By contrast, it was a struggle for the low core students to complete the basic problem set each day.

This difference challenged my patience and attitude in teaching with the low core group. To maintain order and complete the daily lesson, I was forced to become more strict and rigid in my approach to discipline. I was less likely to utilize cooperative learning and group assignments because many students had difficulty staying on task. Teaching involved more seatwork and direct instruction than in the high core class. Thus, the learning experience for the low core was not as rich as it could have been. Also, if I had to teach a low core class for the entire school year, I believe it would have eroded my enthusiasm for teaching this group.

Research supports my experiences. Students in lower-ability groups experience less interactive teaching, less teacher enthusiasm and encouragement, and have more and longer periods of seatwork (Cotton 1991; Oakes 1985; Slavin 1990).

The student quickwrites about the second quarter, in which they were heterogeneously grouped, were much more positive. To illustrate
my point, I wish to compare quickwrite samples from the same students I used earlier in my discussion. These are quotes from five low-core students showing their responses to the first and second quarter of instruction, respectively:

Student A: “Our class was bad, we were always messing around.” (1\textsuperscript{st} quarter);
“I felt the mood was OK.” (2\textsuperscript{nd} quarter)
Student B: “My classmates would often play around and talk…”;
“I think we were noisy and rambunctious.”
Student C: “I didn’t like my class for first quarter because we would always get in trouble…”;
“I think we are actually (some of us) maturing…”
Student D: “I didn’t like the class I was in because we had all the screw offs and it seemed like we were not getting anything done.”;
“I think the mood of the classes was OK.”
Student E: “I didn’t feel like I learned much last quarter because the rowdy kids would always mess around in class.”;
“I think the mood of the classes was good.”
These samples show the class environments, although not perfect, improved when the students were in heterogeneous groups. Revisiting tables E and F shows that each core class for the second quarter had 17 positive responses out of 24. The net result was a situation in which I had two core classes which were well-behaved and manageable. This contrasts to the situation during first quarter when I had one excellent class and one poor class. Thus, the heterogeneous groups eliminated the poorly-behaved classroom and established a better overall environment for all students.
CHAPTER 6
CONCLUSION AND RECOMMENDATIONS

Student diversity in American public schools is a reality both now and in the future. This diversity is not limited solely to race and class, but also includes learning styles and academic ability. Thus, teachers today often have the daunting challenge of educating students of varied backgrounds and talents within a single classroom setting. The traditional practice of tracking children may not provide the best solution for educating these students. Research on the benefits of tracking are mixed, and does not prove tracking improves academic achievement. In fact, some of the research shows tracking has negative effects on some students’ attitudes about school.

The limitations of this study prevent me from conclusively stating that tracking students for instruction should be eliminated from our schools, or is bad educational practice. Nevertheless, the results of this study show no significant changes in academic achievement as measured by the chosen variables of GPA’s and math test scores. The results also show students’ perceptions towards the homogeneous grouping practice to be significantly lower than toward the heterogeneous grouping. Despite the inherent limitations of my being both the researcher and the classroom teacher in this study, this experience has influenced my belief toward the idea that the teacher to student interaction may be the most important
factor for educational success. Therefore, my suggestion for further studies on the topic of grouping students for instruction would focus on teacher training and programs of enrichment.

Upon reflection on my own practice, I have not mastered the art of teaching a classroom with children of wide-ranging abilities, and am therefore not adequately prepared. From my own observations and experiences, I believe many other teachers are also not prepared. Therefore, I would like to see more studies conducted where teachers receive professional development targeted towards teaching all kinds of learners in a heterogeneous environment. These studies could compare academic achievement, student attitudes toward school, and also teacher attitudes to determine the success of the treatment.

I recommend school administrators focus more attention on teacher training for improving student performance. Having teachers attend one or two workshops on the subject is inadequate. Schools need to have ongoing professional development for teachers on differentiating instruction in the classroom. Training must be consistent and have a measure of accountability. In addition, teachers need feedback on their programs of instruction in order to improve their practice. Meaningful sources of feedback may come from other teachers, administrators, and from the students. Thus, a consistent professional development program with a
strong evaluation component would go a long way toward improving classroom instruction.

The second key area for further study should examine the effects of providing high-quality materials and programs of enrichment for teachers to use with their students. My own experience has been that teachers are often asked to teach with little or no instructional materials. Many are asked to use old and outdated materials, while those fortunate enough to have new materials are usually not trained on their proper use to maximize their effectiveness. Thus, the burden is often put on individual teachers to find or develop their own curriculum. Given the time constraints inherent to teaching, this is an unrealistic expectation. In light of these realities, it is easily seen why the simplicity of tracking children becomes attractive to educators.

A general recommendation for schools is increased administrative support for heterogeneous classes. Tracking students is a deeply rooted practice that will not be easy to change. Parents want the best education possible for their children. Therefore, our schools must relate the message to them that tracking is not in their children’s best interest. Administrators, especially site principals, are the key to good public relations with the parents and community. School administrators can also work with the universities to merge educational research with teacher practice in the classroom. Finally, the lack of adequate training is often
compounded by the practice of assigning teachers to areas outside of their expertise. Districts need to be aggressive in recruiting top candidates for positions in their schools, making every effort to acquire sufficient numbers in all subject areas.

It appears Americans have decided the most sensible response to the differences among students in their public schools is to separate them for instruction. The most important issue which tracking raises is equity. Although it was not a factor in this study, there is clearly a pattern of unequal distribution of educational resources when one examines the difference between high and low-ability groups. Therefore, it is not surprising that many studies show these different groups attain differentiated levels of achievement. Those in the high-ability groups tend to do better, and those in lower groups do worse. Perhaps the most disturbing aspect of tracking is the relegation to the lower tracks of ethnic minority groups, people of color, and those from disadvantaged circumstances. Studies show most of the students who make up the lower-tracks come from these groups. Students who are tracked tend to remain in whatever grouping they are assigned to, and this process begins at an early stage in a student's educational career. Thus, there is evidence schools may be propagating a continuing cycle of underachievement among students of color and impoverished backgrounds.
Because of the diverse nature of American students, the many learning styles and abilities they bring to the classroom, and their large numbers in a single classroom, homogeneous grouping appeals to the common sense of many people. Teaching is at best a difficult undertaking, and tracking does somewhat lighten the burden on classroom preparation and instruction. However, continuing a practice which may be flawed simply because it is easier to employ does not make for sound educational practice. Instead, reform efforts need to continue to focus on teacher training, improved instructional materials and programs of enrichment, and administrative support for heterogeneous grouping of students.
References


