



12-1-2002

## The Effects of Class Size on Student Performance in Reading in the Primary Grades

Dotti S. Dixon

Follow this and additional works at: <https://commons.und.edu/theses>

---

### Recommended Citation

Dixon, Dotti S., "The Effects of Class Size on Student Performance in Reading in the Primary Grades" (2002). *Theses and Dissertations*. 1071.  
<https://commons.und.edu/theses/1071>

This Dissertation is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact [zeineb.yousif@library.und.edu](mailto:zeineb.yousif@library.und.edu).

THE EFFECTS OF CLASS SIZE ON STUDENT PERFORMANCE  
IN READING IN THE PRIMARY GRADES

by

Dotti S. Dixon  
Bachelor of Science, North Dakota State University, 1977  
Master of Education, North Dakota State University, 1987

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota  
December  
2002

UMI Number: 3088042

**UMI**<sup>®</sup>

---

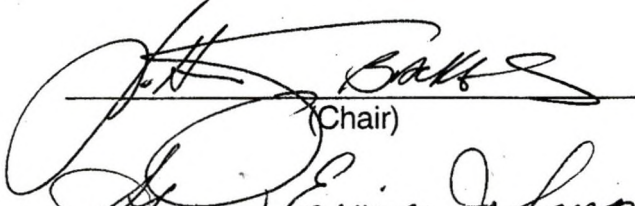
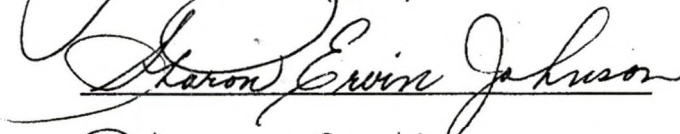
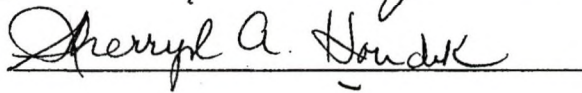
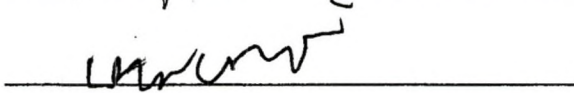
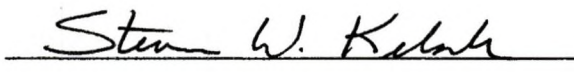
UMI Microform 3088042

Copyright 2003 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

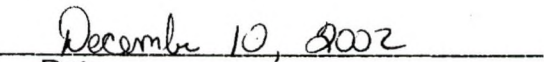
ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

This dissertation, submitted by Dotti S. Dixon in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

  
\_\_\_\_\_  
(Chair)  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_

This dissertation meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

  
\_\_\_\_\_  
Dean of the Graduate School

  
\_\_\_\_\_  
Date

## PERMISSION

Title            The Effects of Class Size on Student Performance in Reading  
                    in the Primary Grades

Department    Educational Leadership

Degree         Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in his absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Signature

*Walter J. Dixon*

Date

*Nov. 5, 2002*

## TABLE OF CONTENTS

LIST OF TABLES.....	vi
ACKNOWLEDGMENTS.....	viii
ABSTRACT.....	x
CHAPTER	
I.    INTRODUCTION.....	1
Statement of the Problem.....	7
Significance of the Study .....	9
Delimitations of the Study.....	11
Definition of Terms .....	11
Summary and Overview .....	13
II.   REVIEW OF LITERATURE IN CLASS SIZE AND PRIMARY STUDENTS' ACHIEVEMENT IN READING .....	15
Historical Perspectives .....	15
State Class Size Reduction Initiatives .....	22
State Studies Showing No Correlation Between Class Size and Student Achievement .....	39
Summary .....	48
III.  RESEARCH DESIGN .....	56
Population .....	56
Instruments .....	59
IV.  ANALYSIS OF THE DATA .....	62
Class Size and Reading Achievement in the Primary Grades .....	64

Class Size and Reading Achievement Gains Across Grades .....	71
Class Size and Reading Achievement Gains Made by Students Who Are Economically Disadvantaged .....	73
Class Size and Reading Achievement by Gender .....	75
Class Size and Reading Achievement of Native American vs. Caucasian Students .....	77
Teacher Questionnaire Results, Grades, 1, 2, and 3, School Year 2001-02 .....	79
Summary.....	87
V. SUMMARY, DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS .....	90
Summary of the Study .....	90
Discussion of Research Questions .....	92
Discussion of Teacher Survey .....	99
Limitations .....	102
Conclusions .....	103
Recommendations for Action .....	105
Recommendations for Further Study .....	107
APPENDICES.....	110
REFERENCES.....	129

## LIST OF TABLES

Table	Page
1. Characteristics of All Small and Large Classroom Students From 1999-2002 by Frequency and Percent .....	65
2. Number of Students in Small and Large Classrooms in Bismarck Public Schools by Gender, Ethnicity, Socioeconomic Status, and Grade From 1999-2002 .....	66
3. Total Reading Extended Pre and Post Scale Scores and Normal Curve Equivalency Pre and Post Scores (Means and Standard Deviations) for First Grade Students From 1999-2002.....	68
4. Total Reading Extended Pre and Post Scale Scores and Normal Curve Equivalency Pre and Post Scores (Means and Standard Deviations) for Second Grade Students From 1999-2002 .....	69
5. Total Reading Extended Pre and Post Scale Scores and Normal Curve Equivalency Pre and Post Scores (Means and Standard Deviations) for Third Grade Students From 1999-2002 .....	70
6. Differences of Mean Pre-Post Composite Total Reading Extended Scale Scores and Normal Curve Equivalency Scores in Primary Grade Levels in Small vs. Large Size Classrooms Over Three Years (1999-2002) .....	72
7. Total Reading Scale Scores and Normal Curve Equivalency Scores of Economically Disadvantaged Students Compared to Non-Economically Disadvantaged Students (Grades 1, 2, and 3) on the Gates-MacGinitie Reading Test .....	74
8. Total Reading Scale Scores and Normal Curve Equivalency Scores From 1999-2002 for Male and Female Students in Small vs. Large Classrooms Compared in Grades 1, 2, and 3.....	76
9. Total Reading Scale Scores and Normal Curve Equivalency Scores of Caucasian Students Compared to Native American Students in Small vs. Large Classrooms in Grades 1, 2, and 3 .....	78



10.	Frequency and Percentage of Teacher Responses to Items Related to Time Allotment and Classroom Directed Efforts for Both Small and Large Classrooms .....	82
11.	Comparison of Mean Teacher Responses in Small vs. Large Classrooms as to Where They Spend Their Time and How They Direct Their Classrooms .....	85
12.	Frequency and Percentage of Teacher Responses to Areas Where They Would Like to Spend Additional Time .....	86

## ACKNOWLEDGMENTS

I would like to thank my husband, Randy Schmeling, and son, Brock, for their support and understanding as I traveled to Grand Forks for class over a four-year period. I would also like to thank my parents and siblings for their continued encouragement as oftentimes I think they thought I would never be done taking classes. A special thank you to my sister, Billie, who provided bed, breakfast, and plenty of conversation during my overnight stays in Grand Forks. Mom, Dad, I think your early childhood advice to “get your education” sunk in deeply and has continued to stay with me!

My doctoral education was partially funded through the generosity of the Delta Kappa Gamma Society. I received two state scholarships, the Adeline Stevenson Nurse Scholarship, and the Hazel B. Nielson Scholarship, and one international gift, the J. Maria Pierce Scholarship, that were extremely helpful in defraying tuition costs. In addition to the monetary support, I want to express my thanks to my local chapter, Delta, and the wonderful support that the membership has given to me as I pursued my studies.

As a member of the first cohort in Educational Leadership to graduate from the University of North Dakota, I want to express my gratitude to the Educational Leadership department and to the University for providing an opportunity for us distant learners to be able to fulfill an educational dream. I also want to thank my cohort colleagues for their friendship, inspiration, and fun. The cohort friendships will last a lifetime, and my life is richer because of each member. In addition, I want to thank my Bismarck traveling companions, Lisa, Tom, Anita, Rick, Donna, and Cindy, for making the travel time go by so quickly. Traveling to and from Grand Forks was a special time because of all of you!

My chairperson, Dr. John Backes, was there when I first started thinking about furthering my education and ended up providing me with the professional guidance and motivation to complete my dissertation. Thank you for agreeing to be there, both at the beginning and at the end of my studies at the University of North Dakota. I would also like to thank the many members of the Educational Leadership faculty who provided a plethora of learning experiences throughout the program--both theoretical and practical. Dr. Steve Kelsch, Dr. Sharon Johnson, Dr. Gerald Bass, Dr. Sherryl Houdek, and Dr. Donald Daughtry, thank you for your insight and unique backgrounds in serving on my committee.

From the Bismarck school district, I would like to thank Dr. Lowell Jensen, Dr. Paul Johnson, Rick Buresh, Jim Potter, and Bill Demaree for allowing me to participate in the doctoral program and for providing the information and help needed to study the effects of class size in the Bismarck school district. A special thanks goes out to the teachers who took the time to meticulously administer student tests and teacher surveys for the dissertation study. The entire staff at Wachter Middle School has been wonderful in providing inspiration to stay focused and finish my degree. Thank you all!

Last, but foremost, my many thanks go out to Sharon Fields, who was the "constant" in the Educational Leadership program. She is not only talented and gifted professionally but is truly a remarkable person as well. Her friendly voice was the encouragement that made the difference, especially when the distance was 240 miles! My sincere gratitude for all the phone calls and administrative help throughout the past four years. I could not have done it without you.

## ABSTRACT

The purpose of this study was to examine the effect of class size on reading achievement of primary grade students in the Bismarck (North Dakota) Public Schools. Data for the study included using scores from the Gates-MacGinitie Reading Test (Fourth Edition). The results on the Gates-MacGinitie Reading Test were analyzed according to grade, gender, socioeconomic status, race, and class size. Students in the study were identified as having been in either small classes (17 or less) or large classes (22 or more). Three years of data (1999-2002) were analyzed as part of the small class study. A second source of data was gathered from a teacher survey instrument that was distributed to all teachers employed by the Bismarck Public Schools for the 2001-2002 school year who taught in classrooms identified as small or large. This survey measures teachers' perceptions regarding instructional practices, classroom management, and time allotment.

Results indicated that there was no significant difference in reading achievement scores of first grade, second grade, or third grade students who were placed in small compared to large classrooms. There was a significant difference in female students' reading achievement when they were placed in small sized classrooms. When the achievement of Native American students was compared to Caucasian students, there was a significant difference in reading achievement scores by Caucasian students in small sized classrooms. There was also a significant difference in reading achievement scores by students not on free-reduced meal plans. Consequently, this study does not suggest that small class size is an equalizer for Native American students, males, or students who are economically disadvantaged. The major finding from the teacher survey showed that teachers in large sized classrooms had had more current professional development on reading strategies, which may have contributed to the class size achievement showing no

significance when in small sized classrooms. By using analysis of variance (ANOVA), the study found that the following items were significantly different when teachers in small sized and large sized classrooms were compared: teaching leans toward students as individuals rather than towards the class in general (.016), time spent on disciplining the class (.008), time working with small groups (.041), time spent with students one on one (.017), time spent working with students on special projects (.004), and time spent developing creative projects for the class (.008).

## CHAPTER I

### INTRODUCTION

Class size is a critical issue for parents, teachers, administrators, and government. Parents and educators have argued that smaller class size leads to more effective teaching and improved learning for students. Students in smaller classes are reported to have higher achievement levels, fewer discipline problems, less retention, and more personal attachment to their teachers and classmates (Achilles, Finn, & Bain, 1997/1998; Word et al., 1990). Researchers have found that small classes provide additional benefits for minority and low-income students, especially in high-poverty areas (Finn & Achilles, 1990; Robinson, 1990; Wenglinsky, 1997). However, various state and federal educational leaders, until the last decade, have argued that substantial reductions in class size are too costly and not effective (Brophy, 2000; Krueger, 2000; National Institute on Student Achievement, Curriculum, and Assessment, 1998; Nye, Boyd-Zaharias, Fulton, & Wallenhorst, 1992; Wang, 2000).

In response to a 1988 Phi Delta Kappa Gallup Poll cited by Folger (1989), 77% of parents and 68% of non-parents reported that having a small class made a difference in student achievement. The issue of what number of students constitutes the ideal class size, however, differs from researcher to researcher. Some believe that 15 is the "magic number," as was the case in the Burke County, North Carolina, initiative to reduce class size in the primary grades that was begun in 1990 and continued to expand in the years that followed. The Wisconsin Reduction Program called the Student Achievement Guarantee in Education (SAGE) Program used 15 to 1 or less in grades K-3. Ferguson (1991) found that district student achievement fell as the student-teacher ratio increased for every student above an 18 to 1 ratio. In 1998, Congress responded to the President's call for a national initiative to lower

class size in the early grades to no more than 18 students as part of the 1999 Department of Education Appropriations Act.

According to data from the U.S. Department of Education (2000), prior to the implementation of the federal Class Size Reduction Program and similar initiatives in several states, more than 85% of United States students were in classes with over 18 children, and about 33% were in classes of 25 or more students. After two years of implementation, the Class Size Reduction Program was providing funding for communities to lower class size in the early grades.

It is a propitious time to consider questions about class size. After years of debate, speculation, and research that yielded only partial and less-than-definitive answers, a major longitudinal study has provided new answers to the question "Do small classes result in greater academic achievement in the elementary grades?" Begun in 1985, Tennessee's Project STAR (Student Teacher Achievement Ratio) set the stage for asking and answering a number of policy related questions. The focus of the Project has been on the effects of small classes in the early elementary grades. There are two reasons for this. First, the most current and best research to date had been conducted in kindergarten through grade 3 and the state of research with respect to small classes in the upper grades was fragmented and even contradictory. Second, there are good reasons for starting research and intervention projects in the early grades based upon the assumption that the early years of schooling lay the foundation for much that follows is explicit and has been substantiated repeatedly by research in the social sciences (Nye, Fulton, Boyd-Zaharias, & Cain, 1995).

In their article addressing exemplary literacy learning programs, Strauss and Irvin (2000) look at the needs of struggling readers. They discovered that much of the academic success in schools is based on students' ability to read and comprehend text, so when students struggle with reading demands, they experience frustration and failure in most classes (Moore, Bean, Birdyshaw, & Rycik, 1999; Strauss, 2000). Small classes taught by certified teachers, who offer students

daily opportunities to read and be involved with text, allow these less confident readers to see themselves as readers (Ivey, 1999).

The question, "Are smaller classes better than larger classes?," continues to be debated among teachers (and their unions), administrators, and parents as well as in the research community. The issue persists because of the powerful common-sense appeal of small classes to alleviate problems indigenous to classrooms. Smaller classes are an integral component of nationally subsidized programs including special education classes for disruptive or learning-disabled students and Title I interventions for children living in poverty. Small classes or small groups working with one teacher or tutor also are a key element of programs targeted most often at students at risk, for example, Success For All (Slavin, 1989; Slavin & Madden, 1995) and Reading Recovery (Pinnell, deFord, & Lyons, 1988).

The issue persists because of the tension between the research findings and the cost of implementation. A great deal of empirical data has been collected. The data have so far been less than convincing and not consistent enough to justify the expense of the additional classrooms and teachers that would be required. Targeted remedial programs are generally less costly and easier to deploy. Remedial programs tend to be adopted for a portion of the school day to address learning problems in one or a small number of subject areas. In contrast, maintaining small classes throughout a grade level or school requires pervasive organizational changes. The proponents would argue that the benefits are also pervasive--being realized throughout the school day and affecting the entire range of school subjects--unlike the band-aid approach of experimenting with one targeted program after another.

Over the past two decades, there have been many summaries of research on the relationship of class size to academic achievement. The Meta-Analysis of Research on the Relationship of Class Size and Achievement (Glass & Smith, 1978), the compilation of studies examined by Educational Research Service (Robinson, 1990; Robinson & Wittebols, 1986), and the Slavin (1990) "best evidence synthesis" review are three summaries worth noting because of their



comprehensiveness and because they planted the seeds for much of the research that followed.

The most widely cited review is the classic Meta-Analysis of Research on the Relationship of Class Size and Achievement (Glass & Smith, 1978). The authors collected and summarized 77 studies of the relationship of class size with academic performance that yielded over 700 class size comparisons on data from nearly 900,000 pupils. The two primary conclusions drawn from this material are that reduced class size can be expected to produce increased academic achievement, and the major benefits from reduced class size are obtained as the size is reduced below 20 pupils.

A compilation of studies examined by Educational Research Service (Robinson, 1990; Robinson & Wittebols, 1986) reviewed more than 100 separate studies. Robinson's (1990) conclusions added an important set of qualifications to the findings of Glass and Smith. He found that research does not support the expectation that smaller classes will of themselves result in greater academic gains for students. The effects of class size on student learning varies by grade level, pupil characteristics, subject areas, teaching methods, and other learning interventions. The review concludes that small classes are most beneficial in reading and mathematics in the early primary grades, and that the research consistently finds that students who are economically disadvantaged or from some ethnic minorities perform better academically in smaller classes.

The third review focused on high-quality research conducted in accordance with accepted scientific standards. Using a procedure termed "best evidence synthesis," Slavin (1990) reviewed only those studies that lasted a minimum of one year; involved a substantial reduction in class size--that is, larger classes were compared to classes that were at least 30% smaller and had 20 students or fewer; and involved either random assignment of youngsters to class sizes or matching to assure that the groups were initially equivalent. Slavin found only eight studies that

met all three criteria. From these studies, he concluded that substantial reductions in class size have a small positive effect on students.

In a brief overview of research, Finn and Voelkl (1994) identified three approaches to studying the issue of class size: the classroom focus approach (Glass & Smith, 1978; Robinson & Wittebols, 1986; Slavin, 1989), the cost-related approach, and the ecological approach. The classroom focus research examines the number of pupils in each classroom, the interactions between the teachers in those classrooms, and the outcomes that were realized by the pupils in those classrooms. It provides the most direct and intensive view of the effects of a small class setting.

The cost-related approach examines the actual or potential costs of implementing small classes and weighs them against the benefits that may accrue. The cost issue is raised by researchers (Tomlinson, 1990) and by state and local policymakers who control the purse strings. The production function approach relies heavily on multiple regression analysis to relate a series of inputs (such as cost factors) to an output (such as student achievement). Hanushek (1986) reviewed 112 studies that used educational production functions to examine the effects of instructional expenditures on student achievement using indicators such as teacher experience, teacher education, and pupil-teacher ratio. Pupil-teacher ratio was statistically significant in only 23 of the 112 studies, only 9 of which were significant in the expected direction. This procedure led Hanushek to conclude that pupil-teacher ratio is not an important correlate of student performance. More sophisticated analyses of the same data have led others to conclude that low pupil-teacher ratios (and other cost-related inputs) are associated with increased pupil performance (Hedges, Laine, & Greenwald, 1994; Laine, Greenwald, & Hedges, 1995).

In an analysis of national survey data at the district level, Wenglinsky (1997) concluded that expenditures to reduce pupil-teacher ratios impact positively on academic achievement at grade 4 but not at grade 8. Ferguson and Ladd (1996) analyzed achievement scores for students in grades 4, 8, and 9 of 131 districts in Alabama. These researchers used average class size in their multi-level regression

models instead of pupil-teacher ratio, concluding that class size does matter in both the earlier and later grades.

Most production function analyses include schools and districts with classes within “normal” ranges--22 to 40 students or so--and the results do not answer the question of what the impact would be if classes were reduced substantially. The step cannot even be taken “in theory” since reductions in class size would change the values of other important inputs as well (Finn, 1998). Another concern is that most production function analyses focus on school-wide or district-wide pupil-teacher ratios rather than actual class size. The Ferguson and Ladd (1996) research is an exception.

The ecological approach views class size in historical or geopolitical perspectives. For example, Tomlinson (1988) examined the changes in median class size in the United States over several decades and related them to changes in standardized test scores. The analysis does not show performance benefits for smaller classes, and it ignores other intervening factors such as population shifts and both cultural and institutional changes over the same period of time. Also, the comparison of class sizes between countries introduces a number of confounding variables including national differences in educational expenditures, educational goals, teacher preparation, and student characteristics. Class sizes may vary dramatically within a country over time or among schools at one point in time (Finn & Voelkl, 1994). Thus, ecological association with pupil performance only obscures the effects of having a smaller or larger number of individuals in a particular class setting.

In this study, the classroom focus approach as identified by Finn and Voelkl (1994) was used, as it best provides direct and intensive review of the effects of a small classroom setting. If there is evidence that reading levels are higher when class size is lower in the early grades, then this has implications for school policy, since the ability to read well is an important factor in later achievement. Thus, if lower class sizes mean higher reading achievement, it would be in the best interest of students to ensure that their development not be handicapped by large classes.

As a study in Madison, Wisconsin, explained in 1973, It is plausible to think that a student's reading achievement is somehow related to the size of the student's class size. After all, fewer people in the student's class means less competition for the teacher's time and attention, which in turn might well have something to do with how well the student learns to read. . . . The reading ability of pupils is fundamental to the learning process. (Madison Public Schools, 1976, p. 3)

#### Statement of the Problem

Today, more than ever before, public school classrooms are composed of a more diverse student population with varied academic, emotional, and social needs. Society is mandating that the public schools assume a greater role beyond the academic education of the school age population. States are mandating more rigorous standards, and more complex curriculum objectives are introduced to students at earlier ages. Demands are placed on teachers to excel in instructional strategies, produce higher achievement test scores, and create within the classroom a harmonious environment which will propagate the next generation of workers for the nation. There are population shifts and growth in some regions of our country, but schools within the state of North Dakota will continue to see a major decline in the number of students entering school for the foreseeable future.

There is a need within the educational community to identify the effect that class size has and to substantially document the benefits that lowering pupil-teacher ratios, or in some cases keeping numbers low, will have on student achievement in public schools. There is a need to examine the effect of smaller class size on reading in the primary grades where the basis for all other learning is developed. Currently, there is not adequate research data on which North Dakota school districts or regions within the state can make decisions regarding the allocation of funds.

The purpose of the study was to determine if there would be achievement differences in mean reading test scores for first, second, and third graders associated with the reduction of class size in the Bismarck Public Schools. It was hypothesized that such reduction of class size would raise significantly students' achievement

scores on a standardized reading test. It was further hypothesized that differences in race, sex, and socioeconomic status traditionally found among achievement scores of first, second, and third grade children would be eliminated since teachers would now have more time for individualized instruction. After three years of comparing small size classes and large size classes in the district, this study attempted to provide answers to the following research questions:

1. What effect does class size have on the reading achievement of first grade students?
2. What effect does class size have on the reading achievement of second grade students?
3. What effect does class size have on the reading achievement of third grade students?
4. Is there a significant difference in achievement made at different grade levels? (Example: Are there greater gains in first grade than in second or third grade?)
5. Are there substantial achievement gains made by students who are economically disadvantaged as a result of small class size?
6. Is there a gender difference between small and large classrooms and student achievement?
7. Do Native American students have greater achievement gains than other students as a result of small class size?

A secondary purpose of the study was to determine what reading instructional characteristics of small size classrooms were used within the Bismarck Public Schools. Questions examined were "Are small sized classrooms organized differently?" and "Is more time spent on reading when class size is reduced?" Adapted from the SAGE qualitative research procedure on effective teaching in reduced-size classes, a self-reporting teacher questionnaire was developed to assist in answering the instructional characteristics aspect of small class size. Teachers who taught in both small and large sized classrooms within the Bismarck Public Schools were selected to participate.

## Significance of the Study

There is evidence that small classes in the primary grades are academically superior to regular size classes. A small class advantage was found for inner-city, urban, suburban, and rural schools; for males and females; and for white and minority students alike. The few significant interactions found each year indicated greater small class advantages for minority or inner-city students. In North Dakota, and specifically in Bismarck, does reduced class size correlate with increased student achievement? By lowering class size in the primary grades, is achievement increasing for Native American students, the largest minority population in the state? By reducing class size, will equity exist among students of color and poverty, as was evidenced in urban setting research on class size?

Without exception, the greatest obstacle to widespread implementation of smaller classes is the expense of additional teachers and classrooms. The cost issue is raised by researchers (Tomlinson, 1990) and by federal, state, and local policymakers. Federal funding to reduce class size has been allocated to public schools since 1999. Current federal legislation will reauthorize funds, as the educational funding allocation took effect July 1, 2002. With the reduction in class size federal legislation, districts will now be allowed to use federal education dollars in many ways. Will the results of lowered class sizes and higher student achievement be compelling enough to continue the effort to reduce student-teacher ratios in the lower elementary grades to 1:18 or 1:20 in the Bismarck Public Schools, or will the money be spent in other areas of education?

If a continued commitment to small class size persists, teachers will need to be hired who fit the teaching and learning style model associated with small class size instruction (Molnar, Zahorik, Ehrle, & Halbach, 2000). Answers to these questions, as provided by the teacher questionnaire, will be vital to the program success in managing resources, providing inservice training, and raising all students' achievement levels.

Short- and long-term benefits, in addition to enhanced performance and academic engagement, have been shown to accrue from small class participation.

The Lasting Benefits Study (LBS), Tennessee's longitudinal study to study Project STAR students as they moved into upper grades, demonstrated that students who had been in small classes were less disruptive than their peers in regular classes (Finn, Fulton, Zaharias, & Nye, 1989). The Success Starts Small project (Achilles, Kiser-Kling, Owen, & Aust, 1994) documented that grade 1 disciplinary referrals dropped over successive years in small classes. Class size research also shows that outcomes associated with small classes are the foundations of safe schools. Smallness promotes familiarity with and knowledge of individuals that can head off violence before it happens (Voelkl, 1996).

A dissertation study was conducted from Project STAR data that focused on pupils who entered kindergarten and grade 1 as retainees (Harvey, 1993). The study concluded that proportionately fewer students were retained in small classes and that pupils in small classes were passed to the next grade with a wider range of scores. If using small class placement became an alternative to grade retention, this would be a major cost saving to school districts.

Ratings of specific engagement dimensions revealed improvements in the expenditures of effort, initiative taking, and reduced disruptive and inattentive behavior in comparison to students in regular classes. With both academic and behavioral advantages, it is possible that small classes could reduce the need for special education placements. This would, of course, represent an important cost savings (Achilles, 1998b).

The past two decades of research on class size have provided possibilities for improving the performance of students in the elementary grades and beyond. Recognizing the potential of small classes, many states, and even the government, have begun initiatives to reduce class size in some or all districts. In North Dakota, small class sizes often exist due to limited local school populations. If lower class size does not increase achievement, should classes be combined? In the larger districts such as Bismarck, will the school board continue to reduce student-teacher ratios in the lower grades in the future? To objectively approach these questions, it is important that the district continue to systematically collect information before,

during, and after class size reduction in order to document student achievement and to make sound economic decisions.

This study has provided information regarding class size and student achievement that relates directly to North Dakota and its student population. It is expected that the data from this study will provide the Bismarck Public Schools with information about elementary student achievement, and the current instructional practices of primary teachers, in order that informed decisions can be made on behalf of current and future students.

#### Delimitations of the Study

The following limitations will be applied to this study:

1. This study will be limited to students in the Bismarck Public Schools, Bismarck, North Dakota. Large and small classes were examined from the 16 elementary schools in the district.
2. The study is limited to three years of statistical information, based on the start of the federal government's Reduced Class Size Initiative, begun in 1999.
3. The results are limited in that they are taken from a single annual reading pre-post test, Gates-MacGinitie Reading Test (Fourth Edition).
4. Some students, due to absences and other uncontrollable factors, submitted incomplete assessments which had to be excluded from the data. Some students exited the school system prior to the posttest while others entered the school system after the pretest. These data were excluded from the final data set. Only test scores from completed tests of students present for both the pretest and posttest were included in the final data set.
5. This study was also limited due to the small sample size of teachers surveyed in the study.

#### Definition of Terms

Class size is the number of students assigned to and enrolled in a specific class under the direction of a specific teacher.

Class Size Reduction (CSR) is the term used when the ratio of teachers to students in a classroom is lowered.



The control group is the group in a research study that receives a different treatment than the experimental group.

ESEA is the Elementary and Secondary Education Act.

The experimental group is the group in a research study that typically receives a new, or novel, treatment, which is under investigation. In this study, the experimental groups are the small class size classroom (less than 17) and the large class size classroom (22 or greater).

Extended Scale Scores (ESS) are a measure of relative position that is appropriate when the test data represent an interval or ratio scale of measurement.

Gates-MacGinitie Reading Test (GMRT) (Fourth Edition), Riverside Publishing, Houghton Mifflin Company, is the reading achievement test used in the Fall and in the Spring to pretest and posttest students in the study.

Lasting Benefits Study (LBS) refers to the longitudinal follow-up of Project STAR in Tennessee that has shown that the achievement benefits for students in grades K-3 small classes are retained years after students leave the small class condition.

Normal Curve Equivalency (NCE) scores are accurate only to the degree to which the distribution is normal.

Pupil-Teacher Ratio (PTR) is the number of children at a site divided by the number of professional educators at the same site.

RAND (a contraction of the term research and development), founded in 1948, is a non-profit institution that helps improve policy and decision making through research and analysis. It is the first organization to be called a "think tank." RAND headquarters are located in Santa Monica, California. (Online: [www.rand.org](http://www.rand.org))

Sigma, the Greek letter "σ," is used to represent the standard deviation (SD) in a population, whereas  $s$  generally indicates that the curve represents the scores of a sample.

In this study, Socioeconomic Status (SES) refers to whether the child does or does not qualify for free or reduced lunch. A student in the public schools can possibly qualify for either free or reduced lunch through the federally subsidized school lunch program. The qualifying guidelines are established by Congress, and a form must be filled out by the parent or guardian and returned to school. The basis for free or reduced lunch is determined by a combination of income and family size.

Student Achievement Guarantee in Education (SAGE) Program was begun statewide in 1996 in Wisconsin to increase the academic achievement of students living in economically disadvantaged circumstances by reducing class size to 15 students per teacher in K-3.

Student Teacher Achievement Ratio (STAR) Project is a four-year longitudinal study of class size. The study was conducted by the Tennessee State Department of Education to determine the effects of reduced pupil-teacher ratio on the achievement of students in public schools.

Total Reading Extended Scale Score (TRESS) is the composite reading score for students on the Gates-MacGinitie Reading Test.

#### Summary and Overview

As a reform measure to improve student achievement, class size reduction has intuitive appeal and mounting support from research. Class size reduction is an initiative that reduces the number of students per teacher in a regular classroom on a daily basis. Parents and teachers like the idea, and policymakers are embracing it. According to the Department of Public Instruction, North Dakota had 187 school districts apply for money from the Class Size Reduction federal grant in 1999, and a few additional districts applied for and received funding the following year. Other states, like California, Florida, Indiana, Nevada, Tennessee, and Wisconsin, launched class size reduction efforts statewide in early elementary grades prior to federal funding (Egelson, Harman, & Achilles, 1996; Viadero, 1998).

Over the past two decades, studies have documented greater achievement gains for students in small classes compared to their peers in larger classes. In particular, members of minority groups and students in socioeconomically

disadvantaged areas do benefit from class size reduction (Achilles & Finn, 2000; Molnar et al., 1999a, 1999b). Results from follow-up studies indicate lasting benefits for students who attended such small class programs as the Student Teacher Achievement Ratio (STAR) Project in Tennessee (Boyd-Zaharias & Pate-Bain, 2000).

Critics of class size reduction say that it is one of the most expensive reforms in education and question whether the benefits are worth the cost. Class size reduction politics compete with other educational reform measures, require a considerable commitment of funds, and influence the availability of qualified teachers (Hruz, 1998). The cost of reducing class sizes and the effects of the reductions on student achievement have received considerable attention, but little has been said about the different ways of implementing class size reduction programs. In North Dakota, assessments have yet to address the issue of student achievement, or the issue of what to expect from class size reduction and how to maximize the benefits of small classes.

This study is divided into five chapters. Chapter I includes an introduction, the statement of the problem, the significance of the study, limitations of the study, definition of terms, and the summary and overview of the study. Chapter II presents a review of the current related literature. Chapter III explains the methodology and design of the study, including the sampling process, the instrument, and process used. Chapter IV provides an analysis of the data. Chapter V includes the findings of the study with conclusions and recommendations.

CHAPTER II  
REVIEW OF LITERATURE IN CLASS SIZE AND  
PRIMARY STUDENTS' ACHIEVEMENT  
IN READING

This chapter reviews literature pertaining to class size and primary students' achievement in reading. It also examines literature related to teachers of small and large class size and the role they play in student achievement. The focus of the first section of this chapter deals with the historical perspective. The second section deals with state class size reduction initiatives. The third section presents state studies that show no correlation between class size and student achievement. The final section of Chapter II summarizes the research related to class size and student achievement at the primary grades in reading.

Historical Perspectives

Class size has always been a concern of education. According to Angist and Lavy (1996), the study and use of class size regarding student achievement began in the 12th century when Maimonides, the great Rabbinic scholar, laid out the principles of class size according to concepts presented in the *Talmud*.

Since the beginning of this century, hundreds of research studies have examined the effect of class size on academic achievement. Until recently, the general findings have been inconclusive. Studies reporting as positive the relationship between small classes and academic achievement well outnumber those supporting the efficacy of large classes, but lack of consensus on this issue has led to the use of contradictory findings to support arguments on all sides of the issue. The literature review focuses on class size from the Glass and Smith (1978) meta-analysis to the present time, or approximately the last 20 to 25 years of study.

Glass and Smith (1978) reported their research by reviewing studies that had been done showing a relationship between reduced class size and elementary pupil achievement. The researchers collected 77 empirical studies that yielded over 700 comparisons of the achievement of smaller and larger classes. The literature search dated back to 1900 and involved more than 900,000 pupils. The results showed that as class size increases, achievement decreases. As an example, the differences in being taught in a class of 20 vs. a class of 40 show an advantage of 6 percentile ranks, with greater gains in classes of 15 or fewer students. Glass and Smith reported achievement differences over 10 percentile ranks in comparisons of groups of 10 and 20 students. Proponents of meta-analysis class size studies stress the fact that small classes (less than 20 students) may be effective primarily because they facilitate other instructional modifications and changes in teacher-student attitudes and behavior which result in improved academic achievement. Glass and Smith suggested in their 1978 original study that 15 was a critical point for class size.

Researchers Smith and Glass soon followed their initial meta-analysis with a second meta-analysis analyzing the relationship between class size and other outcomes. In their 1982 report, Glass and his associates reiterated their findings and noted that of the more than 100 well-controlled comparisons, 81% favored smaller class sizes. They found that small classes (fewer than 20 pupils per classroom teacher) were associated with higher achievement at all grade levels, especially if students were in the small classes for more than 100 hours, and if student assignment was carefully controlled. In their second study, they concluded that small classes were superior in terms of students' reactions, teacher morale, and the quality of the instructional environment (Glass, Cahen, Smith, & Filby, 1982).

The Glass and Smith (1978) paper was followed by two publications from Educational Research Service (ERS) (1978, 1980); by the publication of "An Experimental Study of the Effects of Class Size" (Shapson, Wright, Eason, & Fitzgerald, 1980); the Glass et al. (1982) book; and a book by Cahen, Filby, McCutcheon, and Kyle (1983). Except for the Shapson et al. study, educational progress toward reducing class size came about as a result of analyses of past

studies, a looking backward approach. The interest was driven by analyses of studies years ago and by a growing uneasiness that present-day, poorly researched education practices do not address current problems.

In 1989, Slavin employed a best evidence synthesis strategy to analyze empirical studies that met three specified criteria: (a) A study was included only if class size had been reduced for at least a year; (b) classes of less than 20 students were compared to substantially larger classes; and (c) students in the larger and smaller classes were comparable. Slavin found that of the eight well-designed studies he analyzed, reduced class size had a small, positive effect on students that did not persist after their reduced class experience.

Robinson and Wittebols (1986) published a review of more than 100 relevant research studies using a related cluster analysis approach. Similar kinds of research studies were clustered together, such as studies of the same grade level, subject area, or student characteristics. By focusing on specific problems and issues through isolation of the interaction between class size and relevant achievement factors, cluster analysis provided more useful information about research findings that directly related to areas of concern in evaluating class size policy. They concluded that the clearest evidence of positive effects is in the primary grades, particularly kindergarten through grade 3. The research further showed that by reducing class size, achievement showed promise, especially for the disadvantaged and minority students in the classes. Positive effects were less likely to happen if teachers did not change their instructional methods and classroom procedures in the smaller classes.

Robinson and Wittebols (1986) further concluded that smaller classes result in increased student-teacher contact; reductions in class size to less than 20 students without changes in instructional methods cannot guarantee improved academic success; no single class size is optimal for all grade levels and subjects; classroom management improves in smaller classes; smaller classes result in higher teacher morale and reduced stress; individualization is more likely to occur; and class size

appears to have more influence on student attitudes, attention, interest, and motivation than on academic achievement.

Other research analyses have concluded that class size reduction does not have an appreciable effect. Tomlinson (1988) examined trend data from the 1950s to 1986 in the United States and did not find any consistent relationship between class size and standardized test scores; he concluded that the existing research did not justify a policy to reduce class size, in view of the costs involved and the potential negative impact on the quality of the teaching force. Critics pointed out that this analysis combined students from all grade levels together, that the reliance on student-teacher ratios was an inadequate measure of class size, and that Tomlinson ignored a host of intervening factors and social changes which may have masked the relationship (Achilles, 1996; Finn, 1998).

Odden (1990) reviewed the existing research and argued that a system-wide class reduction policy would produce only modest gains in student achievement and incur an unjustifiable high cost. He opted instead for certain targeted class reduction strategies in conjunction with a series of other interventions, and claimed that his proposals could produce greater benefits with lower costs.

An analysis of the relationship between class size and student achievement in Florida students using 1993-94 school level data found no relationship between smaller classes and student achievement. The study cautioned about drawing conclusions from the analysis based on the limitations of the available data (Florida Department of Education, 1998).

Hanushek (1998) has repeatedly reviewed the available studies that permit a comparison of various school resource inputs--including class size reductions--and student outcomes, and has concluded that reducing class size should not be expected to produce better student performance. His analyses have found that the relationship between various school expenditures--including class size reductions--are remarkably weak, leading him to call for a drastic rethinking of public education policy. Other researchers have used somewhat different analytical techniques to examine the same data and have disputed Hanushek's conclusions,

arguing that the data do show important effects for student achievement, including the influence of smaller classes (Greenwald, Hedges, & Laine, 1996).

Still other researchers have raised questions about the limitations of the basic analytical approach used here because it relies on student-teacher ratios as a measure for class size, it usually groups the data for all grade levels together, and the data represent student achievement at the level of school or school district average scores instead of representing individual students placed in larger or smaller classes (Achilles, 1996; Finn, 1998).

In 1997, Wenglinsky published research findings concerning the relationship between class size and student achievement based on his analysis of data drawn from three national databases. He found that class size served as an important link between school education spending and student achievement at both the fourth and the eighth grade levels. At the fourth grade level, lower student-teacher ratios are positively related to higher achievement. At the eighth grade level, lower student-teacher ratios improve the school social environment, which in turn leads to higher achievement. The largest effects for achievement gains occurred in districts where there were below-average socioeconomic status students, accompanied by above-average teacher costs. Thus, from fairly small beginnings in about 1978 to 1980, it has taken approximately 20 years for class size to be considered seriously, and about 10 years for results of one educational experiment to become relatively widespread in American education.

Finally, in 1998, Congress responded to President Clinton's call for a national initiative to lower class size in the early grades to no more than 18 students. This was a result of the states' research that indicated that classes that are small are effective in helping to improve academic achievement, especially for disadvantaged students. That year, Congress made a bipartisan commitment to provide a down payment on a proposed seven-year phase-in of the Class Size Reduction Program. The fiscal year (FY) 1999 appropriation of \$1.2 billion enabled school districts across the nation to hire an estimated 29,000 new teachers for the 1999-00 school year (U.S. Department of Education, 2000).



In July 2000, the U.S. Department of Education awarded an additional \$1.3 billion in FY 2000 funds to enable states and local school districts to continue their class size reduction efforts. In its first year, 1.7 million young children had the opportunity to learn in smaller, more personalized classrooms. Prior to the implementation of the federal Class Size Reduction Program and similar initiatives in several states, more than 85% of students were in classes with over 18 children and about 33% were in classes of 25 or more students. After one year of implementation, the federal Class Size Reduction Program had reduced the teacher-student ratio from 1:23 to 1:18 in 90,000 classrooms in grades 1-3. Almost one third of the nation's elementary schools, or about 23,000, have hired one or more new teachers. Another 15,000 school districts improved teacher recruitment and hiring or provided professional development to help teachers maximize the benefits of smaller classes. For the year 2001-02, the President's budget proposal asked Congress to provide an additional \$450 million in funding, raising the total to \$1.75 billion. With this funding increase, local communities were able to hire as many as 20,000 additional teachers, for a total of 49,000 teachers hired with Class Size Reduction Program funds (U.S. Department of Education, 2000).

Federal funding for reduced class is distributed to states by a defined formula. All 50 states, the District of Columbia, and Puerto Rico participate in the program. Since needs are greater in the poorest communities, and because research shows that smaller classes provide the greatest benefits to the most disadvantaged students, the program targets funds to high-poverty communities. Each state distributes 80% of the funds to school districts based on the number of poor students in each district. The remaining 20% is distributed on the basis of total enrollment (U.S. Department of Education, 2000).

Class size reduction allocations for North Dakota were \$5,623,097 in 1999, \$6,094,043 in 2000, and was \$8,087,314 for the 2001 school year. In the Bismarck Public Schools, specifically, the entitlement allocation history has been \$411,209 in 1999, \$443,008 in 2000, and was \$556,601 for the 2001 school year. For FY 2002, the Class Size Reduction Program was incorporated into the new Elementary

and Secondary Education Act (ESEA) Title II Teacher Quality block grant. State and local education agencies may use any portion of the nearly \$3 billion in the federal Title II funds to hire qualified teachers to reduce class size (G. Gallagher, personal communication, November 10, 2001; E. Gerhardt, personal communication, November 10, 2001; U.S. Department of Education, 2000).

Class Size Reduction Program funds go directly to the nation's classrooms. Every dollar appropriated by Congress is allocated to local school districts. No funds may be used for federal or state administrative costs. Because small classes make the greatest difference when teachers are well trained, school districts may use up to 25% of the funds for providing professional development to both newly hired and experienced teachers. The remainder of the funds may be used for recruiting and hiring fully qualified regular and special education teachers and teachers of children with special needs, including teachers certified through state and local alternative routes (U.S. Department of Education, 2000).

Because average class size varies considerably from district to district, and often from school to school within a district, districts are encouraged to target program resources to schools with the highest average class sizes and the children most in need of more individualized instruction. The Class Size Reduction Program provides flexibility to accommodate these school districts, as well as the growing number of districts that will reach a class size target of 18 students as a result of the program. Districts that have already reduced a class size in the early grades to 18 students have flexibility. They may use program funds to make further reductions in those grades, to reduce class size in other grades, or to take other steps to improve the quality of teaching in small classes (U.S. Department of Education, 2000).

Since the first year, modifications have been made to make class size reduction even more effective. Some of the modifications include providing districts that receive allocations less than the amount necessary to hire an additional teacher flexibility in the uses of their funds; including kindergarten as one of the early grades; placing even more emphasis on ensuring that teachers hired with program funds are fully qualified; allowing states and districts to substitute state or local class size

reduction goals for the national goal; and inserting new public reporting requirements for states, participation districts, and schools (U.S. Department of Education, 2000).

#### State Class Size Reduction Initiatives

Data from several recent initiatives have added considerably to the research evidence concerning class size reduction in the United States in the early primary grades. Efforts in Indiana, Tennessee, North Carolina, and Wisconsin have reported initial findings with the Tennessee projects currently providing the most complete and well-designed study of class size reduction efforts.

Beginning in 1984, Indiana's PRIME TIME project allocated money (\$19 million) to support the reduction of class size to 18 in first, second, and then kindergarten and third grade classrooms. In proposing the program, Gov. Robert Orr said, "Children spend their first few school years learning to read, and the rest of their lives reading to learn" (Bain & Jacobs, 1990, p. 2). Project PRIME TIME was intended to get Indiana school children off to the best possible start.

Implementation of PRIME TIME was not rigorously controlled, and the results were mixed. It was primarily a project, and not research, although it did have provisions for evaluation. The outcomes of PRIME TIME have been summarized in numerous publications (Center of School Assessment, 1986; Chase, Mueller, & Walden, 1986; Malloy & Gilman, 1989; McGiverin, Gilman, & Tillitski, 1989; Mueller, Chase, & Walden, 1988).

An evaluation of the PRIME TIME project analyzed achievement scores for first and second grade students for the school year 1984-85, comparing mean class scores in reading and mathematics from 10 school districts for tests administered in the first year of the project. In these districts, class sizes ranged from 15 to 22 students. The prior year, 1983-84, class size ranged from 15 to 35 students per teacher. A total of 11,878 scores were obtained from 2,924 students during the years 1983-84 and 1984-85. Student scores on the Iowa Test of Basic Skills, Stanford Achievement Test, Metropolitan Achievement Tests, Gates-MacGinitie Reading Test, Art is Fundamental, and locally prepared basic skills tests were analyzed to determine whether differences existed between classes before and

after the introduction of Project PRIME TIME. Results indicated that, of the 73 statistical test scores analyzed, 39 showed increased student achievement in basic skills, while 30 showed no significant differences. Only 4 test scores showed significant positive differences in favor of larger classes. Tests of student achievement found that for, students in the smaller classes, the reading scores for 1st grade students showed the greatest improvement (Gilman & Antes, 1985).

PRIME TIME did not implement a single, well-defined, small class intervention. While the average class size of 18 pupils was viewed as a target, actual class sizes ranged from 12 to 31; classes of 24 pupils with a teacher aide were considered to be small despite the number of pupils in the classroom. As a result, the evaluations of PRIME TIME could not be interpreted as confirming or refuting a class size effect (Bain & Jacobs, 1990).

Indiana did continue to support reduced class size, spending \$66.5 million in 1987-88 to reduce class sizes to a ratio of 1:18 in grades kindergarten and 1 and to 1:20 in grades 2 and 3. The state reimbursed the local school system \$21,000 for each additional teacher needed to reduce class size. By 1988-89, all 302 Indiana school systems had chosen to participate in PRIME TIME (Bain & Jacobs, 1990).

In 1990, Gilman and Tillitski began four studies to ascertain the long-term effects of reducing class size in primary grades in Indiana. Study one compared the effect of class size reduction on one school system when all students' scores were compared. Study two was a repeated measures cohort study in which the scores of 193 students who had attended reduced-size classes in the same school district for three years were compared to those of a similar number of students who had attended larger classes. Study three examined the effect of reducing the student-teacher ratios in 27 school districts in southwestern Indiana. Study four examined the effect of smaller classes on a statewide basis by comparing the mean scores on the state competency test. The clear and consistent indications of these comparisons of reduced-size and regular size classes suggest that the long-term effects of a state-sponsored class size program are negligible when examining student achievement. It is concluded that, although this study and others have found

no educational benefits for reduced-size classes that can be translated to gains in achievement test scores, the Indiana experience does not necessarily imply that all state class size reduction programs are doomed to fail. Hastily implemented, PRIME TIME was not a well-conceived class size reduction program.

Project PRIME TIME is noteworthy because it demonstrates important principles for the research that followed, namely, the feasibility of a statewide initiative and the need to conduct an intervention of this type over a period of years. Virtually all class size research that preceded PRIME TIME was cross-sectional in nature.

A small experimental study of class size effects in two metro Nashville, Tennessee, schools was conducted 1983 to 1985 (Whittington, Bain, & Achilles, 1985; Bain, Achilles, & Witherspoon-Parks, 1988). The study was initiated by Helen Bain, who had been the president of the National Education Association (NEA) where one of her main interests was to get class sizes to a reasonable level so teachers could teach and children could learn. The 1985 DuPont study was important, as it started Tennessee policymakers thinking about class size research. Results of the DuPont study became available in journal form (Bain, Achilles, Dennis, Parks, & Hooper, 1988; Whittington et al., 1985); and, although small in size, the results had significant impact. These early works helped to build a solid base for a major statewide class size experiment conducted in Tennessee, 1985 to 1989. The cost of this four-year research exceeded \$14 million.

Late in its 1984 session, the Tennessee legislature funded a four-year study of the effects of small classes on the achievement and development of early primary youngsters as part of then-Gov. Lamar Alexander's Better Schools Program. House Bill 544 was to be a definitive, experimental study that would provide the legislature information about class size, with no "maybe" or "it depends" answers. Perhaps to ease financial burdens if small classes should produce positive results, the legislature asked the researchers to examine the efficacy of using a full-time teacher aide or assistant in a regular class. The policymakers in Tennessee wanted

this information as a basis for setting state regulations on class sizes (Tomlinson, 1988, 1990).

The Tennessee study, which started during the 1985-86 school year, had two powerful experimental conditions. The control condition was the "regular" class of 1 teacher to 22-26 students or an average class size of 1:25. The two experimental conditions were 1 teacher in a "small" class of 13-17 students, with an average of 1:15, and a regular class (1:25) with a full-time instructional aide. The regular classes were set so small to assure that a student in one would not have a class larger than the Tennessee class size maximum at that time.

The legislature's mandate of "cause and effect" required the four principal investigators of Project STAR to establish an experimental design using random assignment of students and of teachers. In the parsimonious, but strong "in-school" research plan, each school with one or more of the small classes also had one or more regular class and regular-aide class. The in-school design helped control for building and district differences (Nye, Achilles, Boyd-Zaharias, Fulton, & Wallenhorst, 1992).

Begun in 1985, Tennessee's Project STAR set the stage for asking and answering a number of policy questions that were not addressed previously. Project STAR and two associated data collections have made contributions to the quality of research evidence concerning the reduction of class size. Project STAR was a four-year longitudinal study of kindergarten, first, second, and third grade classrooms in Tennessee. Project STAR compared classes of 13-17 students with classes of 22-26 students both with and without an additional instructional aide in the larger classes. Participating teachers did not receive any professional training focusing on teaching in reduced-size classes. Project STAR was unusual because it possessed essential features of a controlled research experiment design to produce reliable evidence about the effects of reducing class size (Nye et al., 1992).

Project STAR included 79 schools, more than 300 classrooms, and 7,000 students, with students being followed through four years of experience in the given class size. Teachers and students were randomly assigned to the three different

kinds of classes in order to ensure that the study was not biased by who was in which type of class. All participating schools implemented at least one of each of the three types of classes in order to cancel out the possible influences coming from variations in the quality of the participating schools that might affect the quality of the classroom activity (Nye et al., 1992).

The in-school design was an effective way to control for differences among school settings including, but not limited to, the economic status of the student body, per pupil expenditures, and the manner in which schools were administered. The random assignment was monitored carefully by state level evaluators. Both norm-referenced and criterion-referenced achievement data were collected. The norm-referenced tests, based on item-response theory, permitted comparisons of achievement levels from one grade to the next. The design of Project STAR, together with its magnitude and the follow-up research conducted after the four-year period, led Harvard's Frederick Mosteller (1995) to term Project STAR "a controlled experiment which is one of the most important educational investigations ever carried out" (p. 113).

The evidence from student testing in Project STAR demonstrated that the students in the smaller classes outperformed the students in the larger classes, whether or not the larger class teachers had an aide helping them. Project STAR found that smaller class students substantially outperformed larger class students on both standardized (Stanford Achievement Test) and curriculum-based tests (Basic Skills First). This was true for white and minority students in smaller classes, and for smaller class students from inner-city, urban, suburban, and rural schools. These results are similar to the long-range results of the Perry Preschool Project (Barnett, 1985, 1995; Weikert, 1989, 1998).

In the Perry Preschool Project, the positive achievement effect of smaller classes on minority students was double that for majority students initially, and then began to level out as the project continued over time. Another finding was that a smaller proportion of students in the smaller classes were retained in grade, and there was more early identification of students' special educational needs. There

were no significant differences in academic achievement for students in the larger classes with or without an additional instructional aide (Bain & Achilles, 1986).

Of the three conditions--small, regular, and regular with an aide--the small, then the regular, and then the regular with an aide were best in terms of student outcomes. From this and from analyses of other data combined with in-class observations, evidence suggests that a full-time aide in a K-3 classroom does not improve student achievement (Achilles, 1998b). This finding is important because a teacher aide is commonly used for working with youngsters who do not do well. Project STAR was not the first study to show that aides did not help student outcomes (Davidson, Beckett, & Peddicord, 1994). The conclusion from these results is that class size, not teacher aides, influences student outcomes.

Due to the magnitude of the Project STAR longitudinal experiment, the design, and the care with which it was executed, the results were clear. This research leaves no doubt that small classes have an advantage over larger classes in student performance in the early grades (Word et al., 1990).

Two other findings from the Project STAR research were noteworthy. In small classes, teachers identify student learning needs quickly, address these needs, and thus help keep students out of later special education classes. Besides higher test scores, students from smaller classes have far better behavior (as measured by discipline referrals) and far greater participation in school-related activities (clubs and athletics) than do students who started school in larger classes. After the Project STAR positive findings, Tennessee authorized a study to see how long the initial benefits of small classes would persist (Nye et al., 1992).

Subsequent efforts provided important additional evidence on the positive effects of class size reduction. In 1989, the Lasting Benefits Study was started as a longitudinal follow-up study of Project STAR. Finn (1998) and Nye et al. (1992, 1995) began follow-up research to examine whether the effects of the smaller class size experience persisted when students were returned to normal size classes. The study is still ongoing. To date, the research findings include that, in the fourth grade, students from the smaller classes still outperformed the students from the larger



classes in all academic subjects. In fourth grade, students from the smaller classes were also better behaved than the students from the larger classes as measured by student classroom effort, initiative, and disruptiveness. Through eighth grade, a decreasing but still significant higher academic achievement level of the students from the smaller classes persists.

In Project Challenge (1990 to 1993), Tennessee sought to put the Project STAR findings to use by implementing smaller class sizes in 17 of the state's poorest school districts. Beginning in 1990, the state phased in smaller classes at the kindergarten through third grade levels in districts with the lowest per capita income and highest proportion of students in the subsidized school lunch program. The results of this effort were evaluated by examining the effect on the ranking of the school districts according to student performance on a statewide achievement test. An average increase of 5.3 ranks in reading and 6.6 ranks in mathematics in the rankings of Tennessee's 138 (1/2 or a rank of 69 would be considered average) school systems on the Tennessee Comprehensive Assessment Program were recorded. This increase amounted to nearly one quarter (.25) of a standard deviation gain as a result of the Project Challenge effort. The 17 Project Challenge districts moved from near the bottom of school district performance in Tennessee to near the middle in both reading and mathematics for second grade. In addition, in-grade retention of students was reduced in the Project Challenge districts when smaller classes were implemented (Nye et al., 1992).

Krueger and Whitmore (2001), in their long-term follow-up analysis of students who participated in the Tennessee Project STAR experiment, discovered that those students who were assigned to small classes in the primary grades increased their likelihood of taking a college entrance exam, such as the ACT and SAT. The researchers also discovered that, especially among minority students, more minority students took college entrance exams. As these Black students moved to college age, they increased taking the ACT and SAT tests from 31.6% to 41.3%. It was a steeper increase than among white students whose test taking increased from 44.7% to 46.4%.

The continued longitudinal Project STAR reports note that the teen birthrate for those white female students assigned smaller class sizes was one third less than for those in larger classes. The change for Black female students was not statistically significant; but for Black male students, the rate of teen fatherhood dropped by 40% (Krueger, 2001).

Two smaller studies of class size were conducted in North Carolina pursuant to Project STAR. In 1991, educators, citizens, and the school board in Burke County, North Carolina, began a project to reduce the class size to 15 in grade 1, followed by grades 2 and 3 in subsequent years (Achilles, Harman, & Egelson, 1995; Egelson et al., 1996). The Burke County project also included professional development activities covering instruction and assessment, and so the effects are not necessarily simply a function of reducing class size. Evaluation of the initiative has produced the following findings: (a) Compared to a matched group of students in classes that had not been phased into the smaller class initiative, students in the smaller classes outperformed the comparison group in first, second, and third grade on both reading and mathematics achievement tests (quality factor); and (b) based on independent observations of classroom activity, the percentage of classroom time devoted to instruction in the smaller classes increased from 80% to 86% compared to the larger classes, while the percentage of time devoted to non-instructional activities such as discipline decreased from 20% to 14%.

As the longitudinal analyses of the first cohort continue 10 years after the start of small class size in Burke County, North Carolina, the results continue to be positive in academic benefits gained. The academic benefits gained the first through third grade were maintained through fifth grade and continued to be maintained through the seventh grade for the original matched pairs in both reading and math. These positive academic benefits were shown to be maintained four years after returning to larger classrooms (Egelson & Harman, 2000). The full implementation of the small class size initiative in Burke County continues to be supported at a cost of approximately \$2 million a year to hire the additional classroom teachers needed to maintain low class size in grades 1 through 3.

In a related effort, the principal of the Oak Hill Elementary School in the Guilford County, North Carolina, system restructured classes in grades kindergarten through 3 into a small class format (15 students). The initiative was termed Success Starts Small (Achilles et al., 1994; Kiser-Kling, 1995). Oak Hill was fully Chapter I eligible, with 78% of its students in the subsidized lunch program. Matched comparison groups were used in both studies.

The results of both projects favored small classes in academic achievement; small class effect sizes were in the range .4 [Sigma] to .6 [Sigma], in the distribution of class means (Achilles et al., 1994; Achilles et al., 1995). Success Starts Small included systematic comparisons of teaching behavior in small and regular classes. The results, as observational researchers from the Southeastern Regional Vision for Education (SERVE) discovered (1995 to 1999), were that teachers of small classes (less than 18 students per teacher) spent more time on task and less time on discipline or organizational matters compared with teachers of regular sized classes (1 teacher to 24 students). On-task behaviors increased as a percentage of all behaviors between October and April in small classes and decreased over the same time span in the larger classes. Discipline referrals among grade 1 pupils declined in small classes from 38 to 28 to 14 over the four-year period that trained observers studied (Egelson & Harman, 2000).

Both Project STAR and the Lasting Benefits Study provided evidence that small classes in the primary grades are academically more successful than regular size classes. The findings were confirmed for every school subject tested. For example, in "total reading" score at grades K, 1, 2, and 3, small class students were .8, 1.7, 2.7, and 5.4 months ahead, respectively, of those in regular classes. Using grade equivalents, in grades 4, 5, and 7, students were 2.4, 4.8, and 5.8 months ahead. The two sets of grades and scores involve two different tests, but both tests are consistent in showing cumulative gains. Teachers of small classes received no special instructions or training; the outcomes result from class size and from whatever perceptions and advantages accompany having substantially fewer students in a room with one teacher. This is not to say that the effects could not be accentuated if

additional teacher preparation initiatives were provided. For Project STAR, evaluation data included standardized measures of student outcomes and progress, teacher logs, observations in classrooms, student data (attendance, behavior, age, race, sex, and free lunch) (Finn & Achilles, 1998).

A small class advantage was found for inner-city, urban, suburban, and rural schools; for males and females; and for white and minority students alike. Small classes benefit all students, but minority and traditionally hard-to-teach students received approximately twice the benefit from the same investment and treatment (Achilles, 1998b). The few significant interactions found each year indicated greater small class advantages for minority or inner-city students (Bingham, 1994). These studies were based on research suggesting that small class benefits are most likely to occur in the primary grades. The LBS results indicate clearly that the effects carry over into later years. According to Achilles' (1998b) summary of class size Project STAR research, the benefits obtained in K-3 remained with students up through at least grade 9.

In 1996-97, Wisconsin began a class size reduction program called the Student Achievement Guarantee in Education (SAGE) Program. The SAGE Program's objective was to phase in class size reduction in kindergarten through third grade in school districts serving students from low-income families. The SAGE Program was implemented in stages, and its aim was to reduce the class size in the appropriate grade levels to a student-teacher ratio of 15 to 1 or less. In the first annual evaluation of the program, SAGE students' academic learning in first grade classrooms was measured in October 1996 and again in May 1997. The students' scores were compared to those of students in matching comparison schools serving similar populations of students with the following results. SAGE students consistently performed better than comparison students on various areas of the Comprehensive Test of Basic Skills (CTBS). The gap between white and African-American students in achievement did not widen, in contrast to a widening of the gap between white and African-American students in the comparison student groups (Maier, Molnar, Percy, Smith, & Zahorik, 1997).

In this quasi-controlled five-year study, the SAGE evaluation has demonstrated that teachers of smaller classes reported an overall reduction in discipline problems. All SAGE teachers noted increased instructional time, more time for individualization, and more flexibility in choosing among instructional strategies that keep students actively engaged in learning. In addition, SAGE teachers stated that class size reduction increases the likelihood of reaching grade level objectives and covering the content in more depth (Halbach, Ehrle, Zahorik, & Molnar, 2001).

According to the third-year evaluation report of the program, SAGE is fostering an enthusiasm for learning that is boosting student achievement. Results from achievement tests show statistically higher performance for SAGE students across all grade levels when compared to schools with similar characteristics. For example, at the third grade level, using CTBS scores (before they were corrected for pre-existing differences between groups on factors such as prior achievement, attendance, race, and social economic status), there was a mean scale score difference of 8.20 (significant at the .05 level) between SAGE students and comparison students. African-American SAGE students scored lower on a pretest than African-American students in comparison schools but made significantly larger gains and surpassed achievement by African-American students in comparison schools on the posttests. On the CTBS reading test, African-American students had mean scores of 17.55 (significant at the .05 level) higher when in SAGE classrooms. The study is continuing to find that smaller classes provide high levels of classroom efficiency, a positive classroom atmosphere, expanded learning opportunities, and enthusiasm and achievement among both students and teachers (Molnar et al., 2000).

The Wisconsin Department of Public Instruction, with positive evidence of the SAGE Program's success, has prompted the state legislature and the governor to dramatically increase funding to allow 400 to 500 more elementary schools to participate during the 2000-01 school year. To support this expansion, SAGE

funding rose from \$18 million for 1999 to \$58 million just one year later (U.S. Department of Education, 1999).

From the executive summary of the 2000-2001 evaluation results of the SAGE Program, Molnar et al. (2001) reported that SAGE achievement advantage persists. When scores are adjusted for pre-existing differences in socioeconomic status, ethnicity, attendance, and prior knowledge, a SAGE advantage from the beginning of first grade to the end of third grade is shown on all subtests. From the end of second grade to the end of third grade, a SAGE advantage is shown in the third grade reading subtest.

The 2001 SAGE report further states that adding students lowers the average performance of classrooms. Each student added to a classroom beyond the 15:1 SAGE student-teacher ratio results in a decrease of approximately one scale score point in the class average in all academic scores. These results were taken from 1,542 students in 93 classrooms (Molnar, Smith, & Zahorik, 2001).

Large scale efforts to reduce class size have not been limited to Indiana, Wisconsin, and Tennessee. Some states initiated targeted class size reduction policies some time ago, while others are only in the early stages of development and implementation. As of September 2000, the U.S. Department of Education recognized over 20 states across the country as having instituted their own efforts to lower class size. Despite the number of states now enacting Class Size Reduction (CSR) policies, very few have evaluated those policies' impact. Georgia and Massachusetts began just recently; other states have been investing resources for more than a decade (U.S. Department of Education, 2000).

In 1984, Texas passed legislation requiring class size to be limited to 22 students in kindergarten through fourth grade, with the provision going into effect for kindergarten through second grade in 1985-86 and for third and fourth grades in 1988-89 (Texas Education Agency, 1998).

Positive conclusions have been drawn from an analysis of a substantial database about the Texas educational system. Using data from more than 800 districts containing more than 2.4 million students, Ferguson (1991) found significant

relationships among teacher quality, class size, and student achievement. For first through seventh grade, using student-teacher ratio as a measure of class size, Ferguson found that district student achievement fell as the student-teacher ratio increased for every student above an 18 to 1 ratio. Measures of teacher quality, as measured by teacher literacy skills and professional experience, were even more strongly related to higher student scores.

In Austin, Texas, achievement and attendance have remained extremely low at 13 of 15 low-performing schools, while the other 2 schools showed dramatic gains. Those two schools combined CSR with other changes such as new curricula and teaching methods, increased parent involvement, and health services (Murnane & Levy, 1996).

Nevada began a class size reduction program in 1990-91, beginning with a target of a 15 to 1 student-teacher ratio for kindergarten and first grade, then applying that ratio in second grade and third grade, to be followed by efforts to reduce the ratio to 22 to 1 for fourth through sixth grade, and then 25 to 1 for seventh through twelfth grade (Sturm, 1997). Nevada's limited evaluations have been inconclusive. Though researchers recently found evidence of a differential, positive effect on the achievement of English language learners, achievement gains generally have been disappointing and evaluation has not been comprehensive enough to indicate why (Snow & LaMarca, 2001).

Utah has funded class size reduction since 1990, including some targeting of low-income students and flexibility in how districts and schools use the money. A 1997 study of five districts found that the most successful school combined CSR with teacher development, instructional improvement, and productive use of personnel and resources (Evans-Stout et al., 1997).

In 1995, Virginia began an effort to reduce class size in kindergarten through third grade classes for at-risk students, using a strategy in which local systems that devote funds to the voluntary program may receive matching funds from the state (Egelson et al., 1996). This state's legislative effort to reduce class size in K-3 classrooms, with high or moderate concentrations of at-risk students, was in response

to a 1992 study of 31 elementary schools in Fairfax, Virginia. The Fairfax study found that first graders who had been placed in smaller classes (average size 15) had a 75% passing rate in second grade, compared to 54% of those who had been in larger classes (average size 22) (ECS Information Clearinghouse, 1999).

Minnesota began its statewide program in 1993. In the latest two-year budget cycle, \$100 million was allotted to reduce class size, about \$50 million each of the 1999 and 2001 school years. The program emphasizes kindergarten and first grade, with a class size goal of 17 students. Although a few districts hired teachers for the fourth and fifth grades, the vast majority--95%--focused on kindergarten through third grades. As of FY 2003, CSR revenue funds additional teaching staff only in grades K-3 (ECS Information Clearinghouse, 2002). In an e-mail correspondence, Matthew Mohs, Federal Education Programs and Policy Specialist for the Department of Children, Families & Learning, stated that, as of summer 2002, no evaluation has been conducted of the state program to assess its value.

The California Senate Bill 804 chaptered "Class Size Reduction" on August 18, 1997. The program was a response to the continuing poor performance of California students. Size alone gave the initiative significance. With a fiscal year price tag of over \$1 billion, or \$800 for every participating K-3 student, it represented by far the largest educational reform in the history of California or any other state. California established its statewide class reduction program beginning with the 1997-98 school year. Although participation was not mandatory, over 95% of California's districts took part, attesting to the popularity of the initiative. Student achievements in the state's largest school districts improved almost 20% after only one year of class size reductions (Mazzoni, 1998).

An ongoing study of the California program is showing that smaller classes have increased student achievement in communities across the state for the second year in a row (Stecher & Bohrnstedt, 2000). Children throughout California, regardless of their socioeconomic background, race, or ethnicity, are benefiting from being in smaller classes. California's initiative has been followed closely and



coverage of it has appeared in general publications, such as Education Week, U.S. News and World Report, and Time.

Evaluation findings in the 1998-99 report showed that third grade students in smaller classes performed better on achievement tests than third graders in larger classes for the second year in a row. These achievement gains persisted after the students returned to larger classes in fourth grade. The 1998-99 results were obtained from 1.8 million students in 92,000 classrooms (K-3). Over 92% of California students in grades K-3 were in classes of 20 or smaller, and only 9 districts in the state were not participating in the initiative. It should be noted that these results occurred even though the percentage of fully certified teachers in grades K-3 was at 87%, as compared with 98% in 1995 (Stecher & Bohrnstedt, 2000).

In the third of four planned reports released February 2002, the Class Size Reduction Research Consortium finds mixed results for the five-year-old program, which makes classes smaller in grades K-3 and is the largest of its kind in the nation. The study found that, for the first time, school districts with large enrollments of poor and minority students and English learners were just as likely to offer small classes. But for many districts, making classes smaller means sacrifices. Two thirds of districts reported taking money from priorities such as libraries, arts programs, and professional development to cover the cost of reducing class sizes. More than 1 in 5 K-3 teachers were not fully certified at schools where at least 30% of students live in poverty, compared with fewer than 5% of teachers at schools where less than 7.5% of students are poor (Council of State Governments, 2002).

There is no clear link showing the \$6 billion California program to reduce class size improved achievement in elementary schools, according to a recently released report by a consortium of research organizations, including RAND and the American Institute for Research. The latest study of the five-year program showed no clear correlation between class downsizing and academic improvement. Brian Stecher, a senior social scientist, attributes the lack of conclusive evidence to the fact that the state launched so many new educational reform programs at the same time that it is difficult to separate out their effects from the downsizing. Average state test scores

have risen each year for elementary students since 1997, but no strong relationship can be inferred between the improvement and the downsizing of classes. The study goes on to say that it is difficult to say how much of the gain in achievement from test scores is real and how much reflects inflation in scores brought about by teachers learning to “teach to” the test (Stecher & Bohrnstedt, 2000).

Due to the achievement gains correlated to lower class size, other states have joined the effort to implement programs. In 1999, Iowa created the Class Size/Early Intervention Program to reduce class size in kindergarten through third grade to 17 students for basic skills instruction. The overall aim was to provide improvement in reading instruction. The state will phase in the program over four years, allocating \$10 million in the first year, \$20 million in the second, \$30 million in the third, and at least \$30 million in the fourth. The Iowa allocation formula targets low-income districts (ECS Information Clearinghouse, 2002).

After 10 consecutive years of decline in elementary reading comprehension scores, the scores for 2001 showed a slight increase. Student achievement is based on student proficiency levels rather than comparison of test scores or grade level scores. The proficiency levels for fourth grade reading showed that comprehension had gone from 67.7% of the state’s fourth graders reading at or above the proficiency level in 2000 to 67.8% of the state’s fourth graders reading at or above the proficiency level in the year 2001. The gains are modest, but the state is hoping that this is the beginning of continued improvement. The state of Iowa credits the slight improvement on reduced class sizes and targeted literacy initiatives. The state of Iowa still had not reached its initial goal of 17 students per teacher in the primary grades as of the school year 2001-02. The current average class size for the past school year was 18.6 in kindergarten, 18.5 in first grade, 19.4 in second grade, and 20.4 in third grade (Iowa Department of Education, 2002).

Maryland established the Maryland Learning Success Program in 1999, an initiative to reduce class size in grades 1 and 2, particularly in reading, to 20 students. The program, which will be phased in over four years, requires school systems to set specific performance targets and establishes a goal of hiring approximately

1,000 teachers, while reserving additional funds for professional development, supplies, and other implementation costs. The 1999-00 funding was \$11.6 million and was appropriated if, in the opinion of the state superintendent, the plan meets conditions prescribed by the legislature (ECS Information Clearinghouse, 2002).

Due to the extensive reading program that has been developed in Maryland by the State Task Force on Reading, it would be impossible to make a correlation between class size reduction and student performance in the primary grades. Maryland students have shown slow progress in reading achievement. According to National Assessment of Educational Progress (NAEP) scores, Maryland student achievement in reading is slightly below the national average. (Maryland mean was 211, while the national average was 215.) On the CTBS, Maryland students score near the national average. (In 1995, Maryland students in grade 3 were reading at the 53rd percentile, while students in grade 5 were reading at the 48th percentile.) Current efforts are to have students read with fluency, comprehend, integrate, and critically evaluate what they have read. Maryland is currently assessing reading on a yearly basis (Maryland State Task Force on Reading, 2002).

Also in 1999, the state of New York began implementing its class size reduction program, which targets funds for reducing average class size in kindergarten through third grade to 20 students. Funded at \$75 million the first year, the program will be phased in over three years, with the second-year funding expected to be \$150 million, and the third-year funding at \$225 million. Funds may be used for teacher salaries and benefits, as well as for one-time start-up costs for each new classroom; however, funds may not be used for new buildings or professional development. The state targets funds to school districts according to enrollment (Haimson, 2000).

The Independent Educational Priorities Panel completed a study of the first year of the class size reduction program in New York City. Among improvements reported as a result of smaller classes were noticeable declines in the number of disciplinary referrals, improved teacher morale, a focus on prevention rather than remediation, and higher levels in classroom participation by students. The study

further noted that, while it was still too early to make definitive judgments, students placed in smaller classes appeared to be learning faster than when they were in larger classes. The board of education hired only certified teachers, not paraprofessionals, in line with research that shows that the educational benefits of pairing a paraprofessional with a teacher in regular size classes are negligible (Haimson, 2000).

Since 1994, Michigan has funded a pilot program in the city of Flint. It has cost the state approximately \$6 million. The results are significant, with 43% more fourth graders passing the state reading test and 18% more passing the state math test. In 1998, Sen. Joe Conroy helped to develop the next step for Michigan--a \$20 million program for statewide implementation of class size in the neediest districts. In Michigan, the state provides 75% of the funds for poor districts that reduce classes to an average of 17 students with a maximum of 19 (Bell, 1998).

#### State Studies Showing No Correlation Between Class Size and Student Achievement

In 2000, Hoxby studied the effects of class size on student achievement by utilizing population variation techniques. Every school district in Connecticut was surveyed about its maximum and minimum class size rules, teachers' aides, and mixed-grade classes. Both maximum and minimum class size rules varied among the districts, but the modal maximum class size was 25 and the modal minimum class size was 15. All of the data used were obtained from the Connecticut Department of Education or its publications.

Hoxby (2000) conducted her study by using natural variation in the school aged population to identify the effects of class size on student achievement. The approach has three benefits. First, the variation in class size is exogenous. It is not variation generated by parents' choices, choices that are affected by parents' incomes and parents' assessments of the attention their children need. Second, the participants in the natural experiment are not aware of being evaluated or mindful of rewards being contingent upon the outcome. Real policies that reduce class size, such as the 1996 California initiatives and the 1999 federal initiatives, rarely include

evaluations or repercussions (such as funds being taken away if the policy has no effect). Third, natural population variation generates fluctuations in class size that are in the range relevant to current policy.

The Connecticut study demonstrated how population variation can be used to consistently estimate the effect of class size on student achievement. Two independent methods were used. The first method is based on isolating the credibly random component of the natural variation in population for a grade in a school. Random variations in the population generate exogenous variation in class size. The second method is based on exploiting the discontinuous changes in class size that occur when a small change in enrollment triggers a maximum or minimum class size rule and thereby changes the number of classes in a grade in a school. Both methods produce results that are appropriate for considering class size changes in the range of 10 to 30 students. Using both methods, it was found that reductions in class size had no effect on student achievement. The estimates were sufficiently precise so that, if a 10% reduction in class size improved achievement by just 2% to 4% of a standard deviation, statistically significant effects would have been found in achievement. The study also found that there was no evidence that class size reductions are more efficacious in schools that contained high concentrations of low-income students or African-American students. Due to the fact that the results described are not likely to suffer from exogenous bias generated by parents', teachers', administrators', or policymakers' decisions, the results can be trusted to a greater degree, as these evaluations are not tied to incentives. These methods also have the advantage that participants are not aware of being evaluated (Hoxby, 2000).

In interpreting the results of the Connecticut study, the two identification methods are independent and thus provide checks on one another. The results are also sensitive to specification changes. The estimates are based on variations in class size that occur mainly in the range of 10 to 30 students per class. The author points out that it would be a mistake to extrapolate these results to schools in which class size is typically higher than 30, as in many foreign countries, or to extrapolate

these results to class sizes of less than 10. Such tiny classes are too expensive for most American districts to consider because the cost of a one-student reduction increases as class size gets smaller. A five-student reduction from a base of 40 raises costs by 14.3%, but a five-student reduction from a base of 15 raises costs by 50% (Hoxby, 2000).

In a report to the North Carolina Department of Public Instruction, the effectiveness of the Wake County Public School System's plan to reduce class size was evaluated to assess program implementation and the effects of class size reduction on academic achievement. For the 1999-00 school year, an allocation was made to the Wake County Public School System of approximately \$1.1 million. The objective was approached by hiring as many fully qualified teachers as possible, establishing implementation models, and determining the grade levels to target. Twenty-three teachers were supported by class size reduction funds, and they were sent to 23 schools where between 21.6% and 51.1% of students received free or reduced price lunches and between 50 and 117 students per school were considered low achieving. District staff developed four implementation models and schools were used to implement class size reduction in grades 1 or 2, with the preferred model being the introduction of a new class of about equal size to other classes in the target grade. Reduced class sizes affected about 2,473 students. Students did show improvement in academic achievement, with improved growth greatest where class size was smallest. However, low-income students appeared to benefit less academically from class size reduction, even though their achievement improved to some extent (Scudder, 2000).

In another study of a reduced class size program, grades 1 and 2 in Saginaw, Michigan, were evaluated after the 1999-00 school year was complete. At grade 1, 23 rooms of reduced-size classes were maintained at no more than 18 pupils; at grade 2, 5 rooms were limited to 21 pupils. Comparison classes were identified to assess the impact of the reduced class size program. Students in both conditions were included in analyses of reading and math achievement and rates of special education placement, attendance, and promotion were reported. At grade 1, no

significant differences between the groups were found at the beginning or end of the year in reading or math. Similarly, no differences were found in any of the above rates. At grade 2, students in reduced-size classes evidenced greater performance in reading at the end of the year; however, no pretest was conducted on text leveling, as a comparison. As with grade 1, there were no significant differences on other measures (Kurecka & Claus, 2000).

In California, there is no clear link showing that the \$6 billion program to reduce class size has improved achievement in elementary schools. According to a report released by a consortium of research organizations, including RAND and the American Institute for Research, the latest study of the five-year-old program showed no clear correlation between class downsizing and academic improvement. Because the state of California launched so many new educational reforms at the same time, it is difficult to separate out their effects from the downsizing. Average state test scores have risen since 1997, but no strong correlation can be inferred between the achievement improvement and the downsizing of classes. School districts serving most of the state's historically disadvantaged students, those who are minorities, those from low-income families, and English language deficient students, have received fewer benefits through class size reductions. These districts found it more expensive to implement CSR, they saw a disproportionate decline in their average teacher qualifications, and they were forced to take more facilities and resources from other programs to create additional classroom space (Stecher, Bohrnstedt, Kirst, McRobbie, & Williams, 2001).

E. A. Hanushek (2000), professor of economics and public policy at the University of Rochester, provides evidence to politics of the class size debate. He finds three missing elements when policymakers look at class size reduction and student achievement. First, nothing in the current decision process encourages targeting class size reductions to situations where they are effective. Second, class size reductions necessarily involve hiring more teachers, and teacher quality is more important than class size in affecting student outcomes. Third, class size reduction is

very expensive, and little or no consideration is given to alternative and more productive uses of resources.

When student-teacher ratios are analyzed, the results are that, throughout the 20th century, they show a dramatic decline. Between the years of 1960 to 1995, the student-teacher ratios fell by one third, exceeding the magnitude of policy changes that most states are looking at implementing today. From an historical perspective, it would seem that, due to the drastic drop in student-teacher ratios, student performance would have increased dramatically. It is impossible to detect any overall beneficial effects that are related to these sustained increases in teacher intensity (Hanushek, 2000). Hanushek goes on to look at Scholastic Aptitude Test (SAT) scores. When achievement is compared to the lowered student-teacher ratios that have taken place, the correlation goes in the opposite direction expected: Reductions in pupil-teacher ratios are accompanied by falls in the SAT, even when appropriately logged for the history of schooling experience for each cohort of students analyzed. Because the SAT is a voluntary test taken by a select population, a portion of the fall reflects changes in the test taking population instead of real declines in aggregate student performance; but there is general consensus that real declines also occurred (Congressional Budget Office, 1986).

Hanushek (2000) also examined the National Assessment of Educational Progress (NAEP) for indication of student performance. He found that math and reading showed flat performance from earliest testing through 1996, while the comparable science and writing scores have declined. Hanushek (1999a) concluded that the consistent picture from available evidence is that the falling pupil-teacher ratios and commensurately increasing real spending per pupil have not had a discernible effect on student achievement.

Hanushek (2000) reviewed 277 studies that examined the impact of student-teacher ratios on learning and found that (a) only 15% of the studies showed that a lower ratio caused a significantly positive impact on performance, (b) 13% of the studies actually showed a negative effect, and (c) the remaining 72% yielded no conclusive results. The statistically insignificant estimates (those for which there was



less confidence that they indicated any real relationship) were almost evenly split between beneficial and adverse effects. Thus, the overall evidence provided little reason to believe that a general policy of class size reduction would improve student performance.

Because the Tennessee Project STAR reports were not included in the 277 estimates that Hanushek analyzed for correlations between class size and student achievement, he analyzed the Project STAR reports separately. What he expected to find was that the differences in performance would become wider through the grades because the students continued to get more resources (smaller classes) and these resources should keep producing a growing achievement advantage. What Hanushek found was that the small class size advantage is almost exclusively obtained in the first year of being in a small class, which would suggest that the advantages of small classes are not general across all grades. The gains in performance from the experimental education in class size were relatively small (less than .2 standard deviation of test performance), especially in the context of class size reduction (around eight students per class) (Hanushek, 2000).

Hanushek (1999b) casts further uncertainty on the positive Project STAR results due to the uncertainty about the quality of randomization in the experiment. Of the initial experimental group starting in kindergarten, 48% remained in the experiment for the entire four years. How were the replacement students chosen? A second question raised was the choice of teachers in the experiment. While they were to be randomly assigned to treatment groups, there is little description of how this was done. In addition, all teachers in the study knew they were participating in an experiment that could potentially affect the future resources available from the state of Tennessee. The schools were self-selected, not randomly selected. Small schools were excluded from the study, and all participating schools were willing to provide their own partial funding to cover the full costs. As a result, the Project STAR experiment heavily over sampled urban and minority schools where the achievement response to the program is thought to be the largest.

Folger (1989) discovered similar results when he critically analyzed the Project STAR results. He set forth the following five items as lessons learned from Project STAR. First, the maximum effect of reducing class size is on kindergarten and first grade. The effect on achievement levels off and declines in second and third grade even when students remain in smaller classes. Second, the achievement advantage of small class students drops about 50% the first year after they returned to regular sized classes (21-28 students) in the fourth grade. Third, class size reduction appears to be very expensive. The cost of reducing class size is proportional to the size of the reduction (i.e., a one third reduction in class size will increase per pupil costs about one third). Fourth, the high cost of substantial reduction in class size and the most achievement gains that can be expected, even in kindergarten and first grade, suggest that less expensive targeted reductions should be tried. Finally, the most important lesson learned from the Project STAR experiment may be that just changing class size without changing what is taught or how it is taught will probably have modest results, because of all the various factors that influence achievement.

Hruz (2000) further researched the cost and benefits of smaller classes in Wisconsin in much the same way that Hanushek and Folger evaluated the Project STAR and other class size reduction experiments. In his policy research report, Hruz also found that only the positive effects of the program had been disseminated to the public, while the more ambiguous results revealing only minor effects from smaller classes had been suppressed.

When it comes to improving academic achievement, class size reductions achieved through the SAGE Program have not been as significant as is commonly argued and assumed. Hruz (2000) points to the following examples taken from the SAGE Program's annual evaluations that have been conducted to determine the qualitative and quantitative effects of the program, particularly on student achievement. First, he points out that smaller classes in the second and third grades had a minimal impact and, in some cases, had no additional impact on student achievement. The available data reveal that while greater gains are consistently

made by students in smaller classes than students in regular sized classes in the first grade, no such consistent advantage is found from being in smaller classes in the second and third grade. Second, it was found that African-American students in smaller second and third grade classes did not gain relative to their gains made in the first grade or relative to African-American students in regular sized classrooms. This fact seems remarkable given the evidence that African-American students show by far the greatest achievement gains from being in smaller classes in the first grade. Third is the fact that smaller classes appear to not have any effect on students who are not African-American, who constitute the majority of students in the SAGE Program. Fourth, the actual magnitude of the gains experienced by students in the SAGE Program is, on average, relatively meager. On average, students in the SAGE Program are scoring only about 1.5% to 5% higher on tests, depending on grade level and subject. Finally, the data do not separate out findings for Milwaukee and non-Milwaukee public schools; nor did they look directly at the effects of smaller classes by income level. The question remains with the SAGE Program study: "Are aggregate gains made by students being driven solely by students in the Milwaukee Public Schools" (Hruz, 2000, p. 2)?

In addition to the above SAGE Program findings, Hruz (2000) found some possible issues with the evaluation process of the SAGE Program evaluation that are common problems in social science experiments. The Hawthorne effect may have been an issue, as teachers may be inclined to work harder to ensure the program's success, and selection bias may also have effected the results of the SAGE Program. A final concern with the SAGE Program design is why the comparison is made between SAGE Program and non-SAGE Program schools and not between SAGE Program and non-SAGE Program classrooms. Such a procedure would diminish the extent to which between-school factors affect the statistical results. This procedure was followed in the Project STAR evaluation and was one of the most lauded design features of that study.

G. E. Robinson (1990), former president and director of research at Educational Research Service, a non-profit organization that provides objective

research and information on education issues, performed a meta-analysis of the class size research and stated that research does not support the expectation that smaller classes will of themselves result in greater academic gains for students.

A. Odden (1990), professor of educational administration at the University of Wisconsin-Madison, has stated that smaller classes should be used sparingly and strategically and that there are more cost-effective means available to achieve the results of smaller classes, without requiring large amounts of new funds. Odden reviewed data on programs in Tennessee and Indiana, and concluded that these studies show that new and costly state programs that reduce class size to under 20 students do not produce very large gains in student performance.

Brewer (1999), a researcher at the RAND Corporation, estimated the costs of different types of national class size reduction policies and made comparisons of these costs to other educational programs. He concluded that the high monetary costs and probable implementation problems associated with a national class size reduction program suggest a reconsideration of its likely benefits and that reducing class size to 15 students costs twice as much as reductions down to 18.

Johnson (2000), an analyst for The Heritage Center for Policy Analysis, a Washington, DC, based think tank, performed a statistical comparison of performance in 1998 NAEP reading achievement scores between smaller classes (less than 20 students) and larger classes (greater than 30 students). After controlling for income, family background, and other demographics, he found that fourth and eighth grade students in the smaller classes did no better than students in larger classes.

The National Conference of State Legislatures (1998), a non-partisan organization that provides information to all 50 state legislatures, concluded that, although over 1,100 studies examined the relationship between class size and student achievement, no definitive conclusions have been reached. While positive results have been demonstrated in Tennessee and Wisconsin, other research finds little connection between student-teacher ratios and student performance, especially when measured against other types of educational reforms.

Despite the political popularity of overall class size reduction, the scientific support of such policies is weak to nonexistent. The existing evidence suggests that any effects of overall class size reduction policies will be small and expensive. A number of investigations appear to show some effect of class size on achievement for specific groups or circumstances, but the estimated effects are invariably small and insufficient to support any broad reduction policies. Proposed class size reduction policies generally leave little room for localities to decide when and where reduction would be beneficial or detrimental. The existing evidence does not say that class size reductions are never worthwhile and that they should never be taken. It does say that uniform across-the-board policies are unlikely to be effective. A significant problem is that there are few incentives that drive decisions toward ones that improve student performance (Hanushek, 2000).

Class size reduction policies should be made in an informed and efficient manner, such that the public investment in the policy results in a meaningful improvement in education. The data results that show less favorable results of class size to student achievement have not been well disseminated to the public or to government officials who have greatly expanded the program in recent years. Such results suggest that more limited implementation of smaller classes, in only the first grade and in only high-poverty schools, can produce nearly the same results presently experienced, but at far less cost. Funds expended to meet these class size reductions may be much more efficiently used for other programs that help the same students aided by smaller classes, or to improve such educational factors as teacher quality and experience, which have regularly been shown to have a greater impact on student achievement, whether in small or regular sized classes. Spending \$100 million or more a year, as in Wisconsin, to enable classrooms to have more of a family-like atmosphere, without any significant increase in student achievement, will simply not pass the test (Hruz, 2000).

### Summary

Most of the recent research has indicated that smaller class size increases student achievement. Teachers, principals, and parents consistently welcome

proposals to reduce class size. On being assigned to smaller classes, teachers report that the classroom atmosphere enables them to have more flexibility to use different instructional approaches and assignments (Kirkbrush, 1996; Maier et al., 1997; Mosteller, 1995).

Tinbergen (1952), Calhoun (1962), and Hall (1966, 1976) have shown the power of crowding to change behavior in negative ways. In small classes crowding is reduced. Researchers have even suggested that gang behavior may start in crowded, early primary settings.

Class size research shows that outcomes associated with small classes correlate with safe schools: improved student behavior; increased sense of community and family in small classes; and a generally improved school climate where teachers, students, and parents feel less stress than in larger classes and larger schools. Smallness promotes familiarity with and knowledge of individuals that can head off violence before it happens (Klonsky, 2002). The reduction of class size itself changes the classroom situation. There are fewer students to distract each other. Each student in a reduced size class gets more attention on average from the teacher and more time to speak while the others listen (Mitchell, Carson, & Badarak, 1989).

Researchers also have suggested that smaller classes are more likely to be friendlier places, where students develop better relationships with their classmates and with the teacher, encouraging students to become more engaged in classroom learning activities. The smaller the class, the harder it is to escape the positive influence of the classroom educational experience. The research finding that reduced class size is particularly beneficial in the early grades may result from the fact that, in the early grades, children are learning how to be students in classrooms where the number of people is larger than the number of people in their families, and students are learning a new routine (Finn, 1998).

The focus on the early grades also suggests that smaller classes represent a preventive, rather than a remedial, approach. If smaller classes help students start off on the right foot in learning how to adjust to the classroom situation and get

engaged in learning activities, then students avoid the more difficult educational path of falling behind, attracting the appropriate assistance, and catching up to their schoolmates. The research evidence from Project STAR (1985) showed that students in smaller classes (with fewer than 18 students) did better when compared with students in larger classes. Given the variations among individual students and teachers and the way they interact, it is unlikely that there is a single “magic number” below which class size suddenly produces a beneficial effect. But it is fairly evident that class size must be somewhere below 20 in order to make a real difference (Nye et al., 1992).

Reducing the ratio of students to teachers does not necessarily mean a reduction in class size. This issue was a complicating factor in the research studies described earlier, where questions were raised about the adequacy of using the student-teacher ratio as a measure of class size. Some initiatives permit officials to include other education staff besides the classroom teacher in the calculation of the ratio, such as resource teachers in special education, music, and physical education. Consequently, school systems could increase the number of teachers without necessarily reducing class size, and particularly since the number of available classrooms is both a practical and a budget issue, officials may be tempted to solve the ratio problem by adding a second teacher to a larger class. The research findings from Project STAR (1985) are relevant here: The larger classes with instructional aides did not produce the same benefits as the smaller classes (Nye et al., 1992).

School arrangements that reduce class size only for particular students or subjects may achieve greater results with lower costs, depending on how they are organized and what exactly makes the smaller class experience better. It may be more important to reduce class size for reading than for physical education, and the research suggests that minority and economically disadvantaged students benefit most from smaller classes. The class size research helps address the concern that “very little research addresses how schools might organize teaching resources more effectively at the school level” (Miles & Darling-Hammond, 1998, p. 9).

With both academic and behavioral advantages, it is possible that small classes could reduce the need for special education placements. If this were true, it would represent an important cost savings. In the CSR research and evaluations done after five years in California, they have discovered that special education identification or placement has not changed. On the one hand, smaller classes afford teachers more opportunity to observe student behavior, which might lead to increased referrals to special education; and, on the other hand, smaller classes provide more opportunities for teachers to address individual differences, which might lead to fewer special education referrals (Stecher & Bohrnstedt, 2002).

Why has reducing class size had a more positive effect in the early grades? Theories about this phenomenon have fallen largely into two categories. First, most theorists focus on the teacher, reasoning that small classes work their magic because the small class context improves interactions between the teacher and individual students. In the early grades, students first learn the rules of standard classroom culture and form ideas about whether they can cope with education. Many students have difficulty with these tasks, and interactions with a teacher on a one-to-one basis, a process more likely to take place when the class is small, help the students cope. In addition, teachers in small classes have higher morale, which enables them to provide a more supportive environment for initial student learning. This theory might show why students who come from impoverished homes, ethnic groups who have suffered from discrimination or are unfamiliar with United States classroom culture, tend to be helped more by a reduction in class size (Biddle & Berliner, 2002).

Biddle and Berliner (2002) described a second group of theories designed to account for class size effects. These theories focus on the classroom environment and student conduct rather than on the teacher. Discipline and classroom management problems interfere with subject matter instruction. Theorists in this group argue that these problems are less evident in small classes and that students in small classes are more likely to be engaged in learning. Teacher stress is reduced in small classes, so teachers in the small class context can provide more support for student learning. Studies have also found that small instructional groups can provide



an environment for learning that is quite different from that of the large classroom. Small instructional groups can create supportive contexts where learning is less competitive and students are encouraged to form supportive relationships with one another.

These two theories are not mutually exclusive. Both may provide partial insights into what happens in small classes and why small class environments help students. In spite of the theories and the evidence on achievement, there is still considerable debate on adopting small classes at the primary level.

Uses of Project STAR and other findings have generated predictable controversy in the literature and among researchers, politicians, and policy folks such as Burtless (1996), Card and Krueger (1996), Hanushek (1995, 1996), Hedges et al. (1994), and Hedges and Greenwald (1996). A recent wave of added interest in the economics of class size processes and outcomes is evident in the work of Angist and Lavy (1996), Boozer and Rouse (1995), Correa (1993), Krueger (1997), and Wenglinsky (1997). These researchers indicate that there may be more efficient ways to improve student achievement. There are also claims about the lack of efficiency of reducing class sizes in the early grades.

One study that identifies the effects of class size on student achievement by using longitudinal variation in the population associated with each grade in 649 elementary schools in Connecticut indicated that class size does not have a statistically significant effect on student achievement. The estimates from this extensive study indicate that class size does not even have modest effects (2% to 4% of a standard deviation in scores for a 10% reduction in class size) (Hoxby, 2000).

University of Rochester economist, Eric Hanushek, examined 277 separate published studies on the effect of teacher-pupil ratios and class size averages on student achievement in 1997. In his literature summaries (1986, 1996, 1998), Hanushek concludes that there is no strong or consistent relationship between school inputs and student performance. Only 15% suggested that there was a statistically significant improvement in achievement, 72% found no effect at all, and 13% found

that reducing class size had a negative effect on achievement (Council of State Governments, 2000).

California's CSR reform has had both gains and unanticipated losses, as monitored by the CSR Research Consortium, headed by the American Institute for Research and RAND; it also included Policy Analysis for California Education (PACE), WestEd, and EdSource. The data suggest that CSR is having positive effects on parent attitudes and student achievement. The gains after the first few years have come at a substantial cost in terms of equity. School districts serving most of the state's historically disadvantaged students--those who are minorities, those from low-income families, and English language deficient students--have received fewer benefits and may even have been hurt by CSR. These districts found it more expensive to implement CSR, they saw a disproportionate decline in their average teacher qualifications, and they were forced to take more facilities and resources from other programs to create additional classroom space. The latest study of the five-year program shows that there is no clear link showing the \$6 billion California program to reduce class size improved student achievement in elementary schools. Average test scores have risen each year since 1997; but, due to the many initiatives that were started so quickly, they cannot necessarily be attributed solely to class size reduction (Stecher & Bohrnstedt, 2000, 2002).

Wasley (2002) summarizes the class size issue in the perspective of the day, as she explores three current reasons for small classes. First, the standards movement has encouraged the resurgence of the class size debate. Educators and policymakers are looking for strategies that will enable students to succeed on the new assessments and that will enhance students' learning opportunities. Second, class size issues have resurfaced because of the increasing consensus among educators and the public that all students can learn. Cognitive scientists, neurological biologists, and educators determined that all students have the capacity to learn. This new, convincing research means that no student should be left behind in the process. Third, schools have a central responsibility for helping students learn the basic skills of productive citizenry. Class size influences whether teachers are able to

engage students in meaningful discussions of these issues and to help them build these crucial citizenship skills. Schools should strive to develop in students the skills that they will need to examine their differences productively and to coexist peacefully while protecting basic freedoms for all (Goodlad, Soder, & Sirotnik, 1990).

Finn (2002) looks at the reduced class size issue from a slightly different perspective. He believes the reasons that class sizes are finally being reduced at the elementary level are (a) everybody likes the idea of small classes; teachers, parents, policymakers, legislators, and even courts understand the importance of small classes for teaching and learning; (b) research has demonstrated the benefits of small classes in the early grades, especially for students at risk; (c) until recently, education has risen to the top of state and national agencies; and (d) the economy was healthy; ample resources were available to direct toward school improvement.

Finn (2002) sees the future of smaller classes in the elementary grades less clear in the future, as much has changed. "No Child Left Behind" earmarks the federal reduced class size initiative as one of two programs to be eliminated. The recent instability in the economy may leave states and districts less able to hire additional teachers. And, the events of September 11 refocused national attention in a way that may well give lower priority to education issues. It remains to be seen if small class sizes will have become sufficiently important and sufficiently institutionalized that they will continue to be part of basic educational plans (Finn, 2002).

Although over 1,100 studies have examined the relationship between class size and student achievement, no definitive conclusions have been reached. Positive results have been reached in Tennessee and Wisconsin, but other research finds little connection between student-teacher ratios and student performance, especially when measured against other types of educational reform. Critics argue that class sizes have fallen for decades, but a corresponding increase in student performance has not occurred. Also of interest is that the average class size in American schools has dropped from 30 in 1961 to 23 in 1998, without any improvement in standardized test scores (Council of State Governments, 2000).

Findings from current initiatives in a majority of states in the United States seem to point to the following conclusions about small classes. When planned thoughtfully and funded adequately, small classes in the early grades generate gains for students, and those extra gains are greater the longer students are exposed to small classes. Extra gains from small classes in the early grades are larger when the class has fewer than 20 students. Extra gains from small classes in the early grades occur in a variety of academic disciplines and for both traditional measures of student achievement and other indicators of student success. Students whose classes are small in the early grades retain their gains in standard size classrooms and in the upper grades, middle school, and high school. All types of students gain from small classes in the early grades, but gains are greater for students who have traditionally been disadvantaged in education. Initial results indicate that students who have traditionally been disadvantaged in education carry greater small class, early grade gains forward into the upper grades and beyond. The extra gains associated with small classes in the early grades seem to apply equally to boys and girls. Students in small classes led to higher graduation rates and were more likely to pursue college (particularly African-American students) (Achilles, Finn, & Pate-Bain, 2002).

## CHAPTER III

### RESEARCH DESIGN

To determine the impact of class size reductions on student achievement, a quasi-experimental, comparative change design was used. The quasi-experimental design was chosen because it was not possible to randomly assign students and teachers to classrooms, and to control the class size requirement in other ways. Fiscal constraints within the Bismarck, North Dakota, schools prevented the number of small size classrooms from being larger in scope, as only the funding from the federal grant was used in this project to reduce class size in the first, second, and third grades. Class sizes of less than 17 were defined as the small class size group, and classes were considered large if they contained 22 or more students per teacher during the school year.

The purpose of this chapter was to describe the following: (a) the population and samples investigated in this study; (b) the procedures, instruments, and measures used in this study; and (c) the statistical treatment applied in the analysis of the data collected.

#### Population

The school system from which the sample was drawn was Bismarck Public Schools, centrally located and second largest city in the state of North Dakota. The school system has 16 elementary schools scattered across the area of the city, ranging in size from just over 100 students (one class per grade) to over 600 students (three to four classrooms per grade). The sample was based on three consecutive school years: 1999-00, 2000-01, and 2001-02. Large class size was defined as 22 or more students and one teacher, and small class size was defined as one teacher with less than 17 students.

During the school year 1999-00, 38 classrooms were identified to participate in the study based on the number of students in the classroom. These particular classrooms were identified after student numbers had been balanced throughout the district. Participating in the study were 25 classrooms in grades 1, 2, and 3, that each had less than 17 students, and 13 large classrooms in grades 1, 2, and 3, that each had at least 22 students. The study started with 654 students; but, due to absences and other factors, some submitted assessments which had to be excluded from the data. There were 133 first grade students, 304 second grade students, and 217 third grade students. Some students exited prior to the posttest, while some new students entered after the pretest. Only test scores from completed tests from students present for both the pretest and posttest were included in the final data set.

In 2000-01, the sample included 19 large size classrooms, which included student numbers of at least 22 students per teacher in grades 1, 2, and 3, totaling 430 students. There were 15 small size classrooms included in the study, in grades 1, 2, and 3, with student-teacher ratios being less than 17 students per teacher, totaling 241 students. Once again, only test scores from completed tests from students present for both the pretest and posttest were included in the final data set.

In 2001-02, the sample included 13 large classrooms of at least 22 students per teacher in grades 1, 2, and 3 totaling 310 students. There were 21 small size classrooms represented with a total of 323 students in the first, second, third grade population. Classroom size ranged from 9 to 17 students per teacher in the first through third grade classrooms. Only those students who were in attendance both fall and spring were included in the final data set for the third year.

When analyzing student achievement data in both small vs. large size classrooms, the term minority was defined as Native American. This definition was used because Native American students make up the largest minority group in Bismarck and in the state of North Dakota. In this study, Native American students totaled 52 (7.4%) and Caucasian students were at 651 (92.6%) in small classrooms. In large classrooms, Native American students totaled 57 (5.7%) and Caucasian students were at 950 (94.3%). These numbers do not reflect any other minority

populations, as these populations would be too small to draw legitimate statistical conclusions. The total percentage of Native Americans who attend school in the Bismarck Public Schools is comparable to the percentage of students in this study.

Over the three-year period, 11 of the 16 elementary schools were represented in the large class size sample; 13 of the 16 elementary schools were represented in the small class size sample. Eight of the schools had both small and large class size populations included in the study. There were nearly 1,800 primary grade students included in the three-year study, with 106 corresponding classroom teachers participating in the study. The students from small size classrooms and large size classrooms were tested by using the Gates-MacGinitie Reading Test (GMRT) (Fourth Edition), Riverside Publishing, Houghton Mifflin Company. Student achievement was measured by total reading extended scale score gains and normal curve equivalency scores from early October (Fall pretest) to April (Spring posttest) each year.

The teachers who were responsible for teaching these classes during the third year of the study were chosen to take part in the teacher questionnaire. Teachers were sent a letter explaining the study and were requested to participate in the survey during May 2002. (See Appendix A.) The questions were designed to determine how teachers allot their time and also where they direct their efforts, as adapted from the SAGE Program teacher questionnaire in Wisconsin. (See Appendix B.)

The approval from the Institutional Review Board was granted on January 30, 2002. The University of North Dakota requires that any research which involves the use of human subjects be approved by this institutional board. Authorization to conduct research was also sought from the Bismarck Public Schools. This request was granted on November 15, 2001. (See Appendix C.)

The teacher questionnaire included 12 items: (a) more time teaching, (b) covered more content, (c) integrated content, (d) more depth, (e) individualization, (f) more engaging, (g) more hands-on, (h) student's knowledge, (i) problem solving, (j) cooperative groups, (k) more opportunities, and (l) teacher

enthusiasm. Teachers were asked to select the most significant teaching behaviors related to smaller class sizes. From the results of the quantitative analysis of the relationship of teacher behavior to student achievement, correlations of rankings and ratings were determined. Teachers were also asked to respond to an open-ended question: "What is the biggest advantage of being a teacher of small class size or the biggest disadvantage of being a teacher of large class size?" Responses were categorized into themes. (See Appendix D.) As a result of the analysis of the comments, the following eight themes emerged:

1. More quality time can be spent with each child so no one is allowed to slip by.
2. Children get more time to respond and practice skills.
3. Teachers can individualize for each student.
4. Teachers have more time to prepare for small group activities, so a variety of teaching methods can be used, such as small group activities, manipulatives, experiments, and large motor movement.
5. Classroom management is easier; less noise, safer environment; and less time is wasted on discipline so more time can be spent on teaching and learning.
6. Material can be covered faster and more in depth; fewer papers to correct and analyze.
7. Increased sense of belonging; better sense of community; more time to develop social skills; more time to connect with parents through conferences and phone calls.
8. There is more time to connect with specialists who work with students.

#### Instruments

To fulfill the school district's responsibility to provide assessment of the general effectiveness of the CSR grant, the Bismarck Public Schools committed to administering the Gates-MacGinitie Reading Test (GMRT) (Fourth Edition). A pretest was given in the Fall to give a baseline from which to measure and compare achievement gains in reading in the two groups (large and small class size). Then, in April, based on the test's norming dates, the posttest was administered to the same



students in the same large and small sized classrooms. Posttest scores were compared to pretest scores to determine if there were any differences in the amount of reading achievement between students in small vs. large classrooms.

As a part of the Gates-MacGinitie Reading Test being administered in the Fall and again in the Spring, teachers provided a student profile which became part of the sample database. Included in this profile for each student was the following information: test dates (Fall and Spring), school attended, teacher, grade, class count and designation (large or small), student name, gender, date of birth, ethnicity, SES (free or reduced lunch), LEP (limited English proficiency), migrant, and disabilities coded by either a 504 or special education designation.

A teacher questionnaire was administered following the school year. The questionnaire, adapted from the SAGE Program model, was used to determine the type of teaching used by teachers in small and large classroom settings. This instrument obtained teachers' descriptions and judgments of the effects of class numbers on teaching, curriculum, time management, and enthusiasm.

Besides the GMRT's history and research base, this reading test was chosen for its usefulness in achievement assessment. The objective information obtained from the tests is an important basis for selecting students for further individual diagnosis and special instruction, planning instructional emphases, making decisions about grouping students, evaluating the effectiveness of instructional programs, and reporting to parents and the community.

The subtests in the test levels provide information on word decoding, comprehension, word knowledge, and vocabulary. The Gates-MacGinitie Reading Test was scored by the Riverside Scoring Service, providing computer accuracy in scoring and giving a detailed report for all the students. Scoring provided by the Riverside Scoring Service included Normal Curve Equivalent (NCE), Grade Equivalent (GE), Percentile Rank (PR), and Extended Scale Scores (ESS). There are Fall and Spring norming dates established for the Gates-MacGinitie Reading Test.

The study evaluation design utilizes descriptive statistics and multivariate inferential statistics, including linear regression. Descriptive statistics, including means and standard deviations, were computed and subjected to whichever statistical test, t-test or analysis of variance (ANOVA), was appropriate to provide a basis for interpreting the findings. Regression analyses at the individual level are used to enable control variables to be entered in blocks with the variable of interest, thus isolating its effects from the other variables, to allow for a statistical adjustment to equalize the groups on factors where pre-existing differences exist. Such factors would be things like socioeconomic status, ethnicity, and prior knowledge. The reading scores are represented in scaled scores and normal curve equivalents. A scaled score is used to provide a means for comparison across subjects or groups. Multivariate analysis of variance (MANOVA) and analysis of variance (ANOVA) statistics were used to analyze differences in achievement. Hence, the scaled score provides a common yardstick by which scores may be compared reasonably, subject to subject and group to group. The inferential analyses utilize scale scores. For the inferential tests, a significance level of .05 was used and significant results are denoted by an asterisk.

The results of the quantitative analysis of the relationship of teacher behavior to student achievement are reported in Chapter IV and Chapter V. The ratings of teaching behaviors were obtained using a 5-point graduated scale. Each of the 12 teacher questionnaire items were reported (Question 4, items a-l). Comparisons and similarities were reported between the comparative class size teacher participants. With the teacher questionnaire information, additional classroom and environmental factors could be taken into consideration when exploring the value of class size on student achievement in reading.

## CHAPTER IV

### ANALYSIS OF THE DATA

Part one of Chapter IV analyzes the reading achievement data of Bismarck Public Schools students who participated in the small vs. large classrooms while in grades 1, 2, and 3 for the school years 1999-00, 2000-01, and 2001-02. The Gates-MacGinitie Reading Test (Fourth Edition) was the instrument used to measure reading achievement. The pretest was given each Fall and the posttest was administered each Spring of the school year.

Seven questions guided the comparisons made between achievement data. The research questions analyzed were:

1. What effect does class size have on the reading achievement of first grade students?
2. What effect does class size have on the reading achievement of second grade students?
3. What effect does class size have on the reading achievement of third grade students?
4. Is there a significant difference in achievement gains made at different grade levels? (Example: Are there greater gains in first grade than in second or third grade?)
5. Are there substantial achievement gains made by students who are economically disadvantaged as a result of small class size?
6. Is there a gender difference between small and large classrooms and student achievement?
7. Do Native American students have greater achievement gains than other students as a result of small class size?

The second part of Chapter IV is a comparison analysis of the data collected from teachers in the Bismarck Public Schools who taught students in the small vs. large classes during the 2001-02 school year. The questions were designed to determine how teachers allotted their time and also where they directed their efforts. The teacher survey included four open-ended questions that were analyzed for commonalities and differences between the teachers of small vs. large classrooms. The survey also included further questions that were rated by utilizing a graduated scale that ranged from 1 to 5. These teacher rankings were correlated between the teachers of small vs. large class size. The questions were adapted from the Student Achievement Guarantee in Education (SAGE) teacher survey in Wisconsin to obtain the following classroom information:

1. Spend more time on individualized instruction, assessing interests, abilities, needs, and personalities.

2a/b. More teacher enthusiasm for teaching reading, as compared to prior year.

3a. Covered reading content in more depth.

3b. Covered more content in reading.

4a. More time spent diagnosing the needs of individual students.

4b. Spent more time teaching rather than disciplining the class.

4c. Spent more time providing help for individual students.

4d. Spent more time organizing the class into cooperative groups.

4e. Spent more time assessing the progress of individual students.

4f. Spent more time assessing class progress.

4g. Spent more time working with students one-on-one.

4h. Spent more time disciplining individual students.

4i. Involved more students in hands-on activities.

4j. Spent more time working with students on special projects.

4k. Spent more time on creative projects for the class.

To prevent researcher bias, Edward Simanton, Ph.D., at the University of North Dakota's Bureau of Educational Services and Applied Research, was retained

to compile the statistics for the study. The research questions were answered as the data from students and teachers were statistically analyzed.

### Class Size and Reading Achievement in the Primary Grades

Tables 1 and 2 describe the characteristics of all students who were part of the study from 1999-02. Tables 3 through 6 pertain to research questions one, two, three, and four. Extended scale scores and grade equivalency scores were used from the Gates-MacGinitie Reading Test given in the Fall and Spring of each of the past three years.

Table 1 shows the total number and percentage of students in the small vs. large class study during the school years 1999-00, 2000-01, and 2001-02 based on gender, ethnicity, socioeconomic status (SES), and grade level.

The data in Table 1 represent the number of students/classrooms involved in the small vs. large class study in the Bismarck Public Schools during the school years 1999-00, 2000-01, and 2001-02. The total number of students in the study was 1,771, with gender about equally distributed. Grade distribution was 346 in grade 1, 794 in grade 2, and 631 in grade 3. The number of students taught in small classrooms was 739 (41.7%), whereas 1,032 (58.3%) were taught in large classrooms. Of the total student population in the study, 109 (6.2%) indicated a minority status, that being American Indian. Approximately one fourth of the students (472 or 26.7%) received free or reduced lunch.

Table 2 represents student characteristics in small vs. large classrooms that participated in the Bismarck Public Schools study from 1999-02. Frequencies of gender, ethnicity, socioeconomic status, and grade level of students are reported.

When comparing small vs. large classrooms in Table 2, the data indicate no difference in distribution from the total numbers listed in Table 1. In other words, all variables (gender, ethnicity, and free and reduced lunch) remained constant when small and large classrooms were compared to the totals.

Table 1  
Characteristics of All Small and Large Classroom Students From 1999-2002  
by Frequency and Percent

Characteristic	Grade 1		Grade 2		Grade 3		Total	
	N	%	N	%	N	%	N	%
<b>Gender</b>								
Female	178	51.6	380	49.2	296	48.9	853	48.2
Male	167	48.4	392	50.8	308	51.1	867	49.0
Total	345	100.0	772	100.0	603	100.0	1,720	97.1
<b>Race</b>								
Caucasian	314	93.7	720	94.6	567	92.3	1,601	90.4
Native American	21	6.3	41	5.4	47	7.7	109	6.2
Total	335	100.0	761	100.0	614	100.0	1,710	96.6
<b>Sub Lunch</b>								
No	213	68.1	436	67.8	325	66.3	974	55.0
Yes	100	31.9	207	32.2	165	33.7	472	26.7
Total	313	100.0	643	100.0	490	100.0	1,446	81.6
<b>Students</b>								
Small Classrooms	162	46.8	359	45.2	218	34.5	739	41.7
Large Classrooms	184	53.2	435	54.8	413	65.5	1,032	58.3
Total	664	100.0	638	100.0	469	100.0	1,771	100.0

In the analyses to follow, reading achievement was compared from the Gates-MacGinitie pretest (Fall) to the posttest (Spring) scores for the three years studied. Tables 3, 4, and 5 indicate total reading extended scale scores (TRESS) and normal curve equivalency (NCE) scores from students in small vs. large sized classes. The mean, standard deviation, achievement differences, and the significance between small and large sized classroom levels are compared for those students with valid pretests and posttests in grades 1, 2, and 3, respectively.

To answer question one, "What effect does class size have on the reading achievement of first grade students?," data from first grade students were analyzed. Table 3 shows the mean, number, and standard deviation for both the small and large size first grade classrooms.

Table 2

Number of Students in Small and Large Classrooms in Bismarck Public Schools by Gender, Ethnicity, Socioeconomic Status, and Grade From 1999-2002

	Small Classroom				Large Classroom				Small and Large Classrooms
	99-00	00-01	01-02	Total	99-00	00-01	01-02	Total	Total
<b>Gender</b>									
Male	134	121	110	365	88	213	201	502	867
% within size				50.0				50.7	50.4
Female	126	112	127	365	82	218	188	488	853
% within size				50.0				49.3	49.6
Total	260	233	237	730	170	431	389	990	1,720
% within size				100.0				100.0	100.0
<b>Ethnicity</b>									
Native American	24	16	12	52	2	33	22	57	109
% within size				7.4				5.7	6.4
Caucasian	232	209	210	651	198	392	360	950	1,601
% within size				92.6				94.3	93.6
Total	256	225	222	703	200	425	382	1,007	1,710
% within size				100.0				100.0	100.0

Table 2 (Cont.)

	Small Classroom				Large Classroom				Small and Large Classrooms Total
	99-00	00-01	01-02	Total	99-00	00-01	01-02	Total	
<b>Free or Reduced Lunch</b>									
No	117	161	132	410	118	321	125	564	974
% within size				64.7				69.5	67.4
Yes	83	72	69	224	40	110	98	248	472
% within size				35.3				30.5	32.6
Total	200	233	201	634	158	432	233	812	1,446
% within size				100.0				100.0	100.0
<b>Grade</b>									
One	--	62	100	162	--	118	66	184	346
% within size				21.9				17.8	19.5
Two	141	130	88	359	111	201	123	435	794
% within size				48.6				42.2	44.8
Three	125	41	52	218	92	112	209	413	631
% within size				29.5				40.0	35.6
Total				739				1,032	1,771
% within size				100.0				100.0	100.0



Table 3

Total Reading Extended Pre and Post Scale Scores and Normal Curve  
Equivalency Pre and Post Scores (Means and Standard Deviations)  
for First Grade Students From 1999-2002

	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
Pre TRESS	329.1	162	37.3	324.4	184	34.2
Post TRESS	419.4	162	53.3	418.0	184	46.7
TRESS Difference	90.2			93.6		
Significance	.465			.465		
Pre NCE	41.3	160	16.9	39.4	183	16.8
Post NCE	54.8	157	21.5	51.5	177	18.7
NCE Difference	13.5			12.1		
Significance	.553			.553		

Table 3 compared the reading achievement scores of first graders who were in small classrooms (17 or less students per classroom) with students who were in large classrooms (22 or more students) over a three-year period, 1999-02. The reading achievement difference was not significant at the .05 level when the two groups of students were compared. Students in small classes averaged a mean achievement gain of 90.29 in total reading scale scores from Fall to Spring. First graders in large classrooms on average achieved a gain of 93.61, as measured by extended scale scores on the Gates-MacGinitie Reading Test. The mean achievement difference between small and large classrooms at the first grade level was -3.3. As a comparison, average reading achievement gains from Fall to Spring on the GMRT in first grade is 49 TRESS (342 TRESS to 391 TRESS).

When normal curve equivalency pre-post scores in reading achievement were compared with first grade students who were in small classrooms vs. large classrooms, the difference was not significant at the .05 level. Students in small classes averaged a mean achievement gain of 13.5, while students in large classes

averaged a mean achievement normal curve equivalency gain of 12.1, with the mean achievement difference being 1.4.

To answer question two, "What effect does class size have on the reading achievement of second grade students?," data from second grade students were analyzed. Table 4 shows the mean, number, and standard deviation for both the small and large size classrooms. Significance levels were compared between small and large size classrooms at the second grade level.

Table 4

Total Reading Extended Pre and Post Scale Scores and Normal Curve Equivalency Pre and Post Scores (Means and Standard Deviations) for Second Grade Students From 1999-2002

	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
Pre TRESS	425.0	359	42.2	426.4	435	41.0
Post TRESS	461.0	359	48.2	460.1	435	48.8
TRESS Difference	36.0			33.7		
Significance	.370			.370		
Pre NCE	50.0	358	18.5	50.5	430	18.1
Post NCE	56.0	352	18.3	55.5	426	17.4
NCE Difference	6.0			5.0		
Significance	.150			.150		

Table 4 compares the reading achievement scores of second graders who were in small classrooms (17 or less students per classroom) with students who were in large classrooms (22 students or more per teacher) over a three-year period, 1999-02. The reading achievement difference was not significant at the .05 level when students from the two class sizes were compared. Students in the small classes averaged a mean achievement growth of 36.0 TRESS, from Fall to Spring, on the Gates-MacGinitie Reading Test. Second graders in large classrooms on average experienced reading growth of 33.7 TRESS, from Fall to Spring, on the GMRT. The difference in achievement reading scores was 2.3 when second grade

mean scores were compared. Average achievement gain in second grade from Fall to Spring on the GMRT is 23 TRESS (423 TRESS to 446 TRESS).

When normal curve equivalency pre-post scores in reading achievement were compared with second grade students who were in small classrooms vs. large classrooms, the difference was not significant at the .05 level. Students in small classes averaged a mean achievement gain of 6.0 NCE, from Fall to Spring, on the Gates-MacGinitie Reading Test. Second graders in large classrooms on average experienced reading growth of 5.0 NCE from Fall to Spring on the GMRT. The difference in achievement reading scores was 1.0 NCE when second grade mean scores were compared.

To answer question three, "What effect does class size have on reading achievement of third grade students?," data from third grade students were analyzed. Table 5 shows the mean, number, and standard deviation for both the small and large size classrooms at the third grade level. Significant difference is calculated between student achievement in small vs. large classrooms of third graders on the Gates-MacGinitie Reading Test.

Table 5

Total Reading Extended Pre and Post Scale Scores and Normal Curve Equivalency Pre and Post Scores (Means and Standard Deviations) for Third Grade Students From 1999-2002

	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
Pre TRESS	467.5	218	40.5	467.4	413	41.0
Post TRESS	485.6	218	41.7	481.6	413	40.5
TRESS Difference	18.1			14.3		
Significance	.055			.055		
Pre NCE	51.5	214	20.1	53.3	408	19.2
Post NCE	56.5	207	17.5	57.5	397	17.7
NCE Difference	5.0			4.2		
Significance	.128			.128		

Table 5 shows the mean, number, and standard deviation for both the small and large size classrooms. Table 5 compares the reading achievement scores of third graders who were in small classrooms (17 or less students per classroom) with students who were in large classrooms (22 or more students) over a three-year period, 1999-02. The reading achievement difference was not significant at the .05 level, when students from the two class sizes were compared, but was significant at a .055 level.

Students in small classes averaged a growth from Fall to Spring in reading of 18.1 TRESS. Third graders in large classrooms on average achieved a gain in reading of 14.3 TRESS, 3.8 TRESS less than the small size classes.

When normal curve equivalency (NCE) pre-post scores in reading achievement were compared with third grade students who were in small classrooms vs. large classrooms, the difference was not significant at the .05 level. Students in small classes averaged a mean achievement gain of 5.0 NCE, from Fall to Spring, on the Gates-MacGinitie Reading Test. Third graders in large classrooms on average experienced reading growth of 4.2 NCE from Fall to Spring on the GMRT. The difference in achievement reading scores was .8 NCE when third grade mean scores were compared.

#### Class Size and Reading Achievement Gains Across Grades

To answer question four, "Is there a significant difference in achievement gains made at different grade levels? (Example: Are there greater gains in first grade than in second or third grade?)," means were compared within each grade. Table 6 shows the differences in mean achievement scores between the three primary grades on the Gates-MacGinitie Reading Test.

Data from the TRESS in grades 1, 2, and 3 show that greater achievement gains were made in first grade than in second and third grades. Whereas TRESS gains in first grade, 90.3 small classrooms and 93.6 large classrooms, show students in large classrooms outperformed students in small classrooms. Second grade students had TRESS gains of 36.0 in small classrooms and 33.7 in large

classrooms, and third grade students had TRESS gains of 18.2 in small classrooms and 14.3 in large classrooms, showing larger gains in the small classrooms. The NCE gains were greater for all students who were in small classrooms.

Table 6

Differences of Mean Pre-Post Composite Total Reading Extended Scale Scores and Normal Curve Equivalency Scores in Primary Grade Levels in Small vs. Large Size Classrooms Over Three Years (1999-2002)

	Mean Achievement Gain (PR of 50) Fall to Spring	Small Classroom Achievement Gains	Large Classroom Achievement Gains	Mean Achievement Difference
Grade 1				
TRESS Total	49	90.3	93.6	-3.3
NCE Total		13.5	12.1	1.4
Grade 2				
TRESS Total	23	36.0	33.7	2.3
NCE Total		6.0	5.0	1.0
Grade 3				
TRESS Total	13	18.2	14.3	3.8
NCE Total		5.0	4.3	.7
Total				
TRESS		42.6	36.6	6.0
NCE		7.3	6.0	1.3

Median achievement levels (percentile rank of 50) in grades 1, 2, and 3 are represented by total extended scale scores on the Gates-MacGinitie Reading Test corresponding to the median achievement at each grade level. On the TRESS, a score of 500 was set to represent the median achievement level of students in the grade 5 norming group in the Fall. Other TRESSes have no obvious connection to grade level or to achievement level within a grade group. Since the rate of reading growth tapers off during the school years, the gains in TRESSes are not the same from school year to school year. In the first grade, from Fall to Spring, the average achievement gains range from 342 TRESS to 391 TRESS, for an achievement gain

of 49 TRESS. In the second grade, from Fall to Spring, the average achievement gains range from 423 TRESS to 446 TRESS, for an achievement gain of 23 TRESS. In the third grade, from Fall to Spring, the average achievement gains range from 459 TRESS to 472 TRESS, for an achievement gain of 13 TRESS. All gains that were made in both the small vs. large sized classrooms in the Bismarck Public Schools were greater than the average achievement gains made from Fall to Spring at each average grade level represented corresponding to average median achievement gains on the GMRT.

Likewise, the NCEs describe a student's level of achievement in relation to the achievement of other students in the same grade. NCEs are based on percentile ranks that have been transformed statistically into a scale of equal units of reading achievement. NCEs were derived to have an average of 50 and a standard deviation of 21.06. The NCE scale was designed so that the NCEs of 1, 50, and 99 coincide exactly with percentile ranks of 1, 50, and 99.

#### Class Size and Reading Achievement Gains Made by Students Who Are Economically Disadvantaged

To answer question five, "Are there substantial achievement gains made by students who are economically disadvantaged as a result of small class size?," mean scores were compared between economically disadvantaged and non-economically disadvantaged. The term economically disadvantaged was determined using the criterion of student's qualification for free or reduced lunch. Table 7 examines the achievement of economically disadvantaged and non-economically disadvantaged students in small and large classrooms. Economically disadvantaged and non-economically disadvantaged students were compared in small vs. large sized classrooms in grades 1, 2, and 3 on the GMRT total reading scale scores and normal curve equivalency scores.

Non-economically disadvantaged students in small classes had a mean pre TRESS of 419.4 and a post TRESS of 465.7 for a TRESS achievement difference of 46.3. Non-economically disadvantaged students in large classes had a mean pre TRESS of 425.4 and a post TRESS of 466.4 for a TRESS achievement difference

of 41.0. These results were not significant at the .05 level, although significance was at .059. Non-economically disadvantaged students in small classes had a mean pre NCE score of 50.6 and a post NCE score of 58.2 for a difference of 7.6 NCE. Non-economically disadvantaged students in large classes had a mean pre NCE score of 51.7 and a post NCE score of 57.6 for a difference of 5.9 NCE. NCE results between small and large class sizes were significant at the .05 level (.007).

Table 7

Total Reading Scale Scores and Normal Curve Equivalency Scores of Economically Disadvantaged Students Compared to Non-Economically Disadvantaged Students (Grades 1, 2, and 3) on the Gates-MacGinitie Reading Test

Free/Reduced Lunch	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
<b>No</b>						
Pre TRESS	419.4	410	66.8	425.4	564	63.4
Post TRESS	465.7	410	51.9	466.4	564	49.1
TRESS Diff	46.3			41.0		
Significance	.059			.573		
Pre NCE	50.6	406	19.0	51.7	557	18.1
Post NCE	58.2	393	17.9	57.6	544	17.0
NCE Diff	7.6			5.9		
Significance	.007*			.360		
<b>Yes</b>						
Pre TRESS	403.4	224	57.5	403.1	248	65.7
Post TRESS	441.2	224	52.5	438.4	248	51.7
TRESS Diff	46.3			35.3		
Significance	.059			.573		
Pre NCE	41.5	222	17.3	40.6	248	18.5
Post NCE	49.2	223	18.8	46.9	246	18.0
NCE Diff	7.7			6.3		
Significance	.007*			3.60		
<b>Total</b>						
Pre TRESS	413.7	634	64.1	418.6	812	64.9
Post TRESS	457.0	634	53.4	457.8	812	51.5
TRESS Diff	43.3			39.2		

Table 7 (Cont.)

Free/Reduced Lunch	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
Pre NCE	47.4	628	18.9	48.3	805	19.0
Post NCE	54.9	616	18.8	54.3	790	18.0
NCE Diff	7.5			6.0		

\*Denotes significance at the .05 level

Economically disadvantaged students in small classes had a mean pre TRESS of 403.4 and a post TRESS of 441.2 for a TRESS achievement difference of 37.8. Economically disadvantaged students in large classes had a mean pre TRESS of 403.1 and a post TRESS of 438.4 for a TRESS achievement difference of 35.3. These results were not significant at the .05 level. Economically disadvantaged students in small classes had a mean pre NCE score of 41.5 and a post NCE score of 49.2 for a difference of 7.7 NCE. Economically disadvantaged students in large classes had mean pre NCE scores of 40.6 and post NCE scores of 46.9 for a difference of 6.3. These results were not significant at the .05 level.

#### Class Size and Reading Achievement by Gender

To answer question six, "Is there a gender difference between small and large classrooms and student achievement?," male and female students in small vs. large classrooms were compared in grades 1, 2, and 3 on the Gates-MacGinitie Reading Test total scale scores and normal curve equivalency scores. Table 8 examines the reading achievement of male and female students in small and large classrooms. The change in achievement gains was analyzed to see if there was a statistically significant difference between gender and achievement gains.

Male students in small classes had a mean pre TRESS of 413.9 and a post TRESS of 456.0 for a TRESS achievement difference of 42.1. Male students in large classes had a mean pre TRESS of 421.9 and a post TRESS of 458.2 for a TRESS achievement difference of 36.5. Results were not significant at the .05



level. Male students in small classes had a mean pre NCE score of 47.1 and a post NCE score of 54.5 for a difference of 7.4 NCE. Male students in large classes had a mean pre NCE score of 48.6 and a post NCE score of 54.8 for a difference of 6.2 NCE. Again, the results were not significant at the .05 level.

Table 8

Total Reading Scale Scores and Normal Curve Equivalency Scores From 1999-2002 for Male and Female Students in Small vs. Large Classrooms Compared in Grades 1, 2, and 3

Gender	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
<b>Male</b>						
Pre TRESS	413.9	365	63.0	421.9	502	64.1
Post TRESS	456.0	365	53.9	458.4	502	50.8
TRESS Diff	42.1			36.5		
Significance	.062			.062		
Pre NCE	47.1	362	19.8	48.6	496	19.5
Post NCE	54.5	357	19.0	54.8	488	18.1
NCE Diff	7.4			6.2		
Significance	.174			.174		
<b>Female</b>						
Pre TRESS	418.8	365	66.0	425.3	488	63.4
Post TRESS	462.4	365	52.3	463.2	488	50.4
TRESS Diff	43.6			37.9		
Significance	.058			.058		
Pre NCE	49.9	361	18.2	50.6	483	18.4
Post NCE	57.3	350	18.3	56.8	473	17.8
NCE Diff	7.4			6.2		
Significance	.031*			.031*		
<b>Total</b>						
Pre TRESS	416.3	730	64.5	423.5	990	63.7
Post TRESS	459.2	730	53.2	460.8	990	50.6
TRESS Diff	42.9			37.3		

Table 8 (Cont.)

Gender	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
Pre NCE	48.5	723	19.1	49.6	979	19.0
Post NCE	55.9	707	18.7	55.8	961	18.0
NCE Diff	7.4			6.2		

\*Denotes significance at the .05 level

Female students in small classes had a mean pre TRESS of 418.8 and a post TRESS of 462.4 for a TRESS achievement difference of 43.6. Female students in large classes had a mean pre TRESS of 425.3 and a post TRESS of 463.2 for a TRESS achievement difference of 37.9. The results were not significant at the .05 level; they were significant at .058. Female students in small classes had a mean pre NCE score of 49.9 and a post NCE score of 57.3 for a difference of 7.4 NCE. Female students in large classes had a mean pre NCE score of 50.6 and a post NCE score of 56.8 for a difference of 6.2 NCE. These results were significant at the .05 level (.031).

#### Class Size and Reading Achievement of Native American vs. Caucasian Students

To answer question seven, "Do Native American students have greater achievement gains than other students as a result of small class size?," Caucasian students were compared to Native American students in grades 1, 2, and 3 on the Gates-MacGinitie Reading Test total reading scale scores and normal curve equivalency scores. Table 9 examines the reading achievement of Caucasian and Native American students in small and large classrooms.

Native American students in small classes had a mean pre TRESS of 401.3 and a post TRESS of 436.9 for a TRESS achievement difference of 35.6. Native American students in large classes had a mean pre TRESS of 407.4 and a post TRESS of 445.7 for a TRESS achievement difference of 38.3. The results were not significant at the .05 level. Native American students in small classes had a mean

Table 9

Total Reading Scale Scores and Normal Curve Equivalency Scores of Caucasian Students Compared to Native American Students in Small vs. Large Classrooms in Grades 1, 2, and 3

Race	Small Classrooms			Large Classrooms		
	Mean	N	SD	Mean	N	SD
Native American						
Pre TRESS	401.3	52	55.3	407.4	57	70.2
Post TRESS	436.9	52	54.7	445.7	57	55.3
TRESS Diff	35.6			38.3		
Significance	.782			.782		
Pre NCE	38.2	52	15.1	41.7	57	20.1
Post NCE	46.0	52	16.2	48.4	56	19.5
NCE Diff	7.8			6.7		
Significance	.901			.901		
Caucasian						
Pre TRESS	417.3	651	65.5	425.8	950	62.7
Post TRESS	461.1	654	52.5	462.4	950	50.2
TRESS Diff	43.8			36.6		
Significance	.001*			.001*		
Pre NCE	49.1	645	19.1	41.7	57	20.1
Post NCE	56.7	631	18.5	56.2	919	17.7
NCE Diff	7.8			5.9		
Significance	.003*			.003*		
Total						
Pre TRESS	416.1	703	64.9	424.8	1,007	50.6
Post TRESS	459.3	703	53.0	461.4	1,007	50.6
TRESS Diff	43.2			36.6		
Pre NCE	48.3	697	19.1	49.8	996	18.9
Post NCE	55.9	683	18.5	49.8	996	18.9
NCE Diff	7.6			5.9		

\*Denotes significance at the .05 level

pre NCE score of 38.2 and a post NCE score of 46.0 for a difference of 7.8 NCE. Native American students in large classes had a mean pre NCE score of 41.7 and a

post NCE score of 48.4 for a difference of 6.7 NCE. Again, the results were not significant at the .05 level.

Caucasian students in small classes had a mean pre TRESS of 417.3 and a post TRESS of 461.1 for a TRESS achievement difference of 43.8. Caucasian students in large classes had a mean pre TRESS of 425.8 and a post TRESS of 462.4 for a TRESS achievement difference of 36.6. The results were significant at the .05 level (.001). Caucasian students in small classes had a mean pre NCE score of 49.1 and a post NCE score of 56.7 for a difference of 7.6 NCE. Caucasian students in large classes had a pre NCE score of 50.3 and a post NCE score of 56.2 for a difference of 5.9 NCE. These results were also significant at the .05 level (.003). In this study, Caucasian students had greater TRESS gains in the smaller classes than did Native American students; but Native American students had greater NCE gains in the small classrooms. In the larger classrooms, Native American students had greater gains in both the TRESS and NCE scores.

#### Teacher Questionnaire Results, Grades 1, 2, and 3, School Year 2001-02

The teacher survey was completed in May 2002 by teachers who had small and large sized classrooms during the 2001-02 school year. Participation in the survey was voluntary. Twenty-four teachers, out of a possible 34, responded by filling out the survey and returning it. Of the teachers who responded, 41.4% had taught in large classrooms during the year and 58.3% had taught in small sized classrooms. The rate of return was 72.7%. Six of 10 first grade teachers responded to the survey, 8 of 12 second grade teachers responded, and 10 of 13 third grade teachers responded by returning the survey.

In answer to the question, "Have you incorporated new techniques in reading instruction into your classroom in the past year?," 80% of the teachers in the large classrooms answered yes. This included affirmative responses from one first grade teacher, two second grade teachers, and five third grade teachers. Teachers in small sized classrooms also answered affirmative to adding new reading instruction techniques (83.3%). This included affirmative responses from three first grade

teachers, four second grade teachers, and three third grade teachers. Using the Pearson Chi-Square Test and Fisher's Exact Test, there was not a significant difference found amongst teachers in grade level, or between class size, on this reading strategy question.

The teachers were then asked to respond to a follow-up question: "If you incorporated new techniques in reading, were they related to class size?" Of the teachers in large size classrooms, 66.7% said that, yes, it was related. Zero first grade teachers, two second grade teachers, and four third grade teachers from large sized classrooms answered affirmative to there being a relationship between the incorporation of new techniques in reading being related to class size. Thirty-eight percent of the teachers in small size classrooms answered yes. Those teachers answering affirmative included zero teachers from grade 1, two from grade 2, and two from grade 3. The Pearson Chi-Square Test and Fisher's Exact Test were completed and neither test showed a significant difference amongst teachers in grade level, or between class size, on this question relating to new reading techniques being related to class size.

When teachers were asked to respond to the question, "What is the biggest advantage of being a teacher of small class size or the biggest disadvantage of being a teacher of large class size?," the following eight general themes resulted from teacher responses to the survey:

1. More quality time can be spent with each child so no one is allowed to slip by.
2. Children get more time to respond and practice skills.
3. Teachers can individualize for each student.
4. Teachers have more time to prepare for small group activities, so a variety of teaching methods can be used, such as small group activities, manipulatives, experiments, and large motor movement.
5. Classroom management is easier; less noise, safer environment; and less time is wasted on discipline so more time can be spent on teaching and learning.

6. Material can be covered faster and more in depth; fewer papers to correct and analyze.

7. Increased sense of belonging; better sense of community; more time to develop social skills; more time to connect with parents through conferences and phone calls.

8. There is more time to connect with specialists who work with students.

The second part of the teacher survey included responses to 17 items related to time allotment and classroom direction. Teacher responses were rated on a graduated scale from 1 to 5. Table 10 shows the frequency and percentage of teacher responses to items related to time allotment and classroom directed efforts in small vs. large size classrooms during the 2001-02 school year.

Table 10 represents the number of primary teachers involved in the small vs. large class study in the Bismarck Public Schools during the 2001-02 school year. There were 14 primary teachers who had taught in small size classrooms (five first, five second, one combined second/third, and three third grade teachers) and 10 primary teachers who had taught in large size classrooms (one first, two second, and seven third grade teachers) who responded to the survey. The frequency and percentage of teacher responses to each of the 17 items are recorded in Table 10.

When comparing the responses from teachers in small vs. large classrooms during the 2001-02 school year, the following differences were found relating to time allotment and classroom directed efforts. Teachers in small sized classrooms leaned toward the particulars of the students as individuals rather than toward the particulars of the class in general (Q. 1). Teachers of small sized classes spent less time on disciplining the class compared with teachers of large sized classes (Q. 4b) and on disciplining individual students (Q. 4h). Additionally, teachers of small sized classes spent more time working with small groups of students (Q. 4d), working with students one on one (Q. 4g), working with students on special projects (Q. 4j), and creating special projects for the class (Q. 4k).

Table 10  
Frequency and Percentage of Teacher Responses to Items Related to  
Time Allotment and Classroom Directed Efforts for Both Small  
and Large Classrooms

Survey Item	1		2		3		4		5		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
	Students as individuals						Class in general					
1. Class particulars (sm)	1	7.1	7	50.0	3	21.4	3	21.4			14	100.0
1. Class particular (lg)			2	20.0	4	40.0	4	40.0			10	100.0
	Not enthusiastic						Very enthusiastic					
2a. Reading enthusiasm (sm) [past year]							7	50.0	7	50.0	14	100.0
2a. Reading enthusiasm (lg) [past year]					1	10.0	4	40.0	5	50.0	10	100.0
2b. Reading enthusiasm (sm) [current year]					3	21.4	4	28.6	7	50.0	14	100.0
2b. Reading enthusiasm (lg) [current year]					1	10.0	4	40.0	5	50.0	10	100.0
	Not deep						Very deep					
3a. Curriculum content (sm)					1	7.7	11	84.6	1	7.7	13	100.0
3a. Curriculum content (lg)					5	50.0	3	30.0	2	20.0	10	100.0
	Cover grade level						Work into next grade level					
3b. Curriculum content (sm)			2	16.7	5	41.7	5	41.7			12	100.0
3b. Curriculum content (lg)	2	20.0	1	10.0	3	30.0	3	30.0	1	10.0	10	100.0

Table 10 (Cont.)

Survey Item	1		2		3		4		5		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
	Little amount of time				Large amount of time							
4a. Diagnosing individual student needs (sm)			1	7.7	7	53.8	4	30.8	1	7.7	13	100.0
4a. Diagnosing individual student needs (lg)					6	60.0	4	40.0			10	100.0
4b. Disciplining the class (sm)	2	14.3	5	35.7	5	35.7	1	7.1	1	7.1	14	100.0
4b. Disciplining the class (lg)			2	20.0	3	30.0	4	40.0	1	10.0	10	100.0
4c. Help for individual students (sm)					5	35.7	8	57.1	1	7.1	14	100.0
4c. Help for individual students (lg)			3	30.0	2	20.0	4	40.0	1	10.0	10	100.0
4d. Working with small groups of students (sm)			2	14.3	6	42.9	6	42.9			14	100.0
4d. Working with small groups of students (lg)	1	10.0	2	20.0	5	50.0	2	20.0			10	100.0
4e. Assessing the progress of individual students (sm)			1	8.3	5	41.7	6	50.0			12	100.0
4e. Assessing the progress of individual students (lg)			2	20.0	2	20.0	4	40.0	2	20.0	10	100.0
4f. Assessing class progress (sm)					6	42.9	8	57.1			14	100.0
4f. Assessing class progress (lg)					4	40.0	4	40.0	2	20.0	10	100.0
4g. Working with students one on one (sm)					9	64.3	4	28.6	1	7.1	14	100.0
4g. Working with students one on one (lg)			3	30.0	6	60.0			1	10.0	10	100.0
4h. Disciplining individual students (sm)	4	28.6	3	21.4	5	35.7	1	7.1	1	7.1	14	100.0
4h. Disciplining individual students (lg)			4	40.0	3	30.0	3	30.0			10	100.0
4i. Involved more students in hands-on activities (sm)					5	35.7	7	50.0	2	14.3	14	100.0
4i. Involved more students in hands-on activities (lg)					6	60.0	1	10.0	3	30.0	10	100.0



Table 10 (Cont.)

Survey Item	1		2		3		4		5		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
	Little amount of time						Large amount of time					
4j. Working with students on special projects (sm)		6	42.9	4	28.6	4	28.6				14	100.0
4j. Working with students on special projects (lg)				5	50.0	5	50.0				10	100.0
4k. Creative projects for class (sm)		5	35.7	3	21.4	5	35.7	1	7.1		14	100.0
4k. Creative projects for class (lg)		1	10.0	2	20.0	3	30.0	4	40.0		10	100.0
4l. Using direct instruction (sm)		1	7.1	2	14.3	5	35.7	6	42.9		14	100.0
4l. Using direct instruction (lg)				1	10.0	2	20.0	7	70.0		10	100.0

Table 11 shows responses to the question, "Is there a significant difference in classroom teacher perceptions whether they taught in small vs. large classrooms during the 2001-02 school year?" In Table 11, both descriptive statistics (mean, standard deviation) and statistics (mean square, F ratio, and significance level) obtained through the use of analysis of variance (ANOVA) were analyzed from the teachers' scaled responses.

When comparing the responses from teachers in small vs. large classrooms, the following differences in perception were found between teachers who had taught in small vs. large classrooms during the 2001-02 school year. Response items found to be significantly different between teachers of small vs. large class size are denoted with an asterisk (\*) in Table 11. Of the 17 items, the items listed below showed the most teacher response difference:

- Q. 1. Class particulars in general (.016)
- Q. 4b. Disciplining the class (.008)
- Q. 4d. Working with small groups (.042)
- Q. 4g. Working with students one on one (.017)

Q. 4j. Working with students on special projects (.004)

Q. 4k. Creative projects for the class (.008)

Table 11

Comparison of Mean Teacher Responses in Small vs. Large Classrooms as to Where They Spend Their Time and How They Direct Their Classrooms

Survey Item	Large		Small		Total		Mean Square	F	Sig
	M	SD	M	SD	M	SD			
1. Class particulars	3.20	.79	2.7	.94	2.83	.91	4.619	6.226	.016*
2a. Reading enthusiasm (past)	4.40	.70	4.50	.52	4.46	.58	.117	.340	.563
2b. Reading enthusiasm (current)	4.40	.70	4.29	.83	4.33	.75	.152	.264	.610
3a. Curriculum content	3.70	.82	4.00	.41	3.87	.62	1.017	2.763	.104
3b. Curriculum content	3.00	1.33	3.25	.75	3.14	1.03	.682	.644	.427
4a. Diagnosing individual student needs	3.40	.52	3.38	.77	3.39	.65	.432	.006	.938
4b. Disciplining the class	3.40	.97	2.57	1.09	2.92	1.09	8.010	7.731	.008*
4c. Help for individual students	3.30	1.06	3.71	.61	3.54	.82	2.002	3.079	.086
4d. Working with small groups of students	2.80	.92	3.29	.73	3.08	.82	2.752	4.379	.042*
4e. Assessing the progress of individual students	3.60	1.07	3.42	.67	3.50	.85	.367	.503	.482
4f. Assessing class progress	3.80	.79	3.57	.51	3.67	.63	.610	1.553	.219
4g. Working with students one on one	2.90	.88	3.43	.65	3.21	.77	3.260	6.081	.017*
4h. Disciplining individual students	2.90	.88	2.43	1.22	2.63	1.08	2.593	2.265	.139
4i. Involved more students in hands-on activities	3.70	.95	3.79	.70	3.75	.79	.629	.136	.714
4j. Working with students on special projects	3.50	.53	2.86	.86	3.13	.79	4.821	9.079	.004*
4k. Creative projects for the class	4.00	1.05	3.14	1.03	3.50	1.09	8.571	8.313	.006*
4l. Using direct instruction	4.60	.70	4.14	.95	4.33	.86	2.438	3.480	.069

\*Denotes significance at the .05 level

The last question on the teacher survey was for the teachers to indicate on which teaching item they wished that they could spend more time. The teachers had the opportunity to respond with two choices from the teacher survey list (see Tables 10 and 11). There were two blanks for teachers to respond with Choice #1 and Choice #2. Table 12 shows the frequency and percentage of teacher responses to the question of preferred area to spend additional time.

Table 12

Frequency and Percentage of Teacher Responses to Areas Where They Would Like to Spend Additional Time

Choices	Choice #1				Choice #2			
	Large Classroom Teacher		Small Classroom Teacher		Large Classroom Teacher		Small Classroom Teacher	
	N	%	N	%	N	%	N	%
4c. Help for individual students	1	10.0	1	7.1				
4d. Working with small groups of students	2	20.0	3	21.4	2	20.0	1	7.1
4e. Assessing the progress of individual students					1	10.0		
4f. Assessing class progress			1	7.1				
4g. Working with students one on one	5	50.0			3	30.0	3	21.4
4i. Involved more students in hands-on activities	1	10.0	3	21.4	2	20.0	1	7.1
4j. Working with students on special projects			2	14.3	1	10.0	3	21.4
4k. Creative projects for the class			2	14.3	1	10.0	4	28.6
Missing Data	1	10.0	2	14.3			2	14.3
Totals	10	100.0	14	100.0	10	100.0	14	100.0

Teachers in large classrooms indicated that working with students one on one is their most preferred way to spend additional time, followed by working with small groups of students and engaging students in hands-on activities. Teachers in small classrooms indicated preferences, in order, for developing creative projects for their class, working with students on special projects, engaging students in hands-on activities, and working with small groups of students.

### Summary

This chapter presented the results of the analysis of data by research question. Based on the data from the Gates-MacGinitie Reading Test scores, there were 1,771 of a potential 2,575 students who had pre-post scores to compare achievement, providing an adequate representation of the participants in the study. There were student samples from each of the primary grades, 1, 2, and 3, and there were classrooms represented for three consecutive years, 1999-02. The population was taken from all 16 of the elementary schools in Bismarck, with many of the schools having both small and large classes represented in the sample.

When reading achievement was compared using both GMRT total reading extended scale scores (TRESS) and normal curve equivalency scores (NCE), there was not a significant difference (using the .05 level) between student achievement gains in the small classrooms over the large classrooms in any grade level--1, 2, or 3. Tables 3, 4, and 5 included analysis of variance (ANOVA) statistics comparing degrees of freedom, sum of squares, mean squares, and F ratios for each grade level's test scores.

The data analysis, though not statistically significant between small and large class sizes, did show a wide variance in achievement gains made during the primary grades in reading. First graders achieved TRESS reading gains at 90.3 (small) and 93.6 (large) from Fall to Spring, second graders achieved TRESS reading gains at 36.0 (small) and 33.7 (large) from Fall to Spring, and third graders showed reading TRESS gains of 18.2 (small) and 14.3 (large) for the school year. In the study, the largest difference in gains happened between the large and small class size at the third grade level. The GMRT extended scale scores taper off in growth as grade

level increases. The median achievement (percentile rank of 50) for grade 1 from Fall to Spring is 49 TRESS; for grade 2, 23 TRESS; and for grade 3, 13 TRESS.

When comparisons were made between student achievement based on socioeconomic status (SES), the students not on free-reduced meal plans scored significantly better (.007) than the students who were on the subsidized lunch program. Table 7 shows the SES comparisons.

When gender comparisons were made between student reading achievement in small vs. large classrooms, in the primary grades, female students who were in small classrooms scored significantly higher (.031) than their male counterparts. The TRESS change was .058. Table 8 shows gender comparisons.

When Caucasian students at the primary grade level were compared on reading achievement scores between small vs. large classrooms, the Caucasian students achieved more in the small classrooms. The Caucasian TRESS scores were significant at the .001 level and the NCE scores were significant at the .003 level. In the larger classrooms, Native American students had greater gains in both the TRESS and NCE scores.

To further analyze reading achievement in the primary grades, teacher attitudes toward teaching reading in large vs. small classrooms, teacher perception data were collected. This sample was small as it encompassed just the teachers in the final year of the study, 2001-02. The Pearson Chi-Square Test and Fisher's Exact Test were utilized in analyzing data. Because of limited numbers, these statistical tests were of limited use, and the teachers' perceptions from small vs. large classrooms were not found to be statistically significant at the .05 level in most cases. By using ANOVA, there were some significant differences in teacher perceptions. Table 11 has the ANOVA results in detail. Table 12 further substantiates the differences between the small vs. large classroom teachers' perceptions using frequencies and percentages of responses. Of the 17 items listed, the following items showed the most teacher response difference: class particulars in general, disciplining the class, working with small groups, working with students one on one, working with students on special projects, and creating projects for the class.

Differences were also found when teachers were asked their most preferred way to spend additional time. Teachers in large classrooms indicated that working with students one on one is their most preferred way to spend additional time, followed by working with small groups of student, and engaging students in hands-on activities. Teachers in small classrooms indicated preferences, in order, for developing creative projects for their classes, working with students on special projects, engaging students in hands-on activities, and working with small groups of students.

Chapter V presents a summary of the study, discussion of the findings, and conclusions. Also included are an action plan and recommendations for further study.

## CHAPTER V

### SUMMARY, DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Chapter V presents a summary of the study and a discussion of the findings. The chapter also presents the conclusions of the study and recommendations for further action and study.

#### Summary of the Study

The purpose of this study was to examine the effect of class size on reading achievement of primary grade students in the Bismarck Public Schools. Scores for the Gates-MacGinitie Reading Test (Fourth Edition) were obtained from the school district and were categorized according to grade, gender, socioeconomic status, ethnicity, and class size. The students were identified as having been in either small classes (17 or less) or large classes (22 or more). Pretest and posttest scores were obtained for the 1999-00, 2000-01, and 2001-02 school years.

In addition, a survey instrument was distributed to the teachers who taught in small and large classrooms during the 2001-02 school year. The instrument collected teachers' perceptions regarding instructional practices and classroom management and was used to support analyses by class size.

Although research has been conducted in other states into various aspects of class size reduction, with particular focus on the ways it affects student achievement, little or no research has been conducted in the state of North Dakota. Since 1999, school districts have received federal financial support to reduce class size. As of 2002, the federal Class Size Reduction (CSR) Program is no longer available to districts. Federal funding is now in block grants where state and district officials make decisions on whether to maintain additional teachers and smaller classes. The data from this study will be made available to the Bismarck school administration for their

use in making future decisions concerning class size reductions. The findings may also be of interest and value to legislators, state education agency personnel, and local school district leaders.

Students who were in either small or large classrooms during the 1999-02 years in the Bismarck Public Schools were given the Gates-MacGinitie Reading Test (GMRT) in September and again in April of each year. Although a total of 2,575 students were involved during parts of the study, 1,771 (68%) of them were included in this study, as these were the students who had completed both pre and post GMRT.

In order to gather data for this Bismarck Public Schools study, each teacher involved in small vs. large classrooms during the 2001-02 school year was asked to complete the teacher survey in May 2002. Twenty-four of 34 teachers involved in the district program voluntarily responded by filling out the instrument for this small vs. large classroom study. The survey was color coded to identify responses from teachers who had taught in small classrooms and those who had taught in large classrooms. No demographic or personal information was included as part of the survey; therefore, individual teacher responses were not able to be identified with particular subjects. Of the teachers who responded, 41.7% had taught in large classrooms during the year and 58.3% had taught in small classrooms during the year. The rate of return was 72.7%, with six small sized and three large sized classroom teachers choosing not to be a part of the study.

The raw student data from the Gates-MacGinitie Reading Test, measuring reading achievement, and the teacher survey data were coded, analyzed, and interpreted with the assistance of Dr. Edward Simanton at the University of North Dakota. The findings of the research questions are listed and discussed in this chapter in the order as they appear in the study. Statistical Software for the Social Sciences (SPSS) was used in the statistical analysis of the study. Chi-Square and ANOVA tests were utilized to test for significant differences between students and teachers in small vs. large classroom settings.



## Discussion of Research Questions

1. What effect does class size have on the reading achievement of first grade students?

In grade 1, 346 students had valid pre-post GMRT scores over the three-year period. Of those students, 162 were in small sized classrooms and 184 students were in large sized classrooms. The mean pre total reading extended scale scores (TRESS) was 329.1 for students in small classrooms and 324.4 for students in large classrooms. The post TRESS was 410.4 for students in small classrooms and was 419.4 for students in large classrooms at the end of first grade. Class size was not found to be of significance when reading achievement scores were compared.

The students' reading achievement scores, showing that small class size does not improve achievement in the primary grades, would be consistent with the California research (2002) that showed no clear correlation between class downsizing and academic improvement. In the third Class Size Reduction Research Consortium, including RAND and the American Institute for Research, there was no clear link showing that reduced class size improved achievement in California (Council of State Governments, 2002).

The findings for the Bismarck Public Schools CSR study would also be consistent with the Hoxby (2000) class size research study completed in the state of Connecticut. Hoxby found, by using two different methods, that reductions in class size had no effect on student achievement. Her class size measurement is consistent with the class size numbers in the Bismarck Public Schools study, as the research was based on class size in the range of greater than 10 students and less than 30 students per classroom.

Two other research studies have also found no relationship between small class size and increased student achievement. Using 1993-94 school level data, the Florida Department of Education (1998) found that there was not a correlation between greater achievement and reduced class size. Hanushek (1998), a leading

economist, lecturer, and writer, has reviewed numerous studies which indicate that reducing class size should not be expected to produce better student performance.

2. What effect does class size have on the reading achievement of second grade students?

In grade 2, 794 students had valid pre-post GMRT scores over the 1999-02 period. Of those students, 359 were in small classrooms and 425 were in large classrooms. The mean pre TRESS was 425.0 for students in small classrooms and 426.4 for students in large classrooms. The post TRESS was 461.0 for students in small classrooms and 460.1 for students in large classrooms at the end of the second grade. Class size was not found to be of significance when reading achievement scores were compared.

The student achievement scores showing that class size does not improve reading achievement in the primary grades would again be consistent with the California research (Council of State Governments, 2002), the Connecticut statewide research (Hoxby, 2000), the Florida research (Florida Department of Education, 1998), and the Hanushek (1998) studies that were conducted.

3. What effect does class size have on the reading achievement of third grade students?

In grade 3, 631 students had valid pre-post GMRT scores over the 1999-02 period. Of those students, 218 were in small classrooms and 413 were in large classrooms. The mean pre TRESS was 467.6 for students in small classrooms and 467.4 for students in large classrooms. The post TRESS was 485.6 for students in small classrooms and 481.7 for students in large classrooms at the end of third grade. Class size was not found to be significant at the .05 level when reading achievement scores were compared, but was significant at the .055 level.

Once again, student achievement scores showing that class size does not improve reading achievement in the primary grades would be consistent with the studies previously cited for research questions one and two. Odden (1990) reviewed existing research and argued that a system-wide class reduction policy would produce only modest gains in student achievement and incur an unjustifiable

high cost. He opted instead for certain targeted class reduction strategies in conjunction with a series of other interventions, and claimed that his proposals could produce greater benefits with lower costs.

Other research analyses have concluded that class size reduction does not have an appreciable effect. Tomlinson (1988) examined trend data from 1950 to 1986 in the United States and did not find any consistent relationship between class size and standardized test scores.

4. Is there a significant difference in achievement gains made at different grade levels? (Example: Are there greater gains in first grade than in second or third grade?)

Overall, first graders make greater scale score gains than second graders, and second graders make greater scale score gains than third grade students. Because there was not a significant difference in small classrooms, compared to large classrooms, and student achievement at any of the three primary grades studied, it could not be concluded that having smaller classes at one grade level is more important than at a different grade level. The fact that students in the first grade make TRESS gains (90.3, small classroom; 93.6, large classroom) of more than double that of second graders (36.0, small classroom; 33.7, large classroom) and considerably more than third graders (18.0, small classroom; 14.3, large classroom), it might be advisable to have fewer students per teacher in first grade. TRESS average gain from Fall to Spring in the first grade is 49, in the second grade the average TRESS gain is 23, and in the third grade the TRESS average gain is 13. The TRESS average yearly gain tapers off during the school years because reading growth becomes less as students progress in grade level. Based on the Bismarck Public Schools primary grade achievement findings, the differences at the third grade level were greater between small and large classrooms than they were at the first grade level, where greater achievement gains were made by students in large sized classrooms.

In 1984, Indiana's initial CSR effort, PRIME TIME project, was to reduce class size to 18 students per teacher in grades 1 and 2. The class size reduction

effort was then to expand to kindergarten and grade 3. Students in smaller classes in the first grade showed the greatest achievement improvement (Gilman & Antes, 1985). In 1987, PRIME TIME was realigned to reduce class sizes to a ratio of 1:18 in kindergarten and grade 1 and to a ratio of 1:20 in grades 2 and 3 (Bain & Jacobs, 1990). The results of Indiana's CSR effort are contrary to the Bismarck Public Schools findings on class size and student achievement.

5. Are there substantial achievement gains made by students who are economically disadvantaged as a result of small class size?

There were 410 students in small sized classrooms who had valid pre-post test scores who did not qualify for the free-reduced lunch program. There were 224 students in small sized classrooms who did qualify for the free-reduced lunch program. Students who did not qualify for the free-reduced lunch program and were in large sized classrooms totaled 564. Those students who did qualify for the free-reduced lunch program and were in large sized classrooms totaled 248.

There were significant achievement gains by students in small classrooms who did not qualify for the free-reduced lunch program (significant at the .007 level on NCE change and significant at a .059 level on TRESS change). Consequently, small class size does not appear to be an equity factor for socioeconomic status. Students in the free-reduced lunch program started lower on pretests and did not make as many achievement gains as those students who did not receive free-reduced lunch. This is contrary to results in Project Challenge, implemented in Tennessee to raise achievement by lowering class size in the poorest school districts in the state.

In Project Challenge (1990 to 1993), Tennessee implemented smaller class sizes in 17 of the state's poorest school districts, targeting those schools with the lowest per capita income and highest proportion of students in the subsidized school lunch program. The results of this effort were evaluated by examining the effect on the ranking of the school districts according to student performance on a statewide achievement test. The 17 Project Challenge districts moved from near the

bottom of school district performance in Tennessee to near the middle in both reading and math for second grade (Nye et al., 1992).

6. Is there a gender difference between small and large classrooms and student achievement?

In small classrooms, there were 365 male students and 365 female students in the study. Large classrooms had numbers of 502 male students and 488 female students. When using ANOVA to determine if class size had a significant effect on male vs. female student achievement, female gains were found to be significant on the NCE (.031) and at a significance level of .059 on TRESS.

Female students started out with mean average TRESS of 418.8 and an average of 49.9 NCE in small size classrooms, whereas males in small classrooms started out at 413.9 TRESS and NCE score of 47.1. In large classrooms, females started out at TRESS of 425.3 and NCE score of 50.6, and male students started at 421.9 TRESS and NCE score of 48.6. The post TRESS and NCE scores were 462.4 and 57.3, respectively, for females and 456.1 and 54.5, respectively, for males in small sized classrooms. The post TRESS and NCE scores were 463.2 and 56.8, respectively, for females and 458.4 and 54.8, respectively, for males in large sized classrooms.

Gains for both males and females were greater in small classrooms than in large classrooms. In analyzing data for researching question six, it would appear that students were placed in smaller vs. larger classrooms based on reading ability in that pretest scores were lower for both males and females in the smaller classrooms. It should be further noted that the posttest scores of small classroom students (both males and females) generally equaled those of their peers in the large classrooms, making a case for smaller class size as an effective intervention for students with reading deficits.

Project Challenge (1990 to 1993), in Tennessee, was one such project that implemented smaller class sizes in 17 of the poorest school districts in the state to try to improve reading achievement scores. The results of this effort were evaluated by examining the effect on the rankings of the school districts according to student

performance on a statewide achievement test. An average increase of 5.3 in statewide ranking in reading and 6.6 in statewide ranking in mathematics resulted. The increase amounted to nearly a quarter of a standard deviation gain as a result of the Project Challenge effort. The 17 Project Challenge districts moved from near the bottom of school district performance in Tennessee to near the middle in both reading and mathematics for second grade. In addition, in-grade retention of students was reduced in the Project Challenge districts where smaller classes were implemented (Nye et al., 1992).

In the Butler and Handley (1989) study of the difference in achievement for first grade students taught in small classes, it was found that when they reduced class size from 27 to 20 students the achievement went up (significant at the .01 level in all nine achievement areas). The data also showed that no differences in achievement were associated with the independent variables of gender, race, and socioeconomic status.

7. Do Native American students have greater achievement gains than other students as a result of small class size?

Native American students make up the largest minority group in Bismarck and in the state of North Dakota. In this study, Native American students totaled 52 (7.4%) and Caucasian students were at 651 (92.6%) in small classrooms. In large classrooms, Native American students totaled 57 (5.7%) and Caucasian students were at 950 (94.3%). These numbers do not reflect any other minority populations, as these populations would be too small to draw legitimate statistical conclusions. The total percentage of Native Americans who attend school in the Bismarck Public Schools is comparable to the percentage of students in this study.

In small sized classes, pre TRESS and NCE scores were 401.3 and 38.2, respectively, for Native American students and 417.3 and 50.0, respectively, for Caucasian students. In large sized classrooms, the pre TRESS and NCE scores were 407.4 and 41.7, respectively, for Native American students and 425.8 and 50.3, respectively, for Caucasian students. Post TRESS and NCE scores were 436.9 and 46.0, respectively, for Native American students in small classrooms and

461.1 and 56.7, respectively, for Caucasian students in small classrooms. In large classrooms, the post TRESS and NCE scores were 445.7 and 48.4, respectively, for Native American students and 462.4 and 56.2, respectively, for Caucasian students. ANOVA findings were that TRESS and NCE scores were significant at the .001 level and the .003 level, respectively, for Caucasian students in small sized classes.

The fact that Caucasian students benefited more from small class size in the Bismarck Public Schools study was contrary to what was discovered in the SAGE study in Wisconsin, where minority students (African-American) received greater benefits from small sized classrooms. Although African-American SAGE students scored lower on a pretest than African-American students in comparison schools, they made significantly larger gains and surpassed achievement by African-American students in comparison schools on the posttests. On the Comprehensive Test of Basic Skills (CTBS), African-American students had mean scores of 17.6 higher when in SAGE classrooms, significant at the .05 level (Halbach et al., 2001; Molnar et al., 2000). In the Bismarck Public Schools study, small class size does not appear to be the minority equalizer in reading achievement at the primary level that it was for African-American students in Wisconsin.

Both Caucasian and Native American students in small sized classrooms started out with lower pretest scores than students in the large classrooms. Caucasian students caught up during the year when in small sized classrooms; Native American students in small classes did not advance at an equally impressive rate, but stayed behind in reading achievement scores.

Tennessee's Project STAR, begun in 1985, found just the reverse of the findings in this Bismarck Public Schools study. Project STAR found that the positive achievement effect of smaller classes on minority students was double that for majority students, initially, and then was about the same (Word et al., 1990). It should be noted that, in this study, minority (Native American) students had greater TRESS and NCE gains than non-Native American students in the large classrooms,

further complicating the nationwide discussion on minority students and the need for smaller class sizes.

### Discussion of Teacher Survey

Have you incorporated new techniques in reading instruction into your classroom in the past year? If you incorporated new techniques in reading, were they related to class size?

Both teachers in small and large classrooms answered affirmative to this at a rate of 83.3% (small) and 80% (large). There was not a significant difference in the two teacher groups on this question. To the follow-up part of the question, teachers of small sized classes answered yes 38.6% of the time. Teachers of large sized classes answered yes 66.7% of the time.

Class size reduction measures that simply reduced the number of students per teacher and hoped that positive student achievement would result was the case in both the Project STAR and Lasting Benefits Study in Tennessee. Teachers of small classes received no special instructions or training; the outcomes result from lowered class size and from whatever perceptions and advantages accompany having substantially fewer students in a room with one teacher (Finn & Achilles, 1998).

Other class size reduction measures, such as the Burke County Project in 1991 (Achilles et al., 1995), have included professional development activities covering instruction and assessment, and so the effects are not necessarily simply a function of reducing class size. Based on the high percentage of Bismarck, North Dakota, teachers who said they had incorporated new techniques in reading instruction during the past year, this variable would also have been critical to reading achievement results, along with class size. The fact that teachers in large classrooms in this study incorporated new reading techniques as a result of class size more often than teachers of small sized classes may help account for the lack of improved achievement results by students in small classrooms.

What is the biggest advantage of being a teacher of small class size or the biggest disadvantage of being a teacher of large class size?



The following eight general themes resulted from teacher responses to the survey:

1. More quality time can be spent with each child so no one is allowed to slip by.
2. Children get more time to respond and practice skills.
3. Teachers can individualize for each student.
4. Teachers have more time to prepare for small group activities, so a variety of teaching methods can be used, such as small group activities, manipulatives, experiments, and large motor movement.
5. Classroom management is easier; less noise, safer environment; and less time is wasted on discipline so more time can be spent on teaching and learning.
6. Material can be covered faster and more in depth; fewer papers to correct and analyze.
7. Increased sense of belonging; better sense of community; more time to develop social skills; more time to connect with parents through conferences and phone calls.
8. There is more time to connect with specialists who work with students.

The second part of the teacher survey was based on graduated scale responses which included 17 items relating to time allotment and classroom direction. Through the use of analysis of variance (ANOVA), responses from teachers in small classrooms and large classrooms were compared. The following items were found to be statistically significant at the .05 level when the teacher response comparisons were made:

1. Does your teaching lean more toward the particulars of the students as individuals or more toward the particulars of the class in general? (.016 significance level).
- 4b. Time spent disciplining the class (.008 significance level).
- 4d. Time working with small groups (.042 significance level).
- 4g. Time spent with students one on one (.017 significance level).

4j. Time spent working with students on special projects (.004 significance level).

4k. Time spent developing creative projects for your class (.008 significance level).

The last question on the teacher survey was for the teachers to indicate on which two areas they wished that they could spend more time. Teachers in small classrooms indicated preferences, in order, for developing creative projects for their class, working with students on special projects, engaging students in hands-on activities, and working with small groups of students. Teachers in large classrooms indicated that working with students one on one is their most preferred way to spend additional time, followed by working with small groups of students, and engaging students in hands-on activities. Similar to the Bismarck Public Schools teacher survey are the results of studies from Wisconsin, Burke County in North Carolina, Utah, and New York City.

The quasi-controlled, five-year SAGE teacher study in Wisconsin has demonstrated, through evaluation, that teachers of smaller classes report an overall reduction in discipline problems. All SAGE teachers noted increased instructional time, more time for individualization, and more flexibility in choosing among instructional strategies that keep students actively engaged in learning. The study is continuing to find that smaller classes provide high levels of classroom efficiency, a positive classroom atmosphere, expanded learning opportunities, and enthusiasm and achievement among both students and teachers. Most importantly, SAGE teachers stated that class size reduction increases the likelihood of reaching grade-level objectives and covering the content in more depth (Halbach et al., 2001).

This was also true in the Burke County, North Carolina, study where classroom time devoted to instruction went up from 80% to 86%, and the percentage of time devoted to non-instructional activities such as discipline decreased from 20% to 14% as a result of smaller classes (Egelson & Harman, 2000).

Utah has funded CSR since 1990, including targeting some of their resources to low-income students and to allowing districts some flex in how they spend their money. A 1997 study of five districts discovered that the most successful school combined CSR with teacher development, instructional improvement, and productive use of personnel and resources (Evans-Stout et al., 1997).

The Independent Educational Priorities Panel completed a study of the first year of the Class Size Reduction Program in New York City. Among improvements reported as a result of smaller classes were noticeable declines in the number of disciplinary referrals, improved teacher morale, a focus on prevention rather than remediation, and higher levels of classroom participation by students (Haimson, 2000).

Whereas previous studies cited showed a correlation between teacher attitudes/behaviors and higher student achievement in small sized classrooms, this study found teachers reporting the same behaviors and perceptions but limited student achievement gains in the smaller classrooms.

#### Limitations

1. This study was limited in the depth of analysis of reading achievement. Whereas the assessment was limited to the total reading score (TRESS), an assessment of the subtests with reading (vocabulary and comprehension) might have revealed additional information as to the effect of class size.
2. This study was limited in that it did not differentiate between teachers, their students' achievement, and the teacher survey. Because of this limitation, teacher effectiveness and background could not be linked.
3. This study was limited to only pre-post testing using the Gates-MacGinitie Reading Test. Other variables such as attendance, discipline, and teacher quality were not considered.
4. The placement of students in smaller classes does not appear to have been random, causing the pretest data of smaller classrooms to be lower and creating a potential for inflated gains on posttest scores.

### Conclusions

The class size reduction initiative in Bismarck started in 1999 when Bismarck Public Schools received a CSR grant award of \$411,209. This money was used to hire 12 teachers, some of whom were classroom teachers; some were specialists supporting regular classroom teachers and students. The school district, in fulfillment of its responsibility to provide assessment for the general effectiveness of this CSR effort, committed to administering the Gates-MacGinitie Reading Test (Fourth Edition) in the Fall (September) and again in the Spring (April) in order to compare and measure achievement gains between students in large classrooms (22 or more) and students in small classrooms (17 or less). The GMRT assessment of the CSR has continued over the past three years (1999-02).

This study was undertaken to measure the class size reduction efforts in the Bismarck Public Schools over the past three years. Logical questions of school board members, staff, and community members included "Was the money spent to reduce class size in the primary grades over the past three years well spent? How should the district proceed with CSR issues when making future budgetary decisions in the best interest of students and teachers? When the CSR money was no longer strictly specified for reducing class size, should federal block grant money continue to be used for reducing and/or keeping class sizes small, or would the money be better spent in other ways to raise student achievement?"

The hypothesis of this study was that lowering class size would positively correlate with increased student achievement. This was not found to be true. It was also hypothesized that first graders would make greater gains from having been in smaller classes than the gains made by second or third graders. This also proved not to be the case.

Based on a review of selected literature, it would be expected that significant gains would be made by the students who qualified for free-reduced lunch who were in small sized classrooms. An advantage to small class size, as asserted in the literature, is the equalizing effect among economically disadvantaged students. This

was not the case either, as students who did not qualify for the free-reduced lunch program made greater achievement gains than those who did qualify.

In past studies, there was not usually a gender difference in achievement between class sizes. Once again, the results varied from the literature, as female students in this study made significantly greater gains in small sized classrooms than did male students.

Another question posed by administrators in the Bismarck Public Schools was "Would Native American students benefit more from small class size, as minority populations in other areas of the country had shown?" The results in this study were to the contrary, with Native American students making greater achievement gains in the large classrooms. Caucasian students, when compared to Native American students, advanced significantly more in reading achievement after having been placed in small sized classrooms.

The survey results were similar to other teacher survey results in other states. Reduced class size is necessary but not sufficient to increase achievement scores. Teacher quality is more important than class size. In an analysis of teacher quality in Texas, it was found that by just looking at the variations in student performance that arise from differences in teacher quality within a typical school, the variation is large. Moving from an average teacher to one at the 85th percentile of teacher quality (moving up one standard deviation in teacher quality) implies that the teacher's students would move up more than seven percentile ranks in a year. This would also imply that a one standard deviation change in teacher quality leads to a .18 standard deviation in student achievement. This compares to one standard deviation reduction in class size where a .01-.03 standard deviation improvement in student achievement is typical (Rivkin, Hanushek, & Kain, 2000). Teacher quality and teaching methods are as much a factor, if not more a factor, than simply reducing class size. Continuing with small classes may or may not be the right thing to do as long as resources allow it, but professional development in the area of reading instruction must also be considered to make small classrooms a financially viable alternative to raising student achievement.

This study sought to demonstrate that class size reduction in reading would result in student achievement above that of normal classrooms within the primary grades in the Bismarck Public Schools. The results of this study show no significant difference in student achievement gains in reading between small and large classrooms.

#### Recommendations for Action

The findings of this study illustrate the need for ongoing assessment, so that decisions can be based on factual, research-based information rather than simply generally held beliefs of educators. Although a number of studies have shown that lowering class size increases student achievement, the most notable being Project STAR in Tennessee and the SAGE study in Wisconsin, this proved not to be the case in the Bismarck Public Schools. The variable that seemed to make the difference in national studies was the teacher and professional development in reading instruction for the teacher. The Bismarck Public Schools are currently moving in the direction of a unified and more extensive professional development plan. The past and current administration and school board plan is to continue with reduced sized classrooms at the elementary level. Reading specialists in each elementary building have also been the norm in the Bismarck Public Schools for the past decade. School board president, Sonna Anderson, was quoted in the Bismarck Tribune as saying, "The school board is committed to making elementary classrooms as small as possible with what we can afford. When there are way too many children per classroom, it's difficult for the teachers and unfortunate for the children. We want to provide education that is effective" (Van Dyke, 2002, p. B2). With the leadership of the school board chairperson at this time, the district plans on sustaining small class size at the primary level. The results of this study raise questions as to that decision. Staff development providing teachers with the opportunity to learn strategies that are proven effective in teaching in small class size settings might be a prudent first step in the board's effort to improve student achievement.

Class size reduction works best when coupled with professional development opportunities for teachers. Educators should be trained in new teaching techniques that take advantage of smaller class sizes. Chester Finn of the Hudson Institute think tank, when asked about President Clinton's initiative, reported that for \$12 billion you could retrain today's teachers so they knew their subjects. You could give each of the nation's 2.7 million teachers a \$1,000 tuition grant to learn effective techniques for teaching reading (Finn, 2002).

Economists measuring the effect of a \$500 investment per student found that spending the money on teacher education had the greatest impact on student achievement. Lowering the student-teacher ratio was found to have a smaller effect than increasing teacher education, experience, and salaries (Stecher et al., 2001). In Bismarck Public Schools, there are plenty of classrooms and an adequate pool of qualified elementary teachers to choose from when there is a teacher vacancy. The question in Bismarck becomes determining the best use of limited resources. Possibly the issue of professional development should be looked at with as much intensity as class size.

The opportunities that smaller classes provide in the areas of curriculum, instruction, and assessment must be approached differently. The following is a list of suggestions that come from Finn (2002) and Brophy (2000):

1. Mentor new teachers and provide "current best practices" for teachers who could benefit from updating and relearning skills that work well with small classes but were too arduous to practice in large classes.
2. Cover curriculum both in broader and greater depth and take advantage of the increased sense of community that is typical of small classes. Make sure that the content is coherent, structured, and connected.
3. Support specific needs of small class teachers and follow the principles of effective programming.
4. Help students become better students in small classes by offering more engaged learning, broader range of high-quality learning activities and assignments, more social activities, more student pressure to participate, more individualized

monitoring and assistance, initiate intimate seating, utilize practice and application activities, provide a greater sense of community, and more opportunities for students to be responsive to other students in the class.

5. Provide a supportive classroom climate where the teacher knows each child and his or her parents.

6. Shift the emphasis from recitation and assessment of knowledge to discussion and construction of knowledge.

7. Use goal orientated assessment where the teacher can interact more often, including laboratory activities, essay assignments, portfolios, swift teacher feedback, and remediation, when necessary.

8. Hold students accountable to the achievement expectations of the class.

Another recommendation that should be considered is an insightful economic theory for class size laid out by Lazear (1999). Lazear states that students who attend a smaller class learn more because they experience fewer student disruptions during class time, on average. Lazear then assumes that disruptions require teachers to suspend teaching, creating a "negative externality" that reduces the amount of learning for everyone in the class. It is possible that students who spend time in small classes learn to behave better with close supervision, leading to a reduced propensity to disrupt subsequent classes (Lazear, 1999). If Lazear's theory is correct, then the "optimal" class size is larger for groups of students who are well behaved. Schools, therefore, have an incentive to assign weaker, more disruptive students to smaller classes. This concept would be using class size reduction as a remedial strategy for those students with reading difficulties.

#### Recommendations for Further Study

Based on the research conducted through this study and from research conducted by Brophy (2000), Finn (2002), Lazear (1999), Stecher et al. (2001), and other researchers of class size, the following recommendations for further study should be considered. Bismarck Public Schools, to increase student achievement, may choose to implement some or all of the following recommendations for further study.



1. A study should be conducted using attendance data when measuring achievement gains. This study used only students with pre-post test scores when analyzing reading achievement. Student mobility and other attendance variables likely have an impact on achievement performance of students in primary classrooms.

2. A study should be done to determine the types of staff development programs that are most likely to augment the benefits of class size reduction. What are the best instructional practices for small classes?

3. A longitudinal study should be undertaken with the leadership within the district to match teachers with student achievement in these small vs. large classrooms. Which classrooms had the most improved reading scores from Fall until Spring? Which classroom had the weakest reading achievement score gains? What reading strategies are working? What strategies appear not to be as strong? By answering some of these questions, teachers could be mentored, professional development could be offered, and new proven reading and teaching methods could be implemented to improve student achievement.

4. A study could be conducted to see if smaller classes improve student behavior. Are there fewer incidences of vandalism, fewer behavioral referrals within and outside of school, and fewer classroom disruptions?

5. A qualitative study should be undertaken to assess the effect of small class size on student motivation, student participation and engagement with the school, the development of interpersonal skills, and a sense of self. These are areas that could be compared in a study of small vs. large classrooms.

6. A study should be conducted to improve the effectiveness of the CSR Program by integrating and aligning it with other programs and reforms in the district.

After researching the effects of class size on student performance in reading in the primary grades, there are still questions associated with achievement benefits and implementation costs. There will be a continued need to study possibilities for improving the performance of students in the primary grades and in later grades as well. Effective teaching strategies to maximize the effectiveness of teaching reading

will need continued research. Reducing class size is not the sole answer to improving student achievement, but when it is combined with effective teaching strategies, students are likely to benefit academically. Issues of class size must continue to be looked at objectively, not impetuously, but with an eye to matters of cost effectiveness, if they are to be understood accurately and implemented effectively.

## APPENDICES

**APPENDIX A**  
**FIRST AND FOLLOW-UP LETTER TO TEACHERS**

1512 River Drive  
Mandan, ND 58554  
April 2002

Dear Teacher:

I am a counselor at Wachter Middle School and a doctoral student in the Department of Educational Leadership at the University of North Dakota. I am conducting research for my dissertation on issues related to the impact of class size on the reading achievement of students in the primary grades. You have been selected as a subject for this study. As such, I would like to ask you to complete the enclosed survey that is focused on teachers' perceptions of selected elements of teaching and learning. The data from this survey will be used in my dissertation, along with data on student achievement in reading that will be obtained from the Bismarck School District. There will be no link between your responses and scores from students you taught last year as the achievement test data will not be identified by school, teacher, or student.

Your participation in this study is voluntary. There will be no sanctions possible against any person who chooses not to complete the survey instrument. However, I do ask for your cooperation so that the findings from the study may provide valid and reliable data that might be used by district officials as they consider the implications of class size on student achievement.

The survey instruments are color-coded to identify if the respondent taught in a large (>22 students) or small (<18 students) classroom last year. The return envelope is coded to identify those who respond to the survey. I will make a follow-up contact with persons who do not return the survey and the code will help me to identify those individuals. The envelopes and the completed surveys will be separated as soon as opened so that I will not be able to identify which survey was completed by which subject. In addition, the list of subject codes and respondents will be destroyed after the follow-up activities have been concluded. So, your responses will be anonymous and your participation will be confidential. By returning a completed survey, I will assume that you are granting me permission to aggregate your responses with those of others and to report those findings in my study.

If you have any questions or concerns regarding this study, please feel free to contact me or my adviser, Dr. Jerry Bass. Contact information is provided below. I hope that you will be willing to participate in this study and that you will return the completed survey to me within the next two weeks. Thank you for your time and attention in support of my project!

Sincerely,

Dotti Dixon Schmeling  
Wachter Middle School  
221-3585  
dotti\_dixon@educ8.org

Dr. Jerry Bass  
University of North Dakota  
(701) 777-4940 or (701) 777-3577  
gerald\_bass@und.nodak.edu

1512 River Drive  
Mandan, ND 58554  
April 2002

Dear Teacher:

I am a counselor at Wachter Middle School and a doctoral student in the Department of Educational Leadership at the University of North Dakota. I am conducting research for my dissertation on issues related to the impact of class size on the reading achievement of students in the primary grades. You were selected as a subject for this study and should have received an earlier mailing in this regard. As of the date of this letter, I had not received a response from you. If you did complete the survey and mail it back to me, let me simply say "Thank you" and ask you to disregard this mailing. If you have not responded to the earlier survey, I would like to ask you to complete the enclosed survey and return it to me as soon as possible. As noted in the previous mailing, the data from this survey will be used in my dissertation, along with data on student achievement in reading that will be obtained from the Bismarck School District. There will be no link between your responses and scores from students you taught last year as the achievement test data will not be identified by school, teacher, or student.

Your participation in this study is voluntary. There will be no sanctions possible against any person who chooses not to complete the survey instrument. However, I do ask for your cooperation so that the findings from the study may provide valid and reliable data that might be used by district officials as they consider the implications of class size on student achievement.

The survey instruments are color-coded to identify if the respondent taught in a large (>22 students) or small (<18 students) classroom last year. The return envelope is coded to identify those who respond to the survey. I will make a follow-up contact with persons who do not return the survey and the code will help me to identify those individuals. The envelopes and the completed surveys will be separated as soon as opened so that I will not be able to identify which survey was completed by which subject. In addition, the list of subject codes and respondents will be destroyed after the follow-up activities have been concluded. So, your responses will be anonymous and your participation will be confidential. By returning a completed survey, I will assume that you are granting me permission to aggregate your responses with those of others and to report those findings in my study.

If you have any questions or concerns regarding this study, please feel free to contact me or my adviser, Dr. Jerry Bass. Contact information is provided below. I hope that you will be willing to participate in this study and that you will return the completed survey to me within the next two weeks. Thank you for your time and attention in support of my project!

Sincerely,

Dotti Dixon Schmeling  
Wachter Middle School  
221-3585  
dotti\_dixon@educ8.org

Dr. Jerry Bass  
University of North Dakota  
(701) 777-4940 or (701) 777-3577  
gerald\_bass@und.nodak.edu

APPENDIX B  
TEACHER QUESTIONNAIRE

---

# Teacher Survey

## Teacher Information:

Have you incorporated new techniques in reading instruction into your classroom in the past year?

Y N

If yes, why? \_\_\_\_\_

\_\_\_\_\_

If you incorporated new techniques in reading, were they related to class size?

Y N

If yes, please explain. \_\_\_\_\_

\_\_\_\_\_

## Class Information:

What are the biggest advantages to a teacher of small class size? (< 17 students)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

What are the biggest disadvantages to a teacher of large class size? (>22 students)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_



## Class-size Classroom Study

Below are questions designed to determine how teachers allot their time and also where they direct their efforts. Circle the number that best describes you and your classroom.

**1. A good teacher adjusts instruction according to the particulars of the class. Thinking about such student qualities as interests, abilities, needs, and personalities, would you say your teaching leans more toward the particulars of the students as individuals or more toward the particulars of the class in general?**

students as individuals-----class in general  
 1 2 3 4 5

**2. Teachers vary in enthusiasm from year to year. How would you characterize your enthusiasm for teaching reading?**

a. 2000-2001--one year ago?

not enthusiastic-----very enthusiastic  
 1 2 3 4 5

b. 2001-2002--this past year?

not enthusiastic-----very enthusiastic  
 1 2 3 4 5

**3. Teachers make decisions about their approach to curriculum content, specifically about its depth and breadth.**

a. Thinking about reading content, what would you estimate to be your students' depth of understanding at the end of last year?

not deep-----very deep  
 1 2 3 4 5

b. How much reading content do you typically cover in a year?

cover the grade level content-----work into the next grade level  
 1 2 3 4 5

**4. Budgeting time is another dilemma for teachers. Typically, how much time do you spend on each of these tasks?**

- a. Diagnosing the needs of individual students?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- b. Disciplining the class?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- c. Providing help for individual students?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- d. Working with small groups of students?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- e. Assessing the progress of individual students?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- f. Assessing the progress of the class in general?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- g. Working with students one-on-one?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- h. Disciplining individual students?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5
- i. Engaging students in hands-on activities?  
 a very little amount of time-----a very large amount of time  
 1                      2                      3                      4                      5

- j. Working with students on special projects?  
a very little amount of time-----a very large amount of time  
1 2 3 4 5
- k. Developing creative projects for your class?  
a very little amount of time-----a very large amount of time  
1 2 3 4 5
- l. Using direct instruction, meaning that you explain and give information to the class, model, practice, provide feedback, etc.  
a very little amount of time-----a very large amount of time  
1 2 3 4 5

**Indicate the two areas from a. through l. on which you wish you could spend more time.**

Area #1: \_\_\_\_\_

Area #2: \_\_\_\_\_

**APPENDIX C**  
**AUTHORIZATION FROM BISMARCK PUBLIC SCHOOLS**  
**TO CONDUCT RESEARCH**



November 15, 2001

**Bismarck Public Schools**

400 Avenue E East  
Bismarck, North Dakota 58501  
(701)221-3700  
Fax: (701)221-3711  
www.bismarck.k12.nd.us

**DR. PAUL K. JOHNSON**  
SUPERINTENDENT

**JOHN M. SALWEI**  
ASST. SUPERINTENDENT  
SECONDARY SCHOOLS  
CURRICULUM INSTRUCTION

**RICHARD J. BURESH**  
ASST. SUPERINTENDENT  
ELEMENTARY SCHOOLS  
STUDENT SERVICES

**EDWIN J. GERHARDT**  
BUSINESS MANAGER

**LISA J. KUDELKA**  
HUMAN RESOURCES MANAGER

RE: Authorization for Research

Dear Ms. Schmeling,

Our office has received your proposal to conduct research on "The Effect of Class Size on Student Performance in Reading in the Primary Grades". We share your interest in this study and authorize you to proceed with our full cooperation and encouragement. We understand that we will be providing anonymous data on student performance on reading achievement of students from selected small and large classrooms and that the data will contain no information that could be used to identify individual students. Please let us know how we can be of assistance.

We look forward to hearing about the results of your study.

Sincerely,

A handwritten signature in black ink that reads "Rick Buresh". The signature is written in a cursive style with a large initial "R".

Rick Buresh  
Assistant Superintendent

**APPENDIX D**  
**LIST OF ALL TEACHER RESPONSES TO OPEN-ENDED**  
**SURVEY QUESTIONS BY THEME**

The following were the verbatim teacher responses taken from the teacher surveys when teachers were asked to respond to the question, "What is the biggest advantage of being a teacher of small class size or the biggest advantage of being a teacher of large class size?" It was from these open-ended responses that the eight general themes emerged.

**1. More quality time can be spent with each child so no one is allowed to slip by.**

"The amount of time you spend on each student is important. When you have a large class the time you spend on each child is considerably less. At the age 1st graders need a lot of direction, encouragement, & motivation."

"You can do more help for students."

"You don't get to work one-on-one very often."

"You get to spend time with each student . . ."

"More individual attention"

"There is more opportunity for one to one interaction between teacher & student. Less management and behavior issues."

"Being able to work with students individually."

"Not getting to personally listen to all of them read at least once a day."

"Less time to spend with each student individually."

"It is much easier to give each child one on one help."

"You can get to each child on a one-one basis and not just the ones that are the neediest"

"--more attention to the students"

"teacher has a lot more evaluating/preparing/correcting to do and can't spend time helping students individually."

"more individual attention"

"When I had >22 students I always felt I 'missed' the average kids--the talented could work on their own, the needier ones needed all my time."

"I've never had <17. Not enough time to work with students individually."

"Not able to spend the time that is necessary to help those students that struggle or with those that excel above & beyond what is expected."

"harder to get time for each student"

--Being able to work on one with each child each day."

--not as much one on one instruction"

"more one to one contact"

"Isn't it obvious?! \*I don't know why admin. doesn't 'GET IT!' Smaller class sizes would prevent a lot of learning prob. in later grades. We're building a foundation!"

"You can work with each student more during the day even if it when you are walking around the room."

## **2. Children get more time to respond and practice skills.**

"Not able to get to each child."

"When you have a large class the time you spend on each child is considerably less."

"I am able to listen to children read every day."

"You get to spend time with each student to listen to them read, Help them with math ect . . . It is just a nice size to have."

--feel more able to reach all of them (I've never had 17 or fewer students in the classroom)"

"able to help them more"

"Getting a chance to listen and work with individual student more than once a day in all areas of study"

"You can't work with each one as much as you want or need to."

"being able to work one on one with each child each day."

"You're able to have more quality time with students"

"At the end of the day I feel like I haven't spent time with some students who are independent workers."

## **3. Teachers can individualize for each student.**

"A teacher feels she's neglecting the more needy students or not challenging the more gifted"

"Safe environment that students thrive & learn to their ability, also need to be more structure."

"Knowing a student better also helps adjust expectations of that student"



"It's tough to meet all their needs, on all levels, with a large number of students."

"You can do more help for students."

"It is more difficult to reach all the levels of intelligence and readiness."

"More personality, ability, & needs differences,"

"You are able to help those students that struggle with decoding skills and reading comprehension as well as being able to allow those that excel be able to move forward"

"There are always new and exciting programs/ideas to try . . ."

"Can only get to the needy children"

"The most important is the increased contact time with each student--they received a much more individualized instruction."

"\*individual attention to fit needs"

"--being able to spend time with all of them equally--not let some slip by--not being able to work with them on their own level"

"I have so many different learning abilities in my room I have a big class--feel I can't reach them all by teaching to all at the same time."

"able to help them more"

"interests, abilities, needs & personalities also impacts the rate & quality of teaching/learning"

"harder to get time for each student"

"It is difficult to read a story w/25 kids, not everyone reads enough."

"I have 25 students of all reading levels. I need to be able to reach all my students."

"Really getting to know your students and their strengths and weaknesses"

"At the end of the day I feel like I haven't spent time with some students who are independent workers."

"--I have 15 student.--Six w/extremely high needs--no classroom aide (HELP)"

"Because it's [Guided Reading] small group instruction, it works well in a large class because you take 4-5 [students] at a time. It also allows for individual levels for everyone."

"--the talented could work on their own, the needier ones needed all my time."

**4. Teachers have more time to prepare for small group activities, so a variety of teaching methods can be used, such as small group activities, manipulatives, experiments, and large motor movement.**

"More time to prepare activities"

"\*hands on in depth projects can be tackled with greater room & ease"

"Special projects are easier to do with a smaller group."

"Tried to incorporate different learning modes as I had an individual who was more of a visual learner"

"More small group work, easier to do more hands on activities"

"Physically, the classroom is too crowded & they're in each other's space."

"Also I was much less likely to engage in manipulative activity or movement in the room--the logistics simply did not permit it."

"It's easier to have multiple small group activities do to physical space class management, etc. It's easier to do whole group activities such as science experiments--17 can get closer to you than larger numbers of students."

"Can't do as much"

"Sometimes a large class limits activities that would work well with a smaller group."

"not enough (room) space in class"

"Lack of classroom space."

"Smaller class size lends itself more easily to exploring new techniques."

"to increase reading other than using the basal, for variety"

"I couldn't do what I usually do or try new techniques because my class size was large, & we physically couldn't do a lot of things because of lack of room!"

"to include whole language texts"

"I always like to try new techniques."

"Since I have taught a long time I like to keep up with current trends. We added more computer activities such as the internet."

"--customizing student learning"

"On an elementary level/I see no advantage to large class size./Some projects are not even tackled with a larger class that are more conducive to smaller numbers. ie) art projects/some field trips/even prepared special 'Units' of study"

**5. Classroom management is easier; less noise, safer environment; and less time is wasted on discipline so more time can be spent on teaching and learning.**

“Safe environment that students thrive & learn to their ability, also need to be more structure.”

“Less noise in the classroom. Out of the 40 classes I’ve had only about 5 were ever 17 or less.”

“It’s easier to have multiple small group activities to to physical space class management, etc.”

“--students seem to misbehave more in larger groups”

“Class is quieter”

“behaviors of students also play a part”

“I have a small enough group to be able to manage it. Class management, group size, & time.”

“Noise, general hustle-bustle is on going.”

“Easier to maintain control in classroom & outside of school.”

“Less management and behavior issues.”

“Classroom management gets more challenging.”

“Classroom management.”

**6. Material can be covered faster and more in depth; fewer papers to correct and analyze.**

“More papers to check”

“It is more difficult to keep track of student’s work also finding time to work with individual students.”

“\*”move” farther & faster in material covered”

“\*hands on in depth projects can be handled with greater room & ease”

“--being able to be truly effective in teaching for understanding”

“I would do it based on good teaching strategies, class size may affect the extent of implementation”

“Increased paperwork.”

"More of everything--more conferences, more papers to grade, more desks, more IEP's."

"Teacher has a lot more evaluating/preparing/correcting to do and can't spend time helping students individually."

"Also, correcting papers & providing specific feedback is more difficult [in a large classroom]."

"Our themes are developing and becoming easier to do because of smaller class size. I'm able to help students develop their projects."

**7. Increased sense of belonging; better sense of community; more time to develop social skills; more time to connect with parents through conferences and phone calls.**

"A lot of social skills need to be developed at this age."

"You're able to have more quality time with students"

"Parent-teacher conferences are rushed & stressful & scheduling of students in Title I reading, speech, LD, & counseling services is a nightmare!"

"Each child has the opportunity to interact more often, or for longer time periods, with the teacher."

"Not being able to get to spend time with every student to help them achieve success. Always feeling like you left someone behind each day by not being able to get to them."

"Children don't get the personal attention they may need--"

"More parents to deal with"

"Getting a chance to listen and work with individual students more than once a day in all areas of study"

"More conferences & phone calls."

"With large class size the students & teacher both suffer!"

"Knowing a student well enough to be able to recall specific things when visiting with parents at conferences! Knowing a student better also helps adjust expectations of that student."

"--getting to know each student's individual personality."

"17 can get closer to you than larger numbers of students."

"\*greater cooperation with smaller class size"

"The feeling of camaraderie isn't as strong [in a large class]."

"I feel with a small class, there is an increased sense of belonging, a better sense of 'community' & the teacher gets to know each student & their parents better."

**8. There is more time to connect with specialists who work with students.**

"There are more reading problems than I can deal with--so I have to come up w/ways to reach all students effectively!"

"You will probably need to work with more specialists working with your students, and so scheduling their academics and special needs is more complex."

"More of everything--more conferences, more papers to grade, more desks, more IEP's."

"Parent-teacher conferences are rushed & stressful & scheduling of students in Title I reading, speech, LD, & counseling services is a nightmare!"

## REFERENCES

## REFERENCES

- Achilles, C. M. (1996). Summary of recent class-size research with an emphasis on Tennessee's Project STAR and its derivative studies. Nashville: Tennessee State University, Center of Excellence for Research on Basic Skills.
- Achilles, C. M. (1997). Small classes: Big possibilities. The School Administrator, 54(10), 6-15.
- Achilles, C. M. (1998a, April). If not before: At least now. Paper presented at the meeting of the American Educational Research Association, San Diego, CA.
- Achilles, C. M. (1998b, February). Small-class research supports what we all know (So, why aren't we doing it?). Paper presented at the meeting of the American Educational Research Association, San Diego, CA.
- Achilles, C. M., & Finn, J. D. (2000). Should class size be a cornerstone for educational policy? The CEIC Review, 9(2), 15, 23.
- Achilles, C. M., Finn, J. D., & Bain, H. P. (1997/1998). Using class size to reduce the equity gap. Educational Leadership, 55(4), 40-43.
- Achilles, C. M., Finn, J. D., & Pate-Bain, H. (2002). Measuring class size: Let me count the ways. Educational Leadership, 59(5), 24-26.
- Achilles, C. M., Harman, P., & Egelson, P. (1995). Using research on class size to improve pupil achievement. Research in the Schools, 2(2), 23-30.
- Achilles, C. M., Kiser-Kling, K., Owen, J., & Aust, A. (1994). Success starts small: Life in a small class. Greensboro: University of North Carolina at Greensboro.
- Achilles, C. M., Sharp, M., & Nye, B. (1998, February). Attempting to understand the class size and pupil-teacher ratio (PTR) confusion: A pilot study. Paper presented at the meeting of the American Association of School Administrators, San Diego, CA.
- Angist, J. D., & Lavy, V. (1996, July). Using Maimonides' rule to estimate the effect of class size on children's academic achievement. Mt. Scopus, Jerusalem: Hebrew University.
- Bain, H. P., & Achilles, C. M. (1986). Interesting developments on class size. Phi Delta Kappan, 67(9), 662-665.

Bain, H. P., Achilles, C. M., Dennis, B., Parks, M., & Hooper, R. (1988, Fall). Class-size reduction in metro-Nashville: A three-year cohort study. ERS Spectrum, 6, 30-36.

Bain, H. P., Achilles, C. M., & Witherspoon-Parks, M. (1988, November). Three-year longitudinal study of small class size: The metro-Nashville public schools study: 1984-1987. Paper presented at the meeting of the Mid-South Educational Research Association, New Orleans, LA.

Bain, H. P., & Jacobs, R. (1990, September). The case for smaller classes and better teachers. Streamlined Seminar, 9(1). (Available from the National Association of Elementary School Principals, 1615 Duke Street, Alexandria VA 22314)

Barnett, W. S. (1985). Benefit-cost analysis of the Perry preschool program and its policy implications. Educational Evaluation and Policy Analysis, 7(4), 333-342.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. The Future of Children: Long-Term Outcomes of Early Childhood Programs, 5(3), 25-48.

Bell, J. D. (1998). Smaller = better? State Legislatures, 24(6), 14.

Biddle, B. J., & Berliner, D. C. (2002). Small class size and its effects. Educational Leadership, 59(5), 12-22.

Bingham, S. (1994). An examination of small class as a "gap reduction" strategy for achievement differences in groups of students, K-3. Unpublished doctoral dissertation, University of North Carolina, Greensboro.

Boozer, M., & Rouse, C. (1995, May). Intraschool variation in class size: Patterns and implications (Paper #344). Cambridge, MA: National Bureau of Economic Research. (ERIC Document Reproduction Service No. ED 385 935)

Boyd-Zaharias, J., & Pate-Bain, H. (2000). Early and new findings from Tennessee's Project STAR. The CEIC Review, 9(2), 4.

Brewer, D. (1999). Estimating the cost of national class size reductions under different policy alternatives. Educational Evaluation and Policy Analysis, 21(2), 179-190.

Brophy, J. (2000). How might teachers make smaller classes better classes? The CEIC Review, 9(2), 9-10.

Burtless, G. (1996). Introduction and summary. In G. Burtless (Ed.), Does money matter? (pp. 4-42). Washington, DC: Brookings Institution.



Butler, J. M., & Handley, H. M. (1989, March). Differences in achievement for first grade students taught in small classes. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Cahen, L. S., Filby, N., McCutcheon, G., & Kyle, D. W. (1983). Class size and instruction. New York: Longman.

Calhoun, J. B. (1962). Population density and social pathology. Scientific American, 206(2), 139-148.

Card, D., & Krueger, A. B. (1996). Labor market effects of school quality: Theory and evidence. In G. Burtless (Ed.), Does money matter? (pp. 97-140). Washington, DC: Brookings Institution.

Center of School Assessment. (1986, July). PRIME TIME: The relationship between class size and achievement for first grade students in Indiana 1984-85. Indianapolis: Indiana Department of Education.

Chase, C. I., Mueller, D. J., & Walden, J. D. (1986, December). PRIME TIME: Its impact on instruction and achievement. Final report. Indianapolis: Indiana Department of Education.

Congressional Budget Office. (1986). Trends in educational achievement. Washington, DC: Author.

Correa, H. (1993). An economic analysis of class size and achievement in education. Education Economics, 1(2), 129-135.

Council of State Governments. (2002). California: Smaller classes studied. State Government News, 45(3), 8.

Davidson, C. W., Beckett, F. E., & Peddicord, H. Q. (1994, November). The effects of a statewide teacher aide program on functional literacy examination scores of eleventh-grade students. Paper presented at the meeting of the Mid-South Educational Research Association, Nashville, TN.

ECS Information Clearinghouse. (1999). State class size reduction measures. Sacramento: California Department of Education.

ECS Information Clearinghouse. (2002). State class size reduction measures. Sacramento: California Department of Education.

Educational Research Service. (1978). Class size: A summary of research. Arlington, VA: Author.

Educational Research Service. (1980). Class size research: A critique of recent meta-analyses. Arlington, VA: Author.

Egelson, P., & Harman, P. (2000). Ten years of small class size in Burke County, North Carolina. The CEIC Review, 9(2), 14, 18.

Egelson, P., Harman, P., & Achilles, C. M. (1996). Does class size make a difference? Recent findings from state and district initiatives. Greensboro, NC: Southeastern Regional Vision for Education.

Evans-Stout, K., Fleming, N., Johnson, Jr., B. L., Ehrlinger, E., Gray, J. M., Merrill, R., Parker, M., Roberts, R., & Stewart, T. (1997). The use of class size reduction funds in five Wasatch County districts. Salt Lake City: Utah Education Consortium.

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. Harvard Journal on Legislation, 28(2), 465-498.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), Holding schools accountable: Performance-based reform in education (pp. 265-298). Washington, DC: Brookings Institution.

Finn, J. D. (1998, April). Class size and students at risk: What is known? What is next? (OERI Publication No. AR 98 7104). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Finn, J. D. (2002). Smaller classes in American schools: Research, practice, and politics. Phi Delta Kappan, 83(7), 551.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. American Educational Research Journal, 27(3), 557-577.

Finn, J. D., & Achilles, C. M. (1998, April). Tennessee's class-size study: Questions answered, questions posed. Paper presented at the meeting of the American Educational Research Association, San Diego, CA.

Finn, J. D., Fulton, D., Zaharias, J., & Nye, B. A. (1989). Carry-over effects of small classes. Peabody Journal of Education, 67(1), 75-84.

Finn, J. D., & Voelkl, K. E. (1994). Class size. In T. Husen & T. N. Postlethwaite (Eds.), International encyclopedia of education (2nd ed., pp. 770-775). New York: Pergamon Press.

Florida Department of Education. (1998). The relationship of school and class size with student achievement in Florida: An analysis of statewide data. Tallahassee: Author.

Folger, J. (1989). Projects and class size policy. Peabody Journal of Education, 67(1), 1-33.

Gilman, D., & Tillitski, C. (1990). The longitudinal effects of smaller classes: Four studies. Terre Haute: Indiana State University, School of Education.

Gilman, D. A., & Antes, R. L. (1985). The educational effects of the introduction of a state supported program of smaller classes. A study of the first year of Indiana's Project PRIME TIME and its effects on test results (1984-85). A comprehensive analysis. Terre Haute: Indiana State University, School of Education.

Glass, G. V., Cahen, L. S., Smith, M. L., & Filby, N. N. (1982). School class size: Research and policy. Beverly Hills: Sage Publications.

Glass, G. V., & Smith, M. L. (1978). Meta-analysis of research on the relationship of class size and achievement. San Francisco: Far West Laboratory for Educational Research and Development.

Goodlad, J. I., Soder, R., & Sirotnik, K. A. (Eds.). (1990). The moral dimensions of teaching. San Francisco: Jossey-Bass.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. Review of Educational Research, 66(3), 361-396.

Haimson, L. (2000, April). Smaller is better: First-hand reports of early grade class size reduction in New York City public schools. New York: Educational Priorities Panel.

Halbach, A., Ehrle, K., Zahorik, J., & Molnar, A. (2001). Class size reduction: From promise to practice. Educational Leadership, 58(6), 32-35.

Hall, E. T. (1966). The hidden dimension. Garden City, NY: Doubleday.

Hall, E. T. (1976). Beyond culture. Garden City, NY: Doubleday.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. Journal of Economic Literature, 24, 1141-1177.

Hanushek, E. A. (1995). Moving beyond spending fetishes. Educational Leadership, 53(3), 60-64.

Hanushek, E. A. (1996). School resources and students' performance. In G. Burtless (Ed.), Does money matter? (pp. 43-73). Washington, DC: Brookings Institution.

Hanushek, E. A. (1998, February). The evidence on class size. Rochester, NY: The University of Rochester, W. Allen Wallis Institute.

Hanushek, E. A. (1999a). The evidence on class size. In S. E. Mayer & P. Peterson (Eds.), Earning and learning: How schools matter (pp. 143-163). Washington, DC: Brookings Institution.

Hanushek, E. A. (1999b). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. Educational Evaluation and Policy Analysis, 21(2), 143-163.

Hanushek, E. A. (2000, October). The class size policy debate (Working Paper No. 121). Washington, DC: Economic Policy Institute.

Harvey, B. (1993). An analysis of grade retention for pupils in K-3. Unpublished doctoral dissertation, University of North Carolina at Greensboro.

Hedges, L. V., & Greenwald, R. (1996). Have times changed? The relationship between school resources and student performance. In G. Burtless (Ed.), Does money matter? (pp. 74-92). Washington, DC: Brookings Institution.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of differential school inputs on student outcomes. Educational Researcher, 23(3), 5-14.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. Quarterly Journal of Economics, 115(4), 1239.

Hruz, T. (1998). Beyond smoke and mirrors: A critical look at smaller class sizes. Wisconsin Interest, 7(2), 29-37.

Hruz, T. (2000). The costs and benefits of smaller classes in Wisconsin: A further evaluation of the SAGE program. Thiensville: Wisconsin Policy Research Institute.

Iowa Department of Education. (2002). Class size 2001-2002. Des Moines: Bureau of Planning, Research and Evaluation.

Ivey, G. (1999). Reflections on teaching struggling middle school readers. Journal of Adolescent and Adult Literacy, 42, 372-381.

Johnson, K. A. (2000). Do small classes influence academic achievement? What the National Assessment of Educational Progress shows. Washington, DC: The Heritage Foundation.

Kirkbrush, K. (1996). Class size. Madison: Wisconsin Education Association Council, Professional Development Division [On-line]. Available: [www.weac.org/resource/may96/classize.htm](http://www.weac.org/resource/may96/classize.htm)

Kiser-Kling, K. (1995). A comparative study of life in first grade classrooms of 1:14 and 1:23 teacher pupil ratios. Unpublished doctoral dissertation, University of North Carolina at Greensboro.

Klonsky, M. (2002). How smaller schools prevent school violence. Educational Leadership, 59(5), 65-69.

Krueger, A. (2000). An economist's view of class-size research. The CEIC Review, 9(2), 19-20.

Krueger, A. (2001). Smaller class sizes help Blacks more, study says. Jet, 99(15), 20.

Krueger, A. B. (1997, June). Experimental estimates of education production functions (Paper #6051). Cambridge, MA: National Bureau of Economic Research.

Krueger, A. B., & Whitmore, D. M. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. Economic Journal, 111(468), 1.

Kurecka, P., & Claus, R. N. (2000). A study of a reduced class size program, grades 1 & 2, 1999/2000. Evaluation report. Saginaw, MI: Department of Evaluation Services.

Laine, R. D., Greenwald, R., & Hedges, L. V. (1995). Money does matter: A research synthesis of a new universe of education production function studies. In L. O. Picus & J. L. Wattenbarger (Eds.), Where does the money go? Resource allocation in elementary and secondary schools (pp. 44-70). Thousand Oaks, CA: Corwin Press.

Lazear, E. (1999). Education production. Quarterly Journal of Economics. (in press)

Lewis, E. M., & Baker, L. S. (1997). Class size. The Future of Children: Financing Schools, 7(3), 112-121.

Madison Public Schools. (1976, September). Effects of class size on reading achievement in grades 1-3 in the Madison metropolitan school district (1974-1976). Madison: Author.

Maier, P., Molnar, A., Percy, S., Smith, P., & Zahorik, J. (1997). First year results of the student achievement guarantee in education program. Milwaukee: University of Wisconsin-Milwaukee.

Malloy, L., & Gilman, D. (1989). The cumulative effects on basic skills achievement of Indiana's PRIME TIME--A state sponsored program of reduced class size. Contemporary Education, 60, 169-172.

Maryland State Task Force on Reading. (2002). Reading achievement in Maryland. Annapolis, MD: State Board of Education.

Mazzoni, K. (1998). Are smaller class sizes a key to educational success? (yes). Spectrum: The Journal of State Government, 71(3), 26.

- McGiverin, J., Gilman, D., & Tillitski, C. (1989). A meta-analysis of the relation between class size and achievement. Elementary School Journal, 90, 47-56.
- Miles, K. H., & Darling-Hammond, L. (1998). Rethinking the allocation of teaching resources: Some lessons from high-performing schools. Educational Evaluation and Policy Analysis, 20(1), 9-29.
- Mitchell, D. E., Carson, C., & Badarak, G. (1989, May). How changing class size affects classrooms and students. Riverside: University of California, California Educational Research Cooperative.
- Mohs, M. (2002). [E-mail correspondence on class size reduction evaluation in Minnesota] [On-line]. Available: MatthewMohs@state.mn.us
- Molnar, A., Smith, P., & Zahorik, J. (2001). Evaluation of the student achievement guarantee in education (SAGE) program. Milwaukee: School of Education, University of Wisconsin.
- Molnar, A., Smith, P., Zahorik, J., Halbach, A., Ehrle, K., Hoffman, L. M., & Cross, B. (2001). Executive summary: 2000-2001 evaluation results of the student achievement guarantee in education (SAGE) program. Milwaukee: School of Education, University of Wisconsin.
- Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999a). Evaluating the SAGE program: A pilot program in targeting pupil-teacher reduction in Wisconsin. Educational Evaluation and Policy Analysis, 21(2), 165-177.
- Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999b). 1998-1999 evaluation results of the student achievement guarantee in education (SAGE) program [On-line]. Available: [www.uwm.edu/Dept/CERAI/sage.html](http://www.uwm.edu/Dept/CERAI/sage.html)
- Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (2000). Wisconsin student achievement guarantee in education (SAGE) class size reduction program: Achievement effects, teaching and classroom implementation. The CEIC Review, 9(2), 12-13.
- Molnar, A., Zahorik, J., Ehrle, K., & Halbach, A. (2000). Effective teaching in reduced-size classes. Milwaukee: Center for Educational Research, Analysis, and Innovation, University of Wisconsin.
- Moore, D., Bean, T., Birdyshaw, D., & Rycik, J. (1999). Adolescent literacy: A position statement. Newark, DE: International Reading Association.
- Mosteller, F. (1995). The Tennessee study of class size in the early school grades. The Future of Children: Critical Issues for Children and Youth, 5(2), 113-127.

Mueller, D. J., Chase, C. I., & Walden, J. O. (1988). Effects of reduced class sizes in primary classes. Educational Leadership, 45(7), 48-50.

Murnane, R. J., & Levy, F. (1996). Evidence from fifteen schools in Austin, Texas. In G. Burtless (Ed.), Does money matter? The effect of school resources on student achievement and adult success (pp. 93-96). Washington, DC: Brookings Institute Press.

National Conference of State Legislatures. (1998). Class size reduction [On-line]. Available: [www.ncsl.org/programs/educ/class.htm](http://www.ncsl.org/programs/educ/class.htm)

National Institute on Student Achievement, Curriculum, and Assessment. (1998). Reducing class size: What do we know? Washington, DC: Author. (ERIC Document Reproduction Service No. ED 420 108)

Nye, B. A., Achilles, C. M., Boyd-Zaharias, J., Fulton, B. D., & Wallenhorst, M. P. (1992). Tennessee's bold experiment: Using research to inform politics and practice. Tennessee Education, 22(3), 10-17.

Nye, B. A., Boyd-Zaharias, J., Fulton, B. D., & Wallenhorst, M. P. (1992). Smaller classes really are better. American School Board Journal, 179(5), 31-33.

Nye, B. A., Fulton, B. D., Boyd-Zaharias, J., & Cain, C. (1995). The lasting benefits study: Fifth-year report. Nashville: Center of Excellence for Research in Basic Skills, Tennessee State University.

Odden, A. (1990). Class size and student achievement: Research-based policy alternatives. Educational Evaluation and Policy Analysis, 12(2), 213-227.

Pinnell, G. S., deFord, D. E., & Lyons, C. A. (1988). Reading recovery: Early intervention for at-risk first graders. Arlington, VA: Educational Research Service.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2000). Teachers, schools and academic achievement (Working Paper No. 6691, Rev.). Cambridge, MA: National Bureau of Economic Research.

Robinson, G. E. (1990). Synthesis of research on the effects of class size. Educational Leadership, 47(7), 80-90.

Robinson, G. E., & Wittebols, J. H. (1986). Class size research: A related cluster analysis for decision making. Arlington, VA: Educational Research Service, Inc.

Scudder, D. F. (2000, December). Class-size reduction evaluation, 1999-2000. A report to the North Carolina Department of Public Instruction. Raleigh, NC: Department of Evaluation and Research.

Shapson, S. M., Wright, E. N., Eason, G., & Fitzgerald, J. (1980). An experimental study of the effects of class size. American Educational Research Journal, 17, 141-152.

Slavin, R. (1990). Class size and student achievement: Is smaller better? Contemporary Education, 62(1), 6-12.

Slavin, R. E. (1989). Achievement effects of substantial reductions in class size. In R. E. Slavin (Ed.), School and classroom organization (pp. 247-257). Hillside, NJ: Lawrence Erlbaum Associates.

Slavin, R. E., & Madden, N. A. (1995). Success for all: Creating schools and classrooms where all children can read. In J. Oakes & K. H. Quarts (Eds.), Creating new educational communities. NSSE yearbook. Part 1 (pp. 70-86). Chicago: University of Chicago Press.

Snow, M., & LaMarca, P. (2001). Effects of class configuration and teacher characteristics on achievement in reduced size classes. Carson City: Nevada Department of Education.

Stecher, B., Bohrnstedt, G., Kirst, M., McRobbie, J., & Williams, T. (2001). Class-size reduction in California. Phi Delta Kappan, 82(9), 670.

Stecher, B. M., & Bohrnstedt, G. W. (Eds.). (2000). Class size reduction in California: 1998-99 evaluation findings. Class-size reduction consortium, year 2 evaluation report. Palo Alto, CA: American Institute for Research.

Stecher, B. M., & Bohrnstedt, G. W. (Eds.). (2002). Class size reduction in California: Findings from 1999-00 and 2000-01. Sacramento: California Department of Education.

Strauss, S. E. (2000). Literacy learning in the middle grades: An investigation of academically effective grades schools. Unpublished doctoral dissertation, Florida State University, Tallahassee.

Strauss, S. E., & Irvin, J. L. (2000). Exemplary literacy learning programs. Middle School Journal, 32(1), 56-59.

Sturm, P. (1997). Nevada's class-size reduction program [On-line]. Carson City, NV: Senate Committee on Human Resources. Available: [www.leg.state.nv.us/cb/research/bkground/97-07.htm](http://www.leg.state.nv.us/cb/research/bkground/97-07.htm)

Texas Education Agency. (1998). Update on class size waivers: Bilingual education exceptions and waivers for English as a second language [On-line]. Available: [www.tea.state.tx.us/sboe/schedule/9801/dppef010.html](http://www.tea.state.tx.us/sboe/schedule/9801/dppef010.html)

Tinbergen, N. (1952). The curious behavior of the Strickleback. Scientific American, 187(6), 22-26.



Tomlinson, T. M. (1988, March). Class size and public policy: Politics and panaceas (OERI Rept. No. PIP 88-838). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Tomlinson, T. M. (1990). Class size and public policy: The plot thickens. Contemporary Education, 62(1), 17-23.

U.S. Department of Education. (1996, July). Pocket projections of educational statistics to 2006 (OERI/NCES Rept. No. 96-660). Washington, DC: Author.

U.S. Department of Education. (1997). Teachers' working conditions: From 4th condition of education. 1996 (OERI/NCES Rept. No. 97-371). Washington, DC: Author.

U.S. Department of Education. (1998). Reducing class size: What do we know? (OERI/NCES Rept. No. 98-3027). Washington, DC: Author.

U.S. Department of Education. National Center for Education Statistics. (1999). Snapshot of public schools in the United States. Results from the schools and staffing survey. Washington, DC: MPR Associates, Inc.

U.S. Department of Education. (2000, September). The class-size reduction program: Boosting student achievement in schools across the nation. A first-year report [On-line]. Available: [www.ed.gov/offices/OESE/ClassSize/](http://www.ed.gov/offices/OESE/ClassSize/)

Viadero, D. (1998). Small classes: Popular, but still unproven. Education Week [On-line]. Available: [www.edweek.org/ew/vol-17/23class.h17](http://www.edweek.org/ew/vol-17/23class.h17)

Voelkl, K. E. (1996). Measuring students' identification with school. Educational and Psychological Measurement, 56, 760-770.

Wang, M. (2000). How small classes help teachers do their best: Recommendations from a national invitational conference. The CEIC Review, 9(2), 1-3.

Wasley, P. A. (2002). Small classes, small schools: The time is now. Educational Leadership, 59(5), 6-10.

Weikert, D. P. (1989, June). Quality preschool programs: A long-term social investment (Occasional Paper No. 5). New York: The Ford Foundation.

Weikert, D. P. (1998). Results of a preschool intervention project. In S. Ryan (Ed.), US report on longitudinal evaluations of preschool programs (pp. 20-35). Washington, DC: U.S. Government Printing Office.

Wenglinsky, H. (1997). When money matters: How educational expenditures improve student performance and how they don't. Princeton, NJ: ETS Policy Information Center.

Word, E., Johnson, J., Bain, H., Fulton, B., Zaharias, J., Achilles, C. M., Lintz, N., Folger, J., & Breda, C. (1990). Student/teacher ratio (STAR): Tennessee's K-3 class size study. Final report and final report summary. Nashville: Tennessee State Department of Education.