2007

Implementing differentiated instruction in urban, Title I schools:: Effects of facilitated support groups and program fidelity on student achievement

Deborah W. Hellman
University of South Florida

Implementing Differentiated Instruction in Urban, Title I Schools:

Effects of Facilitated Support Groups and Program Fidelity on Student Achievement

by

Deborah W. Hellman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Exceptional Student Education
College of Education
University of South Florida

Major Professor: Albert J. Duchnowski, Ph.D.
David H. Allsopp, Ph.D.
Elizabeth M. Doone, Ph.D.
William F. Benjamin, Ph.D.

Date of Approval:
October 19, 2007

Keywords: implementation, inclusion, middle school,
professional development, and students with disabilities

Dedication

I would like to dedicate this dissertation to my husband, Greg Hellman. Our love has produced 23-years of a magnificent marriage and two wonderful children. I want to thank you for being my biggest supporter. You are always there for me and you made sure everything at home was taken care of so I could focus on furthering my education. Without you, I could not have completed this arduous task. Words cannot express how much your love and support has meant to me on this journey and in my life. I am who I am because of you.

Acknowledgements

I would like to thank my committee chair and members, Dr. Albert Duchnowski, Dr. David Allsopp, Dr. Elizabeth Doone, and Dr. William Benjamin. Your unwavering support and guidance throughout this process has been invaluable as you helped me conceptualize and implement this study. I would especially like to thank my chair, Dr. Albert Duchnowski, for your assistance, direction, encouragement, and hours of conference time. Your knowledge of evidence-based practice, fidelity studies, and general research knowledge has been indispensable to me.

Others who deserve special thanks are Charles Dixon, Dr. Joseph Brown, Jennifer Apgar, and Dianne Williams and all of the teachers at the school sites for making this study a success. Without your cooperation and your willingness to participate, I could not have completed my study.

I also want to thank my USF family for helping me whenever I asked. Thank you to all of the Exceptional Student Education instructors and office staff, especially Dr. James Paul and Jaye Berkowitz. Finally, yet importantly, this acknowledgement would not be complete if I did not thank my very special cohort family: Teri Crace, Sandra May, Sharon Ray, and Anne Townsend. The two years we spent together in classes as scholars would not have been the same without the mutual support we all shared.

Table of Contents

List of Tables

List of Figures

Implementing Differentiated Instruction in Urban, Title I Schools:

Effects of Facilitated Support Groups and Program Fidelity on Student Achievement

Deborah W. Hellman

ABSTRACT

This study presents the results of a mixed methodology study and pilot that investigated the effects of facilitated teacher support groups and differentiated instruction on student achievement at two urban, Title I middle schools. Both general education and students with special needs being served in a collaborative co-taught setting were included in the study. Implications for research to practice and effective inclusive strategies were addressed and the field-tested Differentiated Instruction: Fidelity Implementation Tool (DI: FIT) used to assess teacher fidelity is included.

During the first year, the principal investigator developed and field-tested the DI: FIT observation tool, field-tested a facilitated support group, and collected student achievement data to determine the feasibility of the implementation of differentiated instruction research design. During this second year, two matched urban, Title I middle schools were purposively selected to serve as research sites. At each of the two school sites, 13 to 15 teachers were selected to participate in the treatment group and 13 to 14 teachers in the control group. The teachers selected were balanced among the three grade levels within each school. A triangulation of data from monthly, 2-hour, facilitated

support group meeting minutes (group's perspective), teacher implementation logs (individual's perspective), and differentiated instruction observations (observer's perspective) were utilized to determine the impact of differentiated instruction on teacher implementation fidelity. Finally, the effects of teacher use of differentiated instruction with fidelity on the reading and mathematics achievement scores of approximately 906 students (461 in the treatment group and 445 in the control group) that were part of the combined sample population at the two school sites were assessed using ANOVA procedures. Cohen's (1977) *f* effect sizes are included.

Chapter One

Introduction

Statement of the Problem

Unquestionably, there are many problems facing education today. Two of these problems directly affect the teachers' ability to increase the academic achievement of students. First, the traditional one-day professional development opportunities provided to teachers are expensive and have not demonstrated a transfer of practices to the classroom (CEPRI, 2005; Greenwood & Abbott, 2001; Gregory, 2003; Guskey, 1986; Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; PCESE, 2002; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003; U.S. DOE, NCES, 2000). Second, many of the traditional modes of instruction currently used by teachers are inadequate to meet the varied needs of learners, especially struggling learners and diverse learners (Brandt, 1998; Chapman & King, 2005; NAS, 2002; PCESE, 2002; Tomlinson, 1995; Tomlinson, 1999; Tomlinson, 2003a; Tomlinson, 2003b; U.S. DOE, NCLB, 2002; U.S. DOE, OPSE, 2005).

The U.S. Department of Education's *President's Commission on Excellence in Special Education* (2002) admits that, "existing continuing education efforts are often inadequate for a number of reasons, including lack of substantive and research-based content, the lack of systematic follow-up necessary for sustainability and the "one-shot"

character of many workshop training programs" (p.55). With an estimated $730 million being spent on professional development, we are doing a disservice to America's future generations (CEPRI, 2005). Klingner, Ahwee, Pilonieta, and Menendez (2003) further state that although we know more about how to conduct professional development than we did twenty years ago, we have not learned how to increase the implementation of evidence-based practices on a broad scale. Leading researchers in the field attribute this gap between research and practice to limited implementation fidelity perpetuated by insufficient administrative support, inadequate follow-up support, little teacher collaboration at individual school sites, the pressures of high stakes testing, and a general lack of time (CEPRI, 2005; Guskey, 1986; Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003; U.S. DOE, NCES, 2000).

An additional factor that educators must address is the increased diversity in the general education environment. Recent data from the National Center for Educational Statistics (U.S. Department of Education, 2004) indicate that America's educators teach a very diverse group of students. Within the same classroom, teachers typically have students from a variety of cultural, language, and religious backgrounds who have different educational requirements and learning preferences. Currently, America's school population is comprised of 60% Caucasian and 40% students of color with 10% of the students receiving services for English language learners, 13% receiving services for students with disabilities, and 36% receiving free or reduced lunch (U.S. Department of Education, 2004). In urban districts, the diversity percentages are usually higher. The population of students in the district selected for this study are 43% Caucasian and 57%

students of color with 13% of the students receiving services for English language learners, 16% receiving services for students with disabilities, and 50% receiving free or reduced lunch (Florida Department of Education, 2007). "Given the diversity of America's students, along with their individual needs and learning styles, teachers must be able to tailor individualized instruction based on proven techniques and sound data" (U.S. Department of Education, 2005, p. 13). Contemporary classrooms are clearly more diverse than they were ten years ago and, with the current standards and accountability mandates, many teachers feel ill equipped to meet the needs of *all* of their students (Grant, 2000).

This is especially relative to educating students with special needs. The No Child Left Behind legislation (U.S. Department of Education, 2002) requires that *all* students, including students with special needs, meet or at least make adequate yearly progress toward uniform benchmarks. In addition, parents, school districts, and the state and federal governments are requiring schools to implement the least restrictive environment provision of the Individuals with Disabilities Education Act (1997). Consequently, more students with disabilities are able to access the general education curriculum and classrooms. With the challenge of meeting the needs of diverse learners, teachers require strategies and an educational philosophy that will help them to meet those requirements. Willis and Mann (2000) state:

> Every child is unique. Although we may rejoice in this fact, it poses a dilemma for educators. When students are diverse, teachers can either teach to the middle and hope for the best, or they can face the challenge of diversifying their instruction. (p. 1)

Some of the evidence-based and promising practices that educators are encouraged to implement in their classroom to meet the needs of the diverse learner are early intervention programs, cooperative learning, peer tutoring, direct instruction, mnemonic strategies, teaching reading comprehension, scientific inquiry, formative evaluation, and differentiated instructional strategies (Joyce & Showers, 2002; Forness, 2001; Klingner, Ahwee, Pilonieta, & Menendez, 2003; USDOE, OPSE, 2005). Of these, one particular instructional philosophy that is gaining support is differentiated instruction (DI). Differentiated instruction is a proactive, student-centered approach for teaching diverse learners in a supported, heterogeneous environment in which assessment drives the instruction. The differentiated instruction philosophy utilizes student assessment data to provide multiple learning opportunities for students that vary the content, process, and product in a blend of whole-class, group, and individual instruction according to the readiness, learning profile, and interests of the students (Tomlinson, 2000). The students must be presented with respectful tasks that are both engaging and challenging in flexible groups that are changed frequently so students are not "tracked."

Differentiated instruction has become a priority topic among educators because of its potential to transform classroom environments and to motivate students (ASCD, 2004). America's teachers desire an instructional philosophy that addresses the differences of auditory, visual, and kinesthetic learners, motivates students, and taps into students' personal interests (Gregory, 2003; Willis & Mann, 2000). They also require strategies to help create a classroom atmosphere that accepts and celebrates diversity (Fullan, 2001; Tomlinson, 2003b). The differentiated instruction philosophy supports the

current school philosophies of teaming, celebrating diversity, community building, and supporting the needs of all children (Lawrence-Brown, 2004; Tomlinson, 2003b).

However, as previously mentioned, in order to assess the impact of differentiated instruction on classroom learners, professional development facilitators and administrators must provide adequate on-going teacher support to ensure sustainability and to evaluate fidelity of implementation in the classroom (Blozowich, 2001; Boyd, 2001; Greenwood & Abbott, 2001; McAdamis, 2001; NAS, 2002; PCESE, 2002; Spencer & Logan, 2003; U.S. Department of Education, 2002). The literature on differentiated instruction and teacher support groups recommend further investigation. This study will contribute to the knowledge base for differentiated instruction, facilitated support groups, and implementation fidelity. This study will explore the effect of facilitated support groups on the degree of teacher implementation of differentiated instruction strategies in middle school, inclusive classrooms. This study will also evaluate the relationship among classrooms that utilize the differentiated instruction philosophy, the degree of teacher fidelity, and high-stakes academic state tests.

Theoretical/ Conceptual Framework

A constructivist's framework guided the investigations of this study. Constructivist researchers strive to gain consensus through their methodology by designing all constructions to be as real and accurate as possible and then by comparing and contrasting the constructions from multiple perspectives. Consequently, this study employed both quantitative and qualitative methods that naturally match the setting and research questions and provide multiple opportunities for the confirmation of findings.

First, it is widely accepted that the traditional professional development workshops that utilize a top-down approach in which educators come together for a day and receive information and materials with no follow-up or support have limited impact on teacher implementation (Greenwood & Abbott, 2001; Gregory, 2003; Guskey, 1986; Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003; U.S. DOE, CPRE, 1995; U.S. DOE, NCES, 2000). Over the last twenty years, professional development research has specifically focused on successful professional development programs in the hope of identifying components that will help close the research to practice gap. Because of their research efforts, there is now a list of factors that researchers generally agree facilitate the implementation of knowledge and skills learned during teacher in-service opportunities. Effective components include on-going teacher support, the use of assessment to inform instruction, time for collaboration, necessary materials, coaching, and feedback on implementation (Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003).

This study was designed to investigate the use of on-going, facilitated, teacher support groups, a practice that in previous research studies has claimed to increase the transfer of practices from the professional development opportunities to the classroom (Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003). Research findings have shown that the use of teacher support programs help schools and districts to ensure that the money, time, and other resources spent on teacher in-service are not

wasted (Blozowich, 2001; Boyd, 2001; Greenwood & Abbott, 2001; McAdamis, 2001; NAS, 2002; PCESE, 2002; Spencer & Logan, 2003; U.S. Department of Education, 2002). An effective teacher support program is "a service structure for teachers that can reduce their stress and burnout, increase retention, and improve teaching effectiveness through the use of best practices" (Westling, Cooper-Duffy, Prohn, Ray, & Herzog, 2005, p. 8).

Implementing a new instructional program or philosophy is difficult. According to Fullan (2001), the use of an on-going teacher support program will help offset the "implementation dip" associated with instructional change. Fullan (2001) found that over half of the teachers involved with change began with enthusiasm and confidence; but when they ran into challenges, their enthusiasm and confidence waned, and they typically gave up. The use of on-going teacher support groups can minimize the implementation dip effect (Osborne, 1993; Richardson, 1997; Sparks, 2001). Tomlinson (2005) points out that changing a teacher's teaching philosophy is both difficult and complex and can take from seven to fifteen years.

Within a teacher support group, members encourage each other, engage in collaborative problem solving, share ideas and successes, and explore available literature and resources. To maximize the benefit of the support group, participation should be voluntary and nonjudgmental (Westling, Cooper-Duffy, Prohn, Ray, & Herzog, 2005). On-going, facilitated teacher support/study groups have the potential to aid in the implementation process, increase teacher's fidelity to the model, and ultimately benefit students (Boyd, 2001; Davis, 2003; Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta,

& Menendez, 2003; Little, 1993; Martin, 2000; Pfaff, 1999; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003).

Quality professional development and the implementation of instructional strategies that address the diverse needs of today's learners are vital and inseparable factors of the educational equation (Brandt, 1998; Chapman & King, 2005; NAS, 2002; PCESE, 2002; Tomlinson, 1995; Tomlinson, 1999; Tomlinson, 2003a; Tomlinson, 2003b; U.S. DOE, NCLB, 2002; U.S. DOE, OPSE, 2005). Differentiated instruction is an instructional philosophy that combines numerous elements of evidence-based practices into a holistic instructional model. This philosophy has been gaining popularity since the 1990s (Hodge, 1997; Tomlinson, 1995; Tomlinson, 1999; Tomlinson, 2000) and is being adopted by many school districts across the country. The *No Child Left Behind Act* (NCLB) (U.S. Department of Education, 2002) and the *President's Commission on Excellence in Special Education* (PCESE) (2002) both call for the increased implementation of evidence-based practices. The U.S. Department of Education, Office of Post Secondary Education (2005) even makes specific references to the recommended use of differentiated instruction in *The Secretary's Fourth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom.* However, while there is much interest in the model, there is a paucity of empirical investigations evaluating the effect of differentiated instruction on student achievement and the teacher's ability to sustain fidelity to the model.

This lack of evidence is partly because the multi-facets of the differentiated instruction philosophy make it difficult to operationalize. Differentiated instruction is a proactive, student-centered approach for teaching diverse learners in a supported,

heterogeneous environment. In addition, assessment drives the instruction, which provides multiple opportunities to vary the content, process, and product in a blend of whole-class, group, and individual instruction according to the readiness, learning profile, and interests of the students. The following widely respected conceptual map (see Figure 1) developed by Tomlinson and Allan (2000) provides a concise visual.

When differentiating their instruction teachers use knowledge about each of their students to help them create lessons that provide two to four learning options. Teachers also employ a balance of strategies and approaches so each learner can be successful, e.g., stations, choice boards, curriculum compacting, cubing (see Appendix A).

A differentiated classroom has some obvious differences when compared to a traditional classroom (see Appendix B). To begin with, traditional classrooms are designed for organized, left-brain learners; whereas, differentiated classrooms support multiple learner profiles (Willis & Mann, 2000). Specifically, differentiated instruction is a proactive, student-centered approach for teaching diverse learners in a supported, heterogeneous environment (see Appendix C) (Heacox, 2002). Assessment drives the instruction, which provides multiple opportunities to vary the content, process, and product in a blend of whole-class, group, and individual instruction (Tomlinson & Allan, 2000; Tomlinson, 2003b). Differentiated instruction supports the constructivist's philosophy. It provides an opportunity for both the student and teacher to learn together and from each other (Lawrence-Brown, 2004; Tomlinson, 2001).

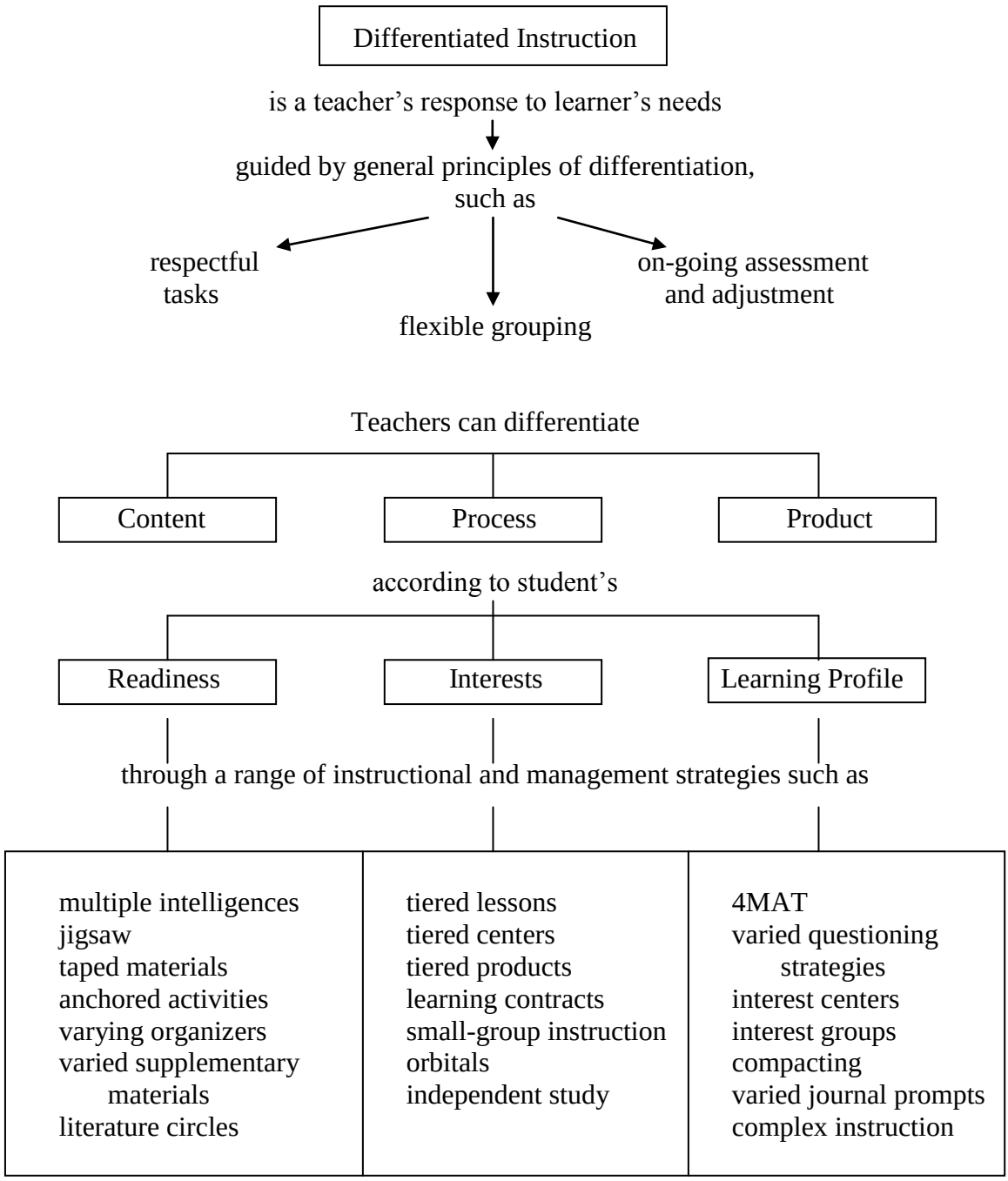


Figure 1. Concept map of differentiated instruction.

Note. From *Leadership for Differentiating Schools & Classrooms*, by C. A. Tomlinson and S. D. Allan, 2000, Alexandria, VA: Association for Supervision and Curriculum Development, p.3. Reprinted with permission from ASCD, February 2006.

In order for individual classrooms and schools to support differentiated instruction, school districts are encouraged to provide differentiated instruction in-service opportunities, administrative support, money for additional resources, teacher support groups, and site-based facilitators (Willis & Mann, 2000). "Teachers who are in collaborative situations with a coaching component that includes study teams and opportunities to problem-solve with supportive colleagues have an 80 to 90 percent chance of applying the innovation into their classroom repertoire" (Gregory, 2003, p. 9).

Finally, there is the issue of implementation fidelity. It is generally supported that for an instructional strategy or intervention to produce the desired effect, the classroom application must closely match the conditions and implementation procedures of the original research study (Cook & Campbell, 1975; Mokrue, Elias, & Bry, 2005; Seachrest, West, Phillips, Redner, & Yeaton, 1979). In the classroom, however, it is very difficult to determine the effectiveness of an instructional delivery program without also assessing the fidelity of implementation. When fidelity to the model is not monitored and maintained, researchers should always question whether the reported effects of the intervention are possibly the result of outside interventions and influences or whether the lack of desired results may possibly be due to an improper or inconsistent implementation of the intervention (Mokrue, Elias, & Bry, 2005).

In order to implement instructional practices adequately, administration and support team members must agree that program fidelity will be observed and documented. The use of an observation measurement tool further helps teachers and administrators measure and maintain instructional fidelity (Gregory, 2003; Tomlinson & Allan, 2000). Conducting on-going assessments of teacher fidelity to the model is critical

to student and program success (Webster-Stratton, 2003). Without this final component of a school program, the school and/or district will never know if the program is truly effective or if the outcomes are the result of chance. The concern for assessing program fidelity is not new and it is definitely not waning (Cook & Campbell, 1975; Greenwood & Abbott, 2001; Greenberg, Weissberg, O'Brien, Zins, Fredericks, Resnik, et al., 2003; Seachrest, West, Phillips, Redner, & Yeaton, 1979; U.S. Department of Education, 2002). Regardless of the intervention, teachers, facilitators, and administrators need to encourage a high degree of program fidelity and to collect and analyze implementation data on all interventions, treatments, and evidence-based practices.

Purpose

The goals of this study were (1) to investigate the effect of differentiated instruction on the mathematics and reading achievement of urban, middle school students; (2) to monitor teacher fidelity to the differentiated instruction model; (3) to assess the effect of facilitated support groups on teacher fidelity; and (4) to evaluate the relationship between the degree of teacher implementation of differentiated instruction and student achievement scores. This study contributes to the limited body of research addressing classroom implementation of the differentiated instruction model.

Research Questions

1. What were the effects of differentiated instruction with teacher support during a five-month period on the academic achievement outcomes of urban, Title I, middle school students?
2. What were the statistical differences among teacher groups who participated in facilitated support groups and those who did not with respect to their

implementation of differentiated instruction as measured by the Differentiated Instruction: Fidelity Implementation Tool (DI: FIT) (see Appendix D) observation tool?

3. What was the relationship between the teachers' differentiated instruction implementation scores as measured by the DI: FIT and the students' achievement change scores?
4. Using qualitative data and feedback provided by the teachers in the treatment groups, what were the teachers' perceptions of the facilitated support group model and their instructional growth?

Significance of the Study

In the context of accountability, schools across the United States need to implement research-based practices (Brandt, 1998; Greenwood & Abbott, 2001; Greenberg, Weissberg, O'Brien, Zins, Fredericks, Resnik, et al., 2003; Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003; U.S. Department of Education, NCLB, 2002). Differentiated instruction is an emerging practice that requires further investigation. Equally important, a gap still exists between research and practice in America's classrooms. For instructional programs, like differentiated instruction, to be accurately monitored, administrators must organize and endorse on-going, facilitated support groups so teachers can work with other professionals and problem-solve site-based solutions to the inevitable challenges of implementation. Finally, because instructional fidelity is difficult to maintain this study incorporates a component of continuous assessment where fidelity to the model is observed and measured.

In this period of emphasis on high stakes testing, administrators and educators are being held accountable for the progress of *all* students, and teachers are desperately searching for effective instructional strategies. Currently, many educators find themselves struggling to teach their increasingly diverse classes. Furthermore, differentiated instruction combines many evidence-based practices for teaching within one educational philosophy, which provides opportunities for increased social interaction, appropriate learning strategies, helpful feedback, and a positive learning environment (Brandt, 1998; Chapman & King, 2005; Hornsby & Diket, 1999; Tomlinson, 2005).

Professional standards warrant that all teachers provide students with a positive, interesting, challenging, collaborative, and supportive learning environment and that they not just teach to the middle. Education is no longer a "one size fits all" world. Teachers can no longer afford not to differentiate their instruction and administrators can no longer afford to take a passive role in supporting their teachers' endeavors. The differentiated instruction model's full potential can only be actualized through school personnel's collaborative efforts, and thus, *all* students will be afforded opportunities to learn.

Operational Definitions

Curriculum. The term curriculum refers to the prescribed content, skills, values, and attitudes that schools and teachers are accountable for teaching (U.S. Department of Education, 2002).

Differentiated instruction. Differentiated instruction is a proactive, student-centered approach for teaching diverse learners in a supported, heterogeneous environment. Assessment drives the instruction, which provides multiple opportunities to vary the content, process, and product in a blend of whole-class, group, and individual

instruction according to the readiness, learning profile, and interests of the students (Tomlinson, 2000).

Facilitated support group. A facilitated support group, also commonly referred to as a teacher study group, is represented by a small group of teachers, usually led by a facilitator, who come together on a regular basis to learn about a particular topic and to provide each other with support, information, and suggestions relating to implementation of a common instructional focus.

Flexible grouping. With flexible grouping, the students are continuously grouped and regrouped, to prevent struggling students from being singled out. Teachers can differentiate student groups by content, process, or product according to student readiness, interest, or learning profile. Students are expected to eventually work with all students in the classroom and teachers should not overuse any particular student grouping.

Implementation fidelity. Implementation fidelity is the degree to which a teacher maintains the integrity of a particular instructional program (reliability) and implements it in the classroom.

Inclusion. Educating students with individual educational plans in the general education classroom for all or part of the school day while providing the appropriate supports as needed.

Mathematics achievement. Mathematics achievement is an estimate of the student's ability to respond to standardized assessment items that measure number sense, number and operations, algebra, geometry, measurement, data analysis and probability,

problem solving, and reasoning as defined by the National Council of Teachers of Mathematics (2005).

Middle school. Middle school refers to the school configuration that includes sixth-grade through eighth-grade. In middle school, students are attached to heterogeneous teams that are taught by an interdisciplinary team of teachers.

Mild to moderate disabilities. Students who are identified as members of this category are students with disabilities that are cognitive, emotional, and/or physical and do not severely limit their ability to benefit from inclusion in the general education environment. The two categories that are the most commonly identified as "mild to moderate" as students with learning disabilities and students with emotional handicaps or emotional/ behavioral disabilities.

Professional development. Professional development is also frequently referred to as staff development or in-service training. It is used to describe the professional training experiences in which teachers participate in order to improve their instructional skills and/ or knowledge. Teachers usually receive stipend pay and in-service points, which help satisfy recertification requirements.

Reading achievement. Reading achievement is an estimate of the student's ability to respond to standardized assessment items that measure phonics, phonemic awareness, vocabulary, comprehension, and fluency as defined by the U.S. Department of Education.

Standards-based test. A test based on student learning standards. They are standardized achievement tests with criterion-referenced interpretations.

Tiered assignments. A lesson in which all students are working toward the same key concept even though they are purposely divided into groups that are adjusted according to student readiness.

Title I schools. According to the No Child Left Behind legislation (U.S. Department of Education, 2002), the public schools in each district that have student percentages that are above the districts' mean of children from low-income families are eligible to receive Title I, federal assistance funds. The school districts usually distribute the money based on the percentage of students eligible for free or reduced lunch at each site beginning with the schools with the highest percentages of eligibility downward until all funds have been expended. The purpose of the additional funds is to target the academic achievement of children from low-income families.

Title II schools. Some schools designated as Title II schools are eligible to receive supplemental funds in addition to their Title I funding (U.S. Department of Education, 2002). These schools have the highest percentages of students who qualify for free or reduced lunch and are in need of academic assistance. Title II schools in the district selected for this research study must have 90% or more of their students eligible for free or reduced lunch. Teachers who chose to work at these schools receive a supplement.

Delimitations

The results of this study may be generalized to diverse, urban middle school student populations in the southern United States and to the teachers who teach these populations using a team and inclusionary approach. The facilitated support group function results are appropriate for facilitators wishing to implement single-site

differentiated instruction support groups with teachers who are differentiated instruction novices.

Limitations

The interpretation of results of this study may be limited by the following possible threats to internal validity. First, there was the factor of history and the fact that some students can be exposed to instructional strategies and information outside of those used by their main teachers that may affect their achievement scores. However, in order to control for this threat, purposive comparison and treatment groups were selected from within each school site, which will help control for any school or neighborhood level factors. In this way, if a school starts a new educational initiative, it will tend to have an even effect on both groups and minimize any differences. Second, even though the data collection period for this study was five months, any maturation changes in the students and teachers was similar for both the comparison and treatment groups and has little effect on the results of this study. Third, because the Stanford Diagnostic Reading Test – Fourth Edition and FCAT Mathematics Predictor/ Benchmark Tests are administered every year to students in Florida, the Florida Department of Education has taken great care to make sure that the different forms of the test have equally high reliability and validity estimates without sensitizing the students to the questions or testing process. Fourth, to combat any reliability issues with the DI: FIT, the observation tool was field tested prior to this study and all observers were methodically trained and monitored with periodic checks for inter-rater reliability to ensure a high degree of reliability and validity. Furthermore, because there was no possible way to control for the loss of students to particular classes, the large number of students and teachers that were a part

of this study helped to minimize the effect of experimental mortality. In addition, the achievement scores from any students who are not present during the entire study were removed from the data set and not included in the final analyses. Both school sites for this research project were carefully selected and matched on multiple factors to ensure that the student populations were as closely matched as possible. For example, the schools were matched according to school size, SES, percentage of minorities, percentage of students with disabilities, region, overall school grade, and percentage of students not making adequate yearly progress. Finally, the bias of the experimenter or experimenter effect was controlled for by having a senior researcher who oversaw the entire study and by conducting frequent member and observer checks.

Chapter Two

Review of the Literature

Overview

Due to the multi-faceted qualities of differentiated instruction, it is difficult to research. Differentiated Instruction is not one entity, but a synthesis of many educational theories and practices. This holistic, student-centered approach combines many of education's best practices. Although there currently is limited research available on differentiated instruction as a whole, many of its common sense components and strategies are well grounded in decades of research on effective instructional practices. For example, the approaches of differentiated instruction include research supported components from Madeline Hunter's Essential Elements of Instruction, Spencer Kagan's Cooperative Learning, Howard Gardner's Multiple Intelligences, Robert Sternberg's Intelligences, David Sousa's Brain Compatible Learning, Lev Vygotsky's Zone of Proximal Development, and Robert Marzano's Dimensions of Learning. At this time, most of the current empirical research that is available has been done on gifted students. However, there is a growing body of individual research cases that demonstrate the effectiveness of differentiated instruction on students with disabilities in an inclusion setting (Baumgartner, Lipowski, & Rush, 2003; Good & Weaver, 2001; McAdamis, 2001; Tomlinson, 2000; Wertheim & Leyser, 2002).

Literature Search

In order to locate any pertinent research that might pertain to this study an exhaustive search was conducted using electronic searches, manual searches, and discussions with experts in the field. First, the university's electronic database was utilized to search Cambridge Scientific Abstracts (CSA) Illumina, SAGE, Educational Resources Information Center (ERIC), Wilson Select, all of the education full-text databases, and the dissertation database in an attempt to locate all articles pertaining to the following terms: "differentiated instruction," "differentiating instruction," "flexible grouping," "professional development," "teacher in-service," "study groups," "support groups," "professional study groups," "teacher study groups," "teacher fidelity," "implementation fidelity," "instructional fidelity," "fidelity," "inclusion," and combinations of these terms. This search yielded several hundred articles, however upon reading the abstracts it was found that most of the articles did not pertain to the focus of this study. All articles that matched any of the facets of this study's design were retrieved and reviewed for possible inclusion into the literature review. Research articles that pertained to middle school and/ or urban students were given priority. Recent issues of popular peer-reviewed research journals that focus on staff development and classroom instruction were also manually searched for the same topics that were previously listed. However, only a few more articles were located using this method. Finally, discussions were conducted with prominent researchers in the field who gave suggestions of books for background information. This researcher also conducted an audit of the references listed for any article selected for inclusion in this review.

Professional Development

Professional development for educators is an integral and costly part of the instructional process. The Council for Educational Policy, Research, and Improvement (CEPRI) (2005), an appointed council under the Office of Legislative Services for the Florida Department of Education, recently reported that during the 2002-2003 school year approximately $182 million was spent on staff development for teachers in Florida, not including the cost of substitutes. At the time, this amount equated to approximately $1,150 per teacher. Based on the percentages, the committee estimates that in 2005 Florida's districts spent over $730 million on professional development. The committee also points out that there is currently no systematic way to assess the benefits of in-service education and its impact on student achievement (CEPRI, 2005).

It is interesting to note that twenty years ago Guskey (1986) researched the major staff development projects that were prominent in the 1980's and found that they were all ineffective and unconnected to student achievement. At that time, he recommended that effective staff development programs need to "recognize that change is a gradual and difficult process for teachers," "ensure that teachers receive regular feedback on student learning progress," and "provide continued support and follow-up after the initial training (p. 9 – 10)." His recommendations have yet to become common practice. Joyce and Showers (2002), who have also been researching this problem for the last twenty years, similarly state that teachers and facilitators tend to underestimate the cognitive aspects of instructional implementation. The ability of staff development personnel and facilitators to effect the transfer of skills from the professional in-service to the classroom continues to be an area of research that is under investigation. They state that, unfortunately, most

professional development programs do not adequately address the teachers' need for on-going support, coaching, and the use of assessment to inform instruction (Joyce & Showers, 2002). The researchers further highlight the use of these effective implementation practices in projects such as Just Read, Read to Succeed, Success for All, The River City Experience, and the Schenley School Project.

In 2000, the U.S. Department of Education's National Center for Education Statistics conducted their own study on teacher preparation and professional development. They surveyed 5,253 teachers from all 50 states using a stratified sample procedure. In the report, they specifically criticized traditional one-day in-service opportunities as inadequate. Further, with regard to the types of in-service attended during the previous year, 80% of the teachers responded that they were "most likely" to have participated in workshops that focused on "state or district curriculum and performance standards" (p.2). They were "least likely" to have attended workshops that provided instructional strategies for working with "students with disabilities" (49%), "students from diverse cultural backgrounds" (41%), and "students with limited English proficiency" (26%) (p. 2). Sadly, only 15% of the teachers responded that their administration encouraged them to apply their new skills in the classroom. This report serves to confirm the lack of quality professional development activities that are available to teachers that encourage instructional strategies that support diversity and the lack of administrative support for professional development.

Lester (2003) in a study of 93 secondary teachers in Louisiana studied the components of effective professional development. Using a questionnaire, interviews, observations, and a reflective writing activity, she concluded that small, on-going

collaborative support groups that met on a regular basis and provided opportunities for reflection were perceived as effective by the teachers in helping them to grow professionally and implement "best practices."

Klingner, Ahwee, Pilonieta, and Menendez (2003) studied professional development components that could possibly help districts and facilitators "scale up" the implementation of evidence-based practices in inclusion classes. They specifically focused on the barriers and facilitators to implementation. For their study, they collected data from 29 teachers from six different elementary schools in Miami, Florida using interviews, teacher journals, and classroom observations. Each of the participants attended a two-week professional development program that targeted the use of four researched-based reading strategies: partner reading, collaborative strategic reading, making words, and phonological awareness. After the initial training, each teacher received on-going support for the remainder of the school year. Particular attention was made to help each teacher adapt the strategies to fit his or her particular teaching style. Researchers used a qualitative analysis of coding the chunks of data into five a priori categories for the interview and observation data and then they analyzed the teacher journals using SPSS 10.1.

In a comparison of high implementers (HI), moderate implementers (MI), and low implementers (LI), the most common barriers to implementation were "lack of instructional time," "students off-task," and "interruptions" (p. 420). Lack of instructional time was the most frequent complaint by the teachers, although it was noted that the teachers who implemented at a low level cited this reason more often than the teachers who were observed to be high implementers (HI). In addition, the low implementers (LI)

often felt overwhelmed and wanted more support and modeling of the strategy. All of the teachers reported that the strategy facilitators were very helpful in supporting the implementation of the strategies. Most of the teachers also responded that when the students liked a particular strategy or when the students performed well that knowledge helped to facilitate the teachers' implementation. The results of the study were mixed with 9 teachers considered to be high implementers (HI), 9 as moderate implementers (MI), and 11 as low implementers (LI). The researchers felt that with a little more time and continued support the moderate implementers (MI) could shift to high implementers (HI). More than a third of the teachers implemented the strategies at a low implementation level, even after a year of support and many teachers modified the strategies, leaving out important components. The factors that did help the teachers transfer the strategy knowledge from the in-service to the classroom with fidelity and helped the district to utilize their in-service budget more effectively were on-going support, modeling, observations, teacher "buy in," and reflection on instruction.

This study provided some very important aspects to consider when implementing evidence-based practices in general education classrooms that also provide inclusion services for students with disabilities. In the discussion, the authors use five examples to point out that even the best professional development studies have mixed results (Klingner, Ahwee, Pilonieta, & Menendez, 2003).

Differentiated Instruction

The support for differentiated instruction is rapidly growing. However, while there exists a great deal of support for each of the educational practices integrated into differentiated instruction, there are a relatively small number of studies that look at the

package as a whole. The following represent the studies that most closely match the intended focus of this study.

Beginning as early as 1990, Tomlinson, et al. (1994) began investigating the impact of differentiated instruction on high achieving, gifted learners, pre-service teachers, and teachers of the gifted. She primarily used qualitative research analysis to investigate the educational requirements of pre-service teachers, the impact of differentiated instruction on students, and the qualities of effective differentiated instruction in-services. First, she interviewed 70 pre-service teachers multiple times over the course of a year as they participated in a training program designed to prepare them to teach in a mixed-ability classroom. She found that pre-service teachers are not sufficiently prepared by their college educational programs to teach in a mixed-ability classroom and that specific training in differentiated instructional strategies is warranted.

The use of multiple reviewers and the multiple interviews over the course of one year added increased credibility to Tomlinson's et al. (1994) study. Although a convenience sample was used, the results are typical of pre-service teachers and programs in all regions of the United States. Consequently, administrators, mentors, and support group facilitators should consider the limited instructional strategy awareness of new teachers when planning staff development and instructional support opportunities.

Tomlinson (1995) followed the study in 1994 on the impact of differentiated instruction on gifted learners, pre-service teachers, and teachers of the gifted with another 18-month qualitative study focused on the ability of in-service facilitators to change the attitude and practices of middle school teachers with respect to differentiated instruction. For this study, she spent the entire year working with a typical mid-sized suburban

school. She also fostered credibility and dependability by collecting data from multiple sources. She presented staff in-services, observed teachers in classrooms, participated in their small learning communities, and interviewed teachers, administrators, and parents. She again concluded that teachers do not differentiate their instruction without specific instruction and support and that on-going administrative support is key to the change. Her research reinforced the growing belief that teachers require specific training and support in order to incorporate differentiated instruction into their instructional philosophy. Clearly, if administrators want to make sure that teachers implement differentiated instruction with fidelity, they should formulate multi-faceted programs that provide staff development, on-going support, classroom observation with feedback, and small learning communities.

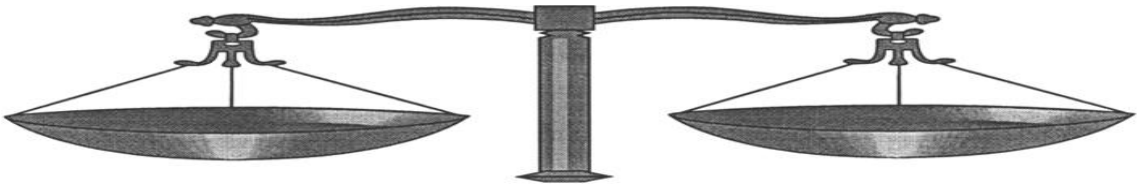
As a culmination to her research, Tomlinson wrote several books outlining the principles of the differentiated instructional philosophy. In a co-authored book, *Leadership for Differentiating Schools & Classrooms,* Tomlinson and Allan (2000) postulate that the following key principles govern effective differentiated instruction. The differentiated instruction concept map, shown earlier (Figure 1), is useful in gaining a holistic picture of the process. The authors state that, primarily, the classroom teacher must be flexible in his/her instruction, procedures, grouping, and assessment. It is only through on-going assessment that the teacher will be able to target the specific needs and interests of the students. This continual assessment process helps the teacher to plan future instruction. Third, the teacher adjusts the content, product, and process according to the student's readiness, interests, and learning profile. Fourth, the flexible grouping aspect of differentiated instruction will provide each student with a wide range of

learning opportunities and encourage the embracing of diversity. There are many grouping combinations possible using a differentiated approach. Lessons can be differentiated by content, process, or product according to student's readiness, interests, or learner profile. By combining two or more of these options based on students' needs, it is possible to create variety for both the students and teacher. Fifth, the teacher must ensure that all students are working on assignments that are both meaningful and engaging. In addition, the skill being taught must be challenging but not frustrating. Finally, the teacher and the students must both collaborate in the learning process. Students should be allowed the opportunity to make choices about their learning; it provides empowerment (Tomlinson & Allan, 2000; Tomlinson, 2003a).

In the same book, Tomlinson and Allan (2000, p. 134) also provide a helpful outline of key components that will help "balance the equation" so educators can successfully implement differentiated instruction in their classroom (see Figure 2). Some of these key components will be integrated into the current study. For example, they suggest that the administration provide "focused staff development," "time and support for collaboration," "generalist/ specialist partnerships for classroom application," "integration with professional growth and accountability," and "assessing student growth."

Using a qualitative design similar to Tomlinson's studies, Fleming and Baker (2002) investigated the interactive role between lesson planning, student teaching preparation, and student teachers' experiences with differentiated instruction. The participants in this study were five pre-service teachers who were placed at three rural Ohio middle schools, and the goal was to collect evidence of their transfer of the

differentiated instruction knowledge gained during their college pre-service methods class into their field teaching experiences.



The What

- High-level, idea-based instruction using key skills to understand and apply the ideas employing key principles of differentiation:
  - Flexible grouping
  - Respectful activities
  - On-going assessment and adjustment
- Modifying content, process, and product based on student readiness, interest, and learning profile using a range of student-centered, meaning-making instructional strategies
- Coaching for individual growth with the goal of moving each student as far and fast as possible
- Assessing student growth at least in significant measure according to personal growth

The How

- Clarity of purpose and vision
- Systemic efforts
- Generalist/ specialist partnerships for classroom application
- Time and support for collaboration
- Structured lesson (curriculum) planning and instructional evaluation
- Focused staff development with plans for transfer
- Incentives for classroom application
- Aligned and focused policies and initiatives
- Coherent leadership
- Integration with professional growth and accountability
- Formative and summative evaluation of efforts and use of findings
- Involvement of parents in understanding and contributing to assessment of change
- Persistence over time

Figure 2. *Balancing the equation to make differentiation work.*

Note. From *Leadership for Differentiating Schools & Classrooms,* by C. A. Tomlinson and S. D. Allan, 2000, Alexandria, VA: Association for Supervision and Curriculum Development, p. 134. Reprinted with permission from ASCD, February 2006.

The participants in the Fleming and Baker (2002) study were purposefully selected based on their completion of the college methods class. For data collection purposes, each participant was asked to submit six lesson plans, three from their college classes and three from their classroom field experience. After these documents were reviewed for evidence of differentiated instruction according to content, process, or product, each participant was observed twice during the last two weeks of their field experience, by two different observers. Once they had finished their student teaching, each participant completed a short survey and was interviewed using four open-ended questions. The participants were not told that the researchers were specifically looking for evidence of differentiated instruction so as not to bias the results. In the final report, pseudonyms were used to reference each teacher to insure confidentiality.

In order to insure dependability and confirmability of the results, both researchers separately reviewed all data prior to collaboration. Each of the researchers also represented different areas of expertise; one taught general education while the other taught special education. The study was conducted over a one-year period to increase credibility of their findings, and both method and data triangulation were utilized which further enhanced the dependability and confirmability of the report findings. Finally, to address transferability, the researchers provided very clear, thick rich descriptions of their methods and results so the results could be replicated and/ or applied to other situations; however, the small sample size and geographic location of the study will greatly limit the generalizability of the results.

The results showed that there still was clearly a gap in the expectations of the university supervisors and the classroom supervising teachers. Several of the pre-service

teachers experienced resistance by the classroom supervising teachers who did not understand the differentiated instruction philosophy and did not require the pre-service teachers to write lesson plans. Most of the pre-service teachers were still confusing differentiated instruction with simple classroom accommodations. It was also noted that the pre-service teachers limited their differentiation to product options, which are the easiest to implement, with an absence of differentiation by content or process. All of the participants commented in the interviews that differentiating in the classroom was harder and more time consuming than they had expected. Again, the importance of providing teachers with support and time for collaborating and developing differentiated lessons is key to implementation. It is also very clear from this study that new teachers, even those who have participated in a formal educational preparation program are ill equipped to differentiate their instruction in a mixed-ability classroom.

Hobson's descriptive study in 2004 provided limited expansion of previously known information for researchers hoping to develop their differentiated instruction school or district program. This mixed-method study used a questionnaire, a focus group, and walk through classroom observations to collect data at a single semi-rural middle school in the Shenandoah Valley in Virginia. It was noted that this district is so small that it only had one middle school, so random selection was not possible.

Even though the school selected for this study was semi-rural, the diversity of the school's population was more heterogeneous than that of previous studies. The students in the sample were 61% Caucasian and 39% minority students. The non-English speaking students represented 34% of the student body and over 40% of the students qualified for free or reduced lunch. A high percentage of the students were gifted (12%), although, no

information was available on the percentage of students who qualified for special education. Fifty-five teachers were part of the targeted convenience pool of participants, however, only 30 teachers returned the questionnaire and became part of the study.

Triangulation of evidence was used in order to increase the internal validity of the study. For the quantitative portion of the study, Tomlinson's *Teacher/ Peer Differentiation Reflection Instrument* (2000) was used as the model for the researcher's questionnaire and observation checklist. Reliability for the questionnaire, which was reviewed by a panel of experts with minimal changes, was established using the split-half method ($\alpha = .93$). Several limitations of these data were found in this study. For the questionnaire, participants were enticed with a cash drawing. For the observation data, 15 teachers were selected for the unannounced walk-through observation, which only lasted 15 minutes. Another limitation to the observation data was that the researcher did not specify how the participants were selected. For the qualitative portion of the study, seven teachers from the staff voluntarily participated in the focus group. Furthermore, the participants in the focus group were enticed with the promise of a cash drawing, a factor that may affect the internal validity of the results. All instruments were field tested first using a modified pilot.

The results of the self-reported questionnaire and observations indicated that most teachers who differentiated their lessons utilized content differentiation. However, the seven teachers who participated in the focus group responded that they used product differentiation more often than content differentiation. Most of the teachers in the focus group agreed that they took the students' readiness and interests into consideration when planning instruction. The majority of these teachers also felt that they were comfortable

using materials other than the textbook for lessons. From a pragmatic standpoint, this study did not yield any particular patterns in teacher use of differentiated instruction that were not expected. Instead, it was found that there is a very wide discrepancy between the types of strategies used and the degree of implementation. As expected, the author recommended continued staff development and support within the school.

In 2002, Vanfleet focused specifically on the use of differentiated instruction as a means to facilitate inclusion for students with mild to moderate disabilities. Using a mixed-method design, she explored the effectiveness of professional development differentiated instruction training with 43 secondary school teachers in a suburban area of Alabama. The quantitative portion of the design consisted of an analysis of data collected using the Data Survey of Secondary School Teachers. The pretest-posttest data from this survey were compared using a t-test. The results showed that the teachers' perceptions as to whether or not they were adequately meeting the needs of students with disabilities in the general education classroom were effected by the differentiated instruction training they received to a statistically significant degree ($p \leq .05$). The qualitative portion of her study revealed that the participants with the most recent formal education had the most collaborative partnerships, improved preparation and attitude toward diversity and students with disabilities, and a greater experience with differentiated instructional strategies.

In a similar study in 1997, Hodge's investigation specifically examined the impact of differentiated instruction staff development on student achievement, perceptions of parents, and teacher attitudes. Using a t-test procedure, she analyzed the reading and math achievement data of students in grades two through six at a suburban, Alabama public

elementary school and found that the students who received differentiated instruction made statistically significant gains in mathematics but not in reading [$t_{(92)} = 2.24$, t-critical = 1.66]. Student academic growth was measured using the Stanford Diagnostic Test: Reading and Mathematics Batteries developed in 1983. Surveys were also administered to 44 teachers and 160 parents with 79% of the parents responding. An analysis of the survey data found no statistical significant differences between the perceptions of the treatment and control groups of parents and the treatment and control groups of teachers with reference to the teachers' ability to meet the requirements of diverse learners in the classroom. However, the study presents limited validity or generalizability for the diverse populations in the urban centers of America. The cluster sample of 160 students (94 treatment, 96 control) used for this quasi-experiment was 98% Caucasian, upper socioeconomic status (median family income of $80,366), with only 10.7% of the students in need of academic remediation. In fact, the mean academic achievement scores of the students at the selected elementary school were well above the national average, based on state assessment results, prior to the study. Furthermore, of the teachers used in the sample, 60% held advanced degrees with a mean of 12.5 years of teaching experience, again above the national averages.

There is much to be learned from the staff development and support that the teachers in Hodge's (1997) study received. The teachers originally volunteered to participate in a 16-hour staff development with Carol Ann Tomlinson in 1996, followed up by seven months of on-going support and quarterly staff development. Throughout the study, the teachers videotaped themselves, watched the tapes, and reflected on ways to

improve their instruction. They also observed teachers within the school that were considered successful and confident with their differentiation of instruction.

Affholder (2003) expanded the research vein of differentiated instruction by investigating its use with all learners in inclusive classrooms. She used a case study design with branching interviews and questionnaires to examine the implementation of differentiated instruction and the factors required to support this approach based on a district-supported initiative. The Blue Valley School District in Kansas began the initiative six years prior and made sure that the appropriate supportive components were in place by providing differentiated instruction staff development, time and resources, opportunities for collaboration, and shared decision-making.

The focus of this study (Affholder, 2003) was the perceptions of 26 elementary school teachers, 12 administrators, and a school board member. In addition to the interview, the selected elementary school teachers also responded to the Stages of Concern Questionnaire developed by Hall and Hord in 2001. All of the teachers in the study participated in a differentiated instruction in-service two years earlier and the researcher collected evidence as to the degree of implementation that each teacher maintained. The branching interviews served as a vehicle to locate those teachers who had the highest levels of implementation. The consensus of the data revealed that the teachers all expressed the desire for on-going support and staff development in addition to time and resources. Specifically, according to Affholder (2003), "Every teacher interviewed in this study mentioned time as a critical factor for the implementation of differentiated instruction, time for lesson planning and preparation, time for collaboration, and a student contact time sufficient for assessment and instruction of

students" (p. 11). It was also found that the teachers with the greatest degree of fidelity had a pronounced sense of responsibility for student growth, familiarity with the curriculum, and a willingness to try new instructional approaches. This study is limited, as was the previous study, by the homogeneity of the sample. The small suburban Blue Valley District is primarily Caucasian (90%) and only 4% of the students qualify for special education, compared with the national average of 13% (U.S. Department of Education, 2004). In the future, if this study is replicated, a more diverse population would help to generalize the results.

McAdamis (2001), the coordinator of staff development for Rockwood School District in Missouri, is one of the few educators who seems to have pulled many of the effective components of instructional implementation into a complete design. In her article in the *Journal of Staff Development,* using a narrative style, she recounts the district's 5-year action research process.

The district began by adopting a policy of supporting differentiated instruction in order to address the needs of all learners. First, a cadre of volunteer teachers was trained in differentiated instruction and peer coaching strategies. The group met five times during the first year, during which they learned new strategies, shared successes and challenges, developed new lessons, and practiced reflection. They also observed each other and practiced giving feedback as a peer coach. These teachers then became the "critical mass" that provided support for other teachers at their individual school site. Next, the district began offering differentiated instruction workshops on a continual basis to all teachers and administrators. In addition, the district added on-going instructional support in the form of release time for teachers to work on lessons, teacher support/ study groups, peer

coaching, and classroom level, action-research support. In the end, each school was required to add a differentiated instruction goal to its school improvement plan.

McAdamis (2001) admits that the change did not happen overnight, the entire process was five years in the making; but in the end, the payoff for everyone involved was well worth it. Because of the district and teachers' efforts, Rockwood students increased their academic performance on the state standardized tests. Overall, the percentage of under-performing students, bottom quartile, decreased by 8% in reading and language arts, 5% in math, and 7% in science. They also increased the percentage of students scoring in the top quartile.

Rockwood is a suburban district, part of the greater St. Louis area. The district employs approximately 3,000 teachers and has an enrollment of 22,000 students who are represented by the following percentages: 83% Caucasian, 14% African-American, and 3% other. The students who qualify for free or reduced lunch represent 15% of the population, and students with special needs represent 8%.

Although this study was not a rigorous quantitative study and it did not represent the population targeted for this study, it does represent the power of differentiated instruction coupled with on-going teacher support, teacher support/ study groups, and observation with feedback.

As evidenced by the previous research reviews, there is limited information on the academic achievement effect of differentiated instruction in American classrooms. Sadly, there are little rigorous data available on its implications with urban middle school students. There is, however, ample evidence that observation and on-going support is key to an effective differentiated instruction school program.

Facilitated Support Groups/ Teacher Study Groups

Having a good facilitated support group is a key ingredient to initiating an effective instructional program (Westling, Cooper-Duffy, Prohn, Ray, & Herzog, 2005). There has been an exceptionally strong push throughout the United States due to both the U.S. Department of Education's NCLB Act of 2001 and the President's Commission on Excellence in Special Education's (PCESE) report, *A New Era: Revitalizing Special Education for Children and Their Families (2002),* for the field of education to include more research-based practices in the classroom. However, one of the primary reasons that many of these practices are absent from the classroom instructional repertoire of teachers is that the majority of teacher in-service programs have little to no follow-up support. Consequently, when challenges arise, if teachers feel unsupported and unsure of their abilities, they will revert to what is comfortable (Osborne, 1993; Richardson, 1997; Sparks, 2001; Spencer & Logan, 2003).

Many of the previously discussed differentiated instruction studies emphasized the characteristics and necessity of having a well-run professional support group for teachers. An additional qualitative case study conducted by Hale (1999) investigated, on a smaller scale, the use of a facilitated support group as a vehicle for professional development. For this study, seven enhancement specialists, from different urban schools, volunteered to be the participants. In order to support and monitor the group's process and progress, a university facilitator met with this formal support group throughout the entire study. The initial task for the group members was to evaluate, plan, and implement research-based strategies as well as finding a productive way to work together and support each other's efforts in their individual schools.

The author, using thick, rich, narrative, documented the entire process as these specialists developed their collaborative process and the effect that the collaborative process had on their work and their relationships. As a result of the collaborative experience, the author was able to identify affective and supportive factors that contributed to the group's facilitative process. The author also proposed that other aspects such as external events, professional background, disposition, and expectations were key factors that effected the participants' perceived gain from the professional study group experience.

Even though this was a small qualitative design with voluntary participants, the study did provide guidance for future researchers regarding the process and possible problems that can be avoided when facilitating professional support groups. For example, the group benefited from an experienced facilitator, group norms need to be established at the first meeting (be on time, respect the input from all members), notes should be taken during each meeting, member checks as to the accuracy of the notes need to be conducted immediately after each session, all members need to agree to make the meetings a priority and to not schedule other activities during the agreed meeting times, make sure the meeting place is comfortable, the group size should be kept small, and the group should meet at least once a month. Although this study may have limited transferability, it provides a very clear picture of the affective factors that must be considered when planning a teacher support group such as the one for this research project.

On a more in-depth and slightly larger scale, Pfaff (1999) further investigated the effect of professional study groups on teacher efficacy. For her study, she used a mixed-method design, primarily qualitative, to answer multiple questions regarding the effect of

a year-long professional study group on the perceptions of seven staff members from a small, rural elementary school located in Carroll County, a north central Maryland district.

Unfortunately, the school was the only site that met the researcher's requirement and was therefore purposefully selected. All of the elementary schools from the district were invited to participate, however, only three principals responded. Of the three, two already had study groups in place, which might bias the teacher participants. The school selected was the smallest elementary school in Carroll County with only 417 students and 25 teachers. Again, the entire staff was invited to participate, however, only eight members volunteered and one teacher dropped out early in the process due to scheduling conflicts. Of the seven remaining members, two were identified as resource specialists and five were classroom teachers. The mean number of years of experience represented by the group was 11.8 years and all but one of the staff members had a Master's Degree. The results of this study therefore have limited transferability due to the small, purposeful, non-typical sample of participants.

The student body of the school was predominantly middle-class and Caucasian. They had no students who were eligible for the Limited English Proficiency program, only 8% who qualified for the free and reduced lunch program, and 15.3% who received special education services.

The professional study group met for 90 minutes per month and each teacher was paid a small in-service stipend for putting extra hours beyond the standard school day. At the end of the year, questionnaires and interviews were utilized to collect qualitative data documenting the changes in the teachers' perceptions of their personal and general

teaching efficacy because of their participation in the professional study group. The results supported the belief that the group participation facilitated a sense of security and confidence in the participants' abilities. The teachers reported that the study group also helped to increase their metacognitive awareness of their own abilities. In addition, using a pretest-posttest, factorial ANOVA analysis, the responses of the teachers in the professional study group on The Teacher Efficacy Scale (TES) were compared to the responses of a control group of teachers selected from the same site. The author provided the reliability and internal validity scores for the TES instrument as reported by the creator of the instrument. The results of this analysis did not show a significant difference between the two groups. This lack of difference was rationalized to be directly related to the fact that the teachers in the treatment group scored high on the instrument during the pretest and therefore had little room for growth. Regression to the mean was a factor to be considered. It was further noted that the teachers in the study group maintained their high efficacy ratings until the end of the year, whereas the mean efficacy rating of the teachers in the control group declined over the course of the year.

The results of this study provide important aspects to consider when creating teacher support/ study groups. The impact on student achievement was missing from this study and is recommended to be added in future studies. In order for the results to be more broadly transferable, the sample of teachers, schools, and student body should also be more representative of the typical classrooms of America.

In another qualitative, longitudinal, case study, Boyd (2001) investigated the implementation of inclusionary practices at a moderately sized, urban middle school in a Central Florida county over the course of four years. The primary focus of her research

was to explore the effect of the climate, best practices, administrative support, and staff attitudes on the inclusion of students with disabilities at the school, which educated an average of 899 students, including 142 varying exceptionality students (15.8%), 178 students with limited English proficiency (LEP) (19.8%), and 458 students on free and reduced lunch (50.9%). The mean enrollment by ethnicity was 37.6% Caucasian (n = 338.5), 14.7% African-American (n = 131.8), 39.5% Hispanic (n = 355), and 8.2% other (n = 73.7). For this study, the researcher increased dependability by triangulating data from multiple sources. She conducted surveys, observations, focus groups, a document review, and interviews from a pool of 38 staff members including 31 teachers, an administrator, 3 deans, 2 guidance counselors, an ESE resource specialist. She also increased her data credibility by collecting these data over a four-year period. Anonymity of all participants was assured for this voluntary study.

Of the 38 possible participants, 34 responded to the survey, providing good credibility of the information. From the pool of 31 teacher participants, the seventh and eighth-grade content area teachers were considered the primary focus of the study. Therefore, only 18 seventh-and eighth-grade teachers were selected to participate in the interviews, 11 seventh-and eighth-grade teachers were selected for observations, and 10 seventh-and eighth-grade teachers participated in the two focus groups.

The author claimed that the 63-question, Likert scale survey was adapted from Van De Mark's research in 1997, although no validity or reliability information was provided. The researcher followed standard development protocol while developing the classroom observation tool. Each teacher was observed nine times for approximately 45 minutes per observation. The open-ended interview questions were purposefully

developed in order to address four main categories: school climate, administrative support and preparation, attitudes of the participants, and best practices. The interview participants were provided with a copy of the questions during the interview to ease their fears and to guide the process. Member checks were utilized to verify the emergent themes. As a final confirmation piece, the focus groups helped to provide clarity of issues raised during the interviews and informal conversations on campus with the researcher.

From her research, she concluded "that the success of the inclusion program was largely dependent upon the stakeholders having ownership in the program" (p. iii). Respondents indicated that over the years the program had deteriorated and they had been left out of the decision-making process. Most of the staff agreed on the survey that the differentiated instruction classroom practices, like cooperative learning, multiple intelligences, activity based learning, portfolio assessment, and peer tutoring were all important; however, when observed, most teachers utilized a purely traditional teaching style. She also found that continuous staff development including peer coaching, mentoring, small learning communities, teacher observations, and a focus on best practices were components of an effective inclusive support program and should be added in the future research studies.

More recently, Davis (2003) conducted a larger, mixed-method research study in which she explored the relationship between the use of study groups and the implementation of knowledge and strategies gained during professional development workshops. For this study, the sample consisted of 57 elementary school teacher participants from 14 study groups within a suburban, Pennsylvanian school district. The transfer of skills from the study group to the classroom was measured using self-reported

data using a pretest-posttest design. The use of self-reported data has limitations and decreases the reliability of the conclusions. The participants also completed a Learning Style Inventory designed by Kolb in 1999 and a Study Group Participant Questionnaire designed by the author. In addition, both participant and facilitator interviews were conducted in an effort to triangulate data relative to "group management and structure, group resources and technical support, group dynamics and interpersonal relationships, and group outcomes addressing the functioning of study groups." Statistical analyses were conducted using regression analysis, ANCOVA, and MANOVA procedures.

The statistically significant ($p \leq .05$) results of the combined analysis revealed that all 14-study groups perceived that they were able to transfer the skills learned during professional development workshops to the classroom. However, the results of the teacher learning style data provided inconsistent results that were not statistically significant and did not support the teachers' perceived increase of new knowledge and skills. It is therefore not sufficient to rely solely on the use of interviews and self-reported data. The transfer of new knowledge and skills into the classroom and a teacher's instructional repertoire needs to be quantified with observational and student data.

Although Murphy and Lick (2001) conducted their research on whole faculty study groups, their pivotal work provides insight and guidelines that would also benefit researchers planning to conduct research on stand-alone collegial teacher study groups and was therefore selected for inclusion in this literature review. In their book, they summarized their sizeable research that was collected over a five-year period and utilized data from over 2,000 study groups at over 200 schools. The authors begin by clearly stating that for study groups to be successful, the individual group members must first

agree on a guiding question, guiding principles, and procedural guidelines. Murphy and Lick (2001) emphasize that these criteria cannot be sacrificed or the group will flounder. In order for the study groups to be effective, they provide the following additional guidelines:

- Keep the size of the group to no more than six
- Don't worry about the composition of the study group
- Establish and keep a regular schedule
- Establish group norms at the first meeting of the study group
- Agree on an action plan for the study group
- Complete a journal entry after each study group meeting
- Encourage members to keep individual journals for their personal reflections
- Establish a pattern of study group leadership
- Give all study group members equal status
- Have a curriculum and instructional focus
- Plan ahead for transitions
- Make a comprehensive list of learning resources, both material and human
- Include training in the study group's agenda
- Evaluate the effectiveness of the study group
- Establish a variety of communication networks and systems (p. 51-59)

The authors further emphasize the key role of the facilitator or leader. They state that the facilitator has an essential role in the functioning of the study group. In order for the study group to be effective, the facilitator must take on the additional roles of organizer, recorder, and liaison.

Based on the research by Murphy and Lick (2001), Martin (2000), an elementary school principal, conducted a qualitative descriptive multi-case study investigating the effects of study groups on teachers' implementation of skills learned during professional development workshops, the interaction of the study group members, and the effects of the facilitators on the study groups. The sample consisted of 24 teachers from a rural-suburban school district in Pennsylvania. Martin (2000) collected data on two treatment study groups of seven and five members each who were matched to another 12 teachers in a control group. The 12 teachers who voluntarily participated in the study groups were matched to teachers in the control group who had attended the same in-service opportunity but chose not to participate in the study group. For her data collection, she utilized interviews, questionnaires, and an innovation checklist. She had originally intended to conduct a secondary statistical analysis on some of the data, but she was unable to do this as her sample size was too small.

Martin (2000) reported that the majority of the teachers in the two study groups believed that their classroom instruction had improved as a direct result of their participation in the study group. The participants also confirmed that the facilitator's role was critical to their groups' success. They stated that the facilitator helped to maintain the organization, on-going communication, and interpersonal functioning of the group. In a summary analysis of the groups' functioning, Martin (2000) stated that the key components necessary for teacher integration of professional development skills were: "regular meetings with adequate structure," "documentation and evaluation of group work," "facilitative leadership," and "support for the groups' efforts."

Martin's study has limited transferability due to its small size and the characteristics of the participants, although her descriptions of the facilitative process of the study groups would be very beneficial to researchers choosing to implement a support/ study group. Because she used self-reported data, the validity of the findings may be limited. She admits that the group would have been more productive if they had kept a formal meeting log to document their progress and accomplishments. Furthermore, observations by a trained observer would have added another source of data verification.

Recently, another pair of researchers (Spencer & Logan, 2003) also investigated the use of on-going teacher study groups, but added components that were missing from some of the previously discussed studies. Spencer and Logan (2003) used a time series experiment with a treatment and a comparison group to study the effects of a school based staff development model that utilized the support of a Research Lead Teacher (RLT) to help general education teachers develop and maintain instructional fidelity to the district mandated "Benchmark Strategy Instruction Process."

For this study, Spencer and Logan (2003) selected a large elementary school (1100 students) in a large suburban school district. The location of the school was not provided. The student body was representative of the U.S. Department of Education's NCES (2004) national averages for elementary schools with 59% Caucasian, 18% African American, 12% Asian, 8% Hispanic, and 3% other. The percentage of students enrolled in special education was unavailable, however, it was noted that 25% of the students qualified for free or reduced lunch. The mean number of years of teaching experience for the intervention group was 9.7 years, while the comparison group averaged 11.3 years. All teachers had between two and four students with mild to

moderate special needs included in the general education environment in addition to several more "at risk" students.

First, all of the 42 general education K-5 teachers in the selected elementary school attended the traditional half-day in-service provided by the Research Lead Teacher (RLT). Then all of the teachers were invited to participate in the RTL Model that included "an on-going teacher study group, coaching, observations, and data based feedback" (Spencer & Logan, 2003, p. 51). From the teachers who volunteered, nine became part of the treatment group and nine became the comparison group.

The data collection began during the 9-weeks before the beginning of the school year, during which the nine teachers in the treatment group attended nine 60-minute, weekly teacher study group sessions as part of a voluntary summer training opportunity. The treatment group of teachers also agreed to be observed on a weekly basis with follow-up feedback provided by the RLT. The RLT also provided in-class modeling and coaching of the Benchmark Strategy Instruction Process for each of the teachers as needed. For their participation, the teachers received ten staff development credit hours. The nine teachers in the comparison group, without on-going support, were told from the beginning the purpose of the study. They also agreed to be observed during the baseline, treatment, and maintenance weeks. The comparison group of teachers knew that their observation data would be compared to teachers who attended the same in-service with the additional coaching support. Therefore, no deception was involved. The researchers felt that because they were aware of the study's purpose they would make sure that they presented lessons to the best of their ability, possible John Henry Effect, thereby illustrating the impact of the RLT Model.

After the initial 9-week training period, all 18 teachers were observed four times during the 3-week baseline period to determine the degree of implementation by each teacher prior to the intervention from specialists and observational feedback. All observations lasted approximately 30 to 40 minutes. This phase was followed by a 9-week treatment period, during which all teachers were observed nine times. The comparison group received no feedback while the treatment group received specific feedback, modeling, and support. During the next 3-week rest period, no observations were conducted. This period was followed by the final 2-week maintenance check in which all teachers were again observed twice. Two trained observers were used to collect the classroom implementation data. The observers were not aware which teachers were receiving the intervention in order to avoid observer bias.

Using the RLT Model, all of the teachers in the treatment group mastered the implementation of the 15 steps of the Benchmark Strategy Instruction Process. The mastery of the instructional strategy did not happen in isolation. All of the teachers in the treatment group required between three and nine weeks of on-going support including modeling, coaching, feedback, and study group participation before they were able to demonstrate mastery. On the other hand, none of the teachers in the comparison group was able to achieve mastery of the strategy and their observation data documented a declining trend in their performance over the study period.

Based on the data, the authors concluded that the RLT Model was an effective method of supporting the implementation of evidenced-based practices. Although the components of the program cannot be separated, all of the teachers felt that the on-going support, observations with feedback, and teacher study group sessions were critical to

their overall implementation fidelity. Unfortunately, no student data were collected during the study. These data would have helped the researchers support their claims of the program's overall effectiveness and improve their overall design. However, it is especially important to note that even though all of the teachers attended the same in-service, the teachers in the comparison group had more experience, and the comparison group teachers were aware of the purpose of the study, they still were not able to successfully implement the strategy without the support of the RLT Model.

As evidence by the above studies, it is vital to provide systematic, on-going support for teachers who are implementing a new instructional strategy using a well-designed teacher support/study group in addition to observations with specific feedback. Although none of the studies specifically addressed the urban, middle school populations, most of the studies do illustrate successful research strategies for investigating facilitated support groups.

Instructional Fidelity

Instructional fidelity is vital to the implementation of any evidence-based instructional program. Without an assessment of fidelity, the reliability of the results will always be in question.

One such qualitative study by Blozowich (2001) investigated the implementation of differentiated instruction strategies in ten sixth-grade, middle school, classrooms in a rural school district in eastern Pennsylvania. The researchers selected a moderately sized middle school for this study with approximately 700 students; the SES and ethnicity percentages were unavailable. The students at this school are placed on interdisciplinary teams in which they were heterogeneously grouped according to their ability except for

mathematics. The students are tracked for mathematics. Students with disabilities represented 11% of the school's population, and students with mild to moderate disabilities were included into the general education classroom for most of their instruction. The researcher collected data using multiple assessments including a professional development survey, a differentiated instruction survey, an unannounced classroom observation using a checklist, and follow-up interview. Although, the researcher developed all of the instruments, no reliability or construct validity information was provided, thus findings must be interpreted with great caution. Through the planned variety of assessment tools, he did however allow for a triangulation of results. He also built into his design member checks with each of the participants to confirm the validity of the results and peer examination of the interpreted data themes to remove the effect of researcher bias.

The results of the study (Blozowich, 2001) revealed disappointing data and themes. First, the survey results did show that although the teacher participants had been exposed to differentiated instruction strategies through county professional development activities, the teachers as a whole did not make any effort to learn any more beyond what they had learned at the workshops. Furthermore, only three of the ten incorporated differentiated instructional strategies into their classroom lesson plans, more than half did not participate in a learning community or collaborate with other teachers, more than half stated they were satisfied with their current teaching strategies, and more than half knew about differentiated instruction and/ or did not wish to learn additional information. From the classroom observations, it was also found that most of the teachers who claimed to be using complex differentiated strategies on the surveys did not demonstrate any evidence

of use when observed. The results from the interviews corroborated this finding. It was found that, despite a school board policy for teachers to include differentiated instructional strategies into their lesson plans, the majority of the ten teachers continued to teach their class like the traditional tracked classroom with very little learner differentiation.

Although this study had a very small sample size with limited reliability and internal validity, it is consistent with previous research findings that indicate if teachers are left on their own without on-going support, despite a district policy, they will continue to teach the same way they always have taught. Thus, researchers should be very cautious of data based on teacher survey responses in relation to the fidelity of instructional practices.

After exhausting the databases for articles on instructional fidelity of research-based practices in the classroom, the literature search was broadened to include mental health studies that focused on the fidelity of behavioral training programs for students. In particular, Webster-Stratton's (2004) chapter on supporting implementation fidelity with The Incredible Years program reported key findings learned from ten years of research using random controlled trials that could easily be adapted for the implementation of instructional strategies in schools.

The focus of the chapter, The Incredible Years Parent, Teacher, and Child Training Program is currently used to promote "positive parent and teacher interactions with children, strengthening children's emotional, social, and self-regulation competence and reducing behavior problems in both prevention and clinic populations" (Webster-Stratton, 2004, p. 1). Throughout the entire chapter, the author stresses the importance of

fidelity when implementing evidence-based interventions, frequently referred to as treatment fidelity. She also references their comprehensive teacher training intervention in which they help teachers implement classroom management and discipline strategies that promote social competence.

As reported by Webster-Stratton (2004) the five key components for effective program implementation with fidelity are:

1. Standardization of treatment delivery using comprehensive clinician manuals, well articulated protocols, videotapes, and materials for parents, teachers, and children
2. Standardized quality training for group leaders delivering the intervention
3. Effective supervision of group leaders
4. On-going fidelity monitoring and certification
5. Agency or administrative support (p. 2)

Although this study represents only one particular area of mental health research, researchers in this field have been investigating implementation fidelity since the mid 1970's. Obviously, the field of education has a lot to learn. In a similar fashion, this proposed study hopes to demonstrate that when teachers are provided adequate resources and support combined with observations and feedback they can effectively implement evidenced-based practices into the classroom.

Another notable mental health study by Mokrue, Elias, and Bry (2005) investigated the effectiveness of a video series, *Talking with TJ*, which is intended to be used by teachers to encourage the development of positive social and emotional skills with urban, predominantly minority, elementary school children. The sample for the

study was 655 second and third graders from 30 classrooms at six urban elementary schools in the Plainfield, New Jersey, School District. The descriptive information from the district revealed that at the time of the study 60% of the students qualified for free or reduced lunch and the district had a high percentage of students of color with 82% African-American, 14% Hispanic, and 4% other.

The instruments used, in addition to a demographic information intake sheet, were the Social Skills Rating Scale (SSRS) developed by Gresham and Elliot in 1990 to measure social competence, the Piers-Harris Children's Self-Concept Scale developed by Piers and Harris in 1984 to measure self-concept, and a Teacher Implementation Survey designed by the authors to measure the degree of fidelity of each instructor. Reliability and validity information was provided for both the SSRS and the Piers-Harris Children's Self-Concept Scale. The SSRS was reported to have an internal consistency reliability of .96, a test-retest reliability of .68 to .87, and validity of .75 to .81 when compared to the Child Behavior Checklist developed by Achenbach and Edelbrock in 1983. The Piers-Harris Children's Self-Concept Scale was reported to have an internal consistency of .85 and a test-retest reliability coefficient of .73.

Because this program was part of the county required curriculum, before data collection could begin passive consent letters were sent home to the families that described the study, asked for the parents' permission, and provided an opt-out option for them if they did not want their child to participate. Only one family chose for their child to not participate in the study. For data collection purposes, the student surveys were administered before and after the video intervention. To assist in the process, two trained research assistants went together to each of the 30 classrooms and administered the

student surveys using scripted directions during two 30 to 45-minute periods so as not to tire the young children. All items were read aloud to the students to accommodate for any reading difficulties in students. This process was then repeated after the students viewed and participated in the intervention program, which lasted approximately four months. The teachers were also asked to complete SSRS forms on each student before and after the intervention. The teachers were paid $22 per hour for completing the surveys on their own time.

Based on the mean of the teachers' self-reported implementation fidelity scores, three of the schools were identified as "high implementation schools" and three were identified as "low implementation schools." The student data were then analyzed using an ANCOVA procedure while controlling for the differences in the teachers' implementation dosage scores. The results from the teachers' data showed that "children in high implementation groups had higher ratings of social skills and lower ratings of problem behaviors during post-assessment period while their counterparts received higher ratings of problem behaviors and lower ratings of social competence" (Mokrue, Elias, & Bry, 2005, p. 68). It is important to note that the initial comparison of the students' self-reported, self-concept scores revealed no significant differences between the two groups. Had the researchers not measured the fidelity of the teachers and conducted a second weighted analysis of the data they would have concluded that the program had no effect on the students' behavior and self-concept.

The necessity for researchers to measure the fidelity of implementation is key to determining the true success of an instructional program. However, the results of this particular study should be interpreted with caution because the researchers relied on self-

reported data. Finally, the results are not generalizable to the majority of schools in America. The results can, in fact, only be generalized to other urban elementary schools.

Summary

From this research synthesis, it is evident there are a limited number of published studies investigating the implementation of differentiated instruction with fidelity on urban middle school students. There does, however, seem to be a growing body of evidence that report success with using teacher support/ study groups to assist teachers with implementing instructional practices learned during professional development inservices. Furthermore, although investigations with differentiated instruction as a complete philosophy and instruction model began with the highly able, gifted learners, recently more researchers have begun to focus on its use and ability to effect the more diverse populations including students with special needs and struggling learners. Currently there is also a predominance of qualitative studies being conducted in this area, precipitating a need for more quantitative and mixed-method studies. Data remain scarce regarding the effect of differentiated instruction on student achievement. Finally, while it may be apparent that teachers are receptive to utilizing a supportive differentiated instructional philosophy, the bottom line is that until there is a body of evidence illustrating its impact on academic achievement and a viable approach to measuring instructional fidelity, district personnel and administrators will be reluctant to support its implementation.

Chapter Three

Method

Overview

The goals of this study were (1) to investigate the effect of differentiated instruction on the mathematics and reading achievement of urban, middle school students; (2) to monitor teacher fidelity to the differentiated instruction model; (3) to assess the effect of facilitated support groups on teacher fidelity; and (4) to evaluate the relationship between teacher implementation of differentiated instruction and student achievement scores. This study incorporated data through a mixed methods design that evaluated the effectiveness of facilitated teacher support groups on the implementation of differentiated instruction in two urban, middle school settings. In addition, this study contributes to the limited body of research that addresses classroom implementation of the differentiated instruction model. This chapter provides information on the study's design, population and sample, variables, measurement tools, data collection procedures, and data analysis.

Pilot Data

Overview of pilot. During the previous year, the principal investigator conducted a pilot to this research study at an urban, middle school in Florida. The purpose of the pilot was to develop and field test the DI: FIT observation tool, field test a facilitated support

group, and collect student achievement data to determine the feasibility of the implementation of the research design. The pilot had a design similar to the current study in that the treatment and control groups were based on interdisciplinary team membership. There were four teams of five teachers each, two seventh grade and two eighth grade. Special effort was made to insure that the groups were similar in size and demographics. The study was conducted over the entire school year and the student achievement data were assessed using the students' individual FCAT Developmental Scale Scores (DSS). These ordinal scores are based on the FCAT Scale Scores and possible scores range from 0 to 3000. As with the current study, teacher fidelity to the differentiated instruction philosophy was encouraged and supported through a facilitated differentiated instruction support group, access to a resource library, and classroom fidelity observations using the DI: FIT once each nine-weeks.

DI: FIT observation tool. The ten teachers in the treatment group were observed four times each by trained observers and the correlation statistic for the DI: FIT observation tool was calculated to be .86 ($p < .0013$) with a 95% confidence interval ranging from .67 to .96 [$CI_{95} = (.67, .96)$]. A correlation statistic of .85 or greater is considered good (Cohen, 1992).

Student achievement. Because the middle school students had a wide range of initial academic abilities, student achievement was measured as their improvement or change over the course of the year with respect to their FCAT Developmental Scale Change Score, posttest minus pretest. With regard to the student academic achievement pilot data, the results showed a great deal of promise for future studies.

Because the reading and mathematics achievement scores are individual, discrete scores, they were analyzed separately. Students with missing data were removed from the data set prior to analyses. The reading and mathematics FCAT Developmental Scale Change Scores were analyzed using a 2x2 factorial ANOVA by treatment level and by grade, with an alpha level of less than .05 considered significant.

Student reading data. The descriptive statistics of the four reading subgroups are presented in Table 1 and illustrate that the four groups were similar in size and distribution with a total sample size of 353. Figure 3 shows the side-by-side comparison of the students' FCAT reading developmental scale mean change score by group, from which it can be easily seen that the students whose teachers were part of the treatment group improved their scores more than the students whose teachers were part of the control group at both grade levels. The analysis of variance data (see Table 2) further support this statement as the treatment effect was found to be statistically significant ($F(1, 349) = 5.41$, $p = .02$) with no significant interaction by grade level. Cohen's (1977, 1988, 1992) effect size (f) was calculated to be .12, which is considered small. Note, Cohen's f is calculated by taking the square root of the product of the degrees of freedom times the F-statistic divided by the total sample size: $f = \sqrt{(df)(F)/N}$. In addition, a Cohen's f of .1 is considered small, .25 medium, and .4 large.

Student mathematics data. As with the reading data, the descriptive statistics of the four mathematics subgroups are presented in Table 3 and illustrate that the four groups were similar in size and distribution with a total sample size of 353. Figure 4 shows the side-by-side comparison of the students' FCAT mathematics developmental scale mean change score by group, from which it can be easily seen that the students

whose teachers were part of the treatment group improved their FCAT Developmental Scale Scores more than the students whose teachers were part of the control group at both grade levels. The analysis of variance data (see Table 4) further support this statement as the model's main effect was found to be statistically significant ($F(1, 349) = 3.27$, $p = .02$). Cohen's (1977, 1988, 1992) effect size (f) was calculated to be .17, which is considered to be a small to medium effect size.

Table 1

Descriptive Statistics of FCAT Reading Developmental Scale Change Scores (N = 353)

Statistic	7th grade with support (n= 92)	7th grade without support (n = 92)	8th grade with support (n = 78)	8th grade without Support (n = 91)
Mean	163.6	122.5	149.3	113.1
Median	158.0	141.5	115.0	104.0
Range	892.0	805.0	704.0	728.0
Interquartile Range	212.5	202.5	166.0	199.0
Standard Deviation	169.2	160.0	148.1	142.9
Skewness	.3	-.3	.3	.2
Kurtosis	.5	-.1	.1	-.1
Standard Error Mean	17.6	16.7	16.7	15.0

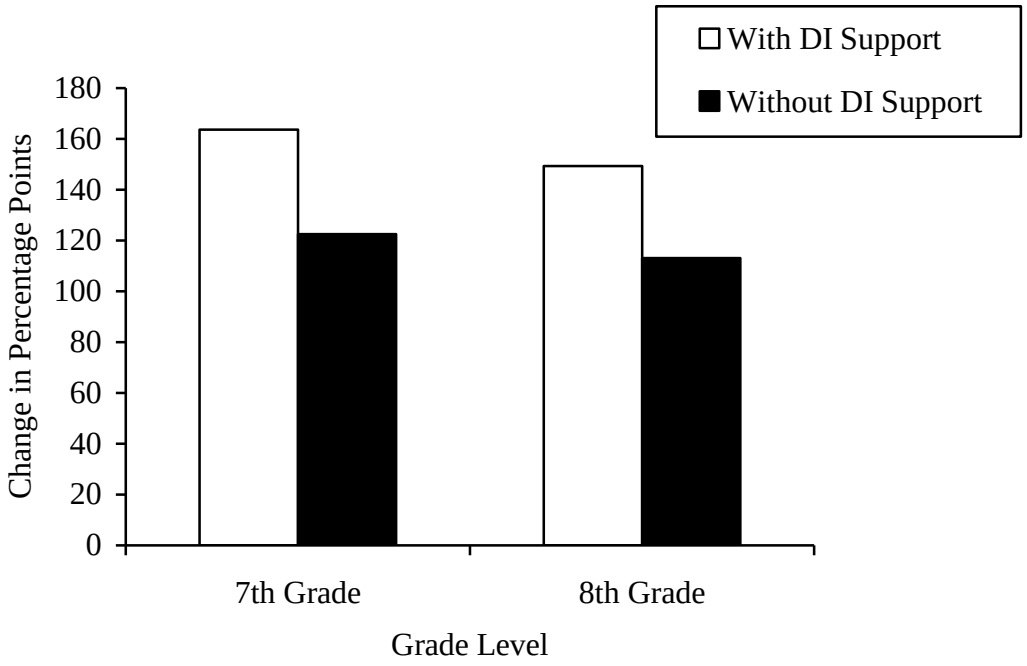


Figure 3. FCAT reading developmental scale change scores by grade level (N = 353).

Table 2

Analysis of Variance Summary Table for FCAT Reading Developmental Scale Change Scores (N = 353)

Source	Df	SS	MS	*F*	*P*
Grade	1	12318.7	12318.7	.51	.48
Treatment	1	131221.3	131221.3	5.41	.02*
Grade x Treatment	1	505.4	505.4	.02	.88
Within Group (Error)	349	8462321.2	24247.3		

*$p < .05$

Table 3

Descriptive Statistics of FCAT Mathematics Developmental Scale Change Scores (N = 353)

Statistic	7th grade with support (n = 92)	7th grade without support (n = 92)	8th grade with support (n = 78)	8th grade without support (n = 91)
Mean	152.5	119.5	123.2	101.3
Median	137.0	124.0	103.5	98.0
Range	557.0	828.0	555.0	551.0
Interquartile Range	112.5	162.5	140.0	107.0
Standard Deviation	103.1	132.0	121.0	89.2
Skewness	.6	-.1	.9	-.3
Kurtosis	1.4	1.1	.4	1.7
Standard Error Mean	10.8	13.8	13.7	9.4

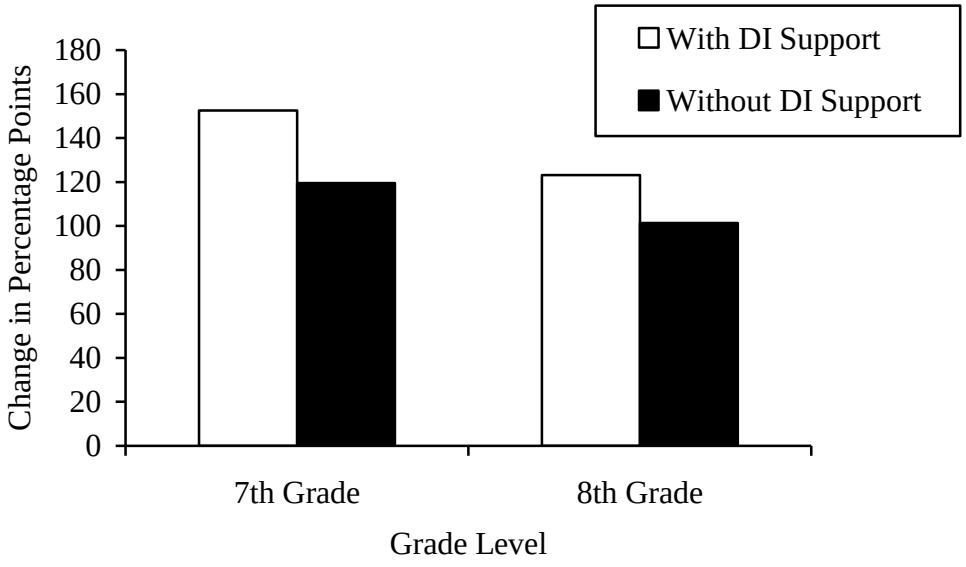


Figure 4. FCAT mathematics developmental scale change scores by grade level (N = 353).

Table 4

Analysis of Variance Summary Table for FCAT Mathematics Developmental Scale Change Scores (N = 353)

Source	Df	SS	MS	*F*	*p*
Model	3	123607.4	41202.5	3.27	.02*
Grade	1	49729.5	49729.5	3.95	.04*
Treatment	1	660.68.5	66068.5	5.24	.02*
Grade x Treatment	1	2704.6	2704.6	.21	.64
Within Group (Error)	349	4396704.2	12598.0		

*$p < .05$

Summary of pilot. The success of the pilot was critical in laying the groundwork for this study. It demonstrated that the DI: FIT was a viable tool for assessing teacher fidelity to the differentiated instruction model, and it provided preliminary evidence that the use of differentiated instruction strategies could affect student achievement. Further, it contributed to the development of the current support group model.

Research Design

A mixed methods design, with a quasi-experimental design in the quantitative component, was utilized to evaluate the multiple themes of this research study over a five-month period. First, qualitative and quantitative methods were used to investigate the impact of facilitated teacher support groups, teacher reflection, and fidelity observations with feedback on the teachers' implementation of differentiated instruction. In addition, a triangulation of data from facilitated support group minutes (group's perspective), teacher

implementation journals (individual's perspective), and differentiated instruction observations (observer's perspective) were utilized to determine the impact of the support group model on teacher implementation fidelity (see Figure 5). The use of multiple sources and perspectives increases the reliability of the study's findings. Finally, at the end of the five-month period, an ANOVA procedure was conducted to quantitatively determine the relationship between the degree of teacher implementation of differentiated instruction and student achievement.

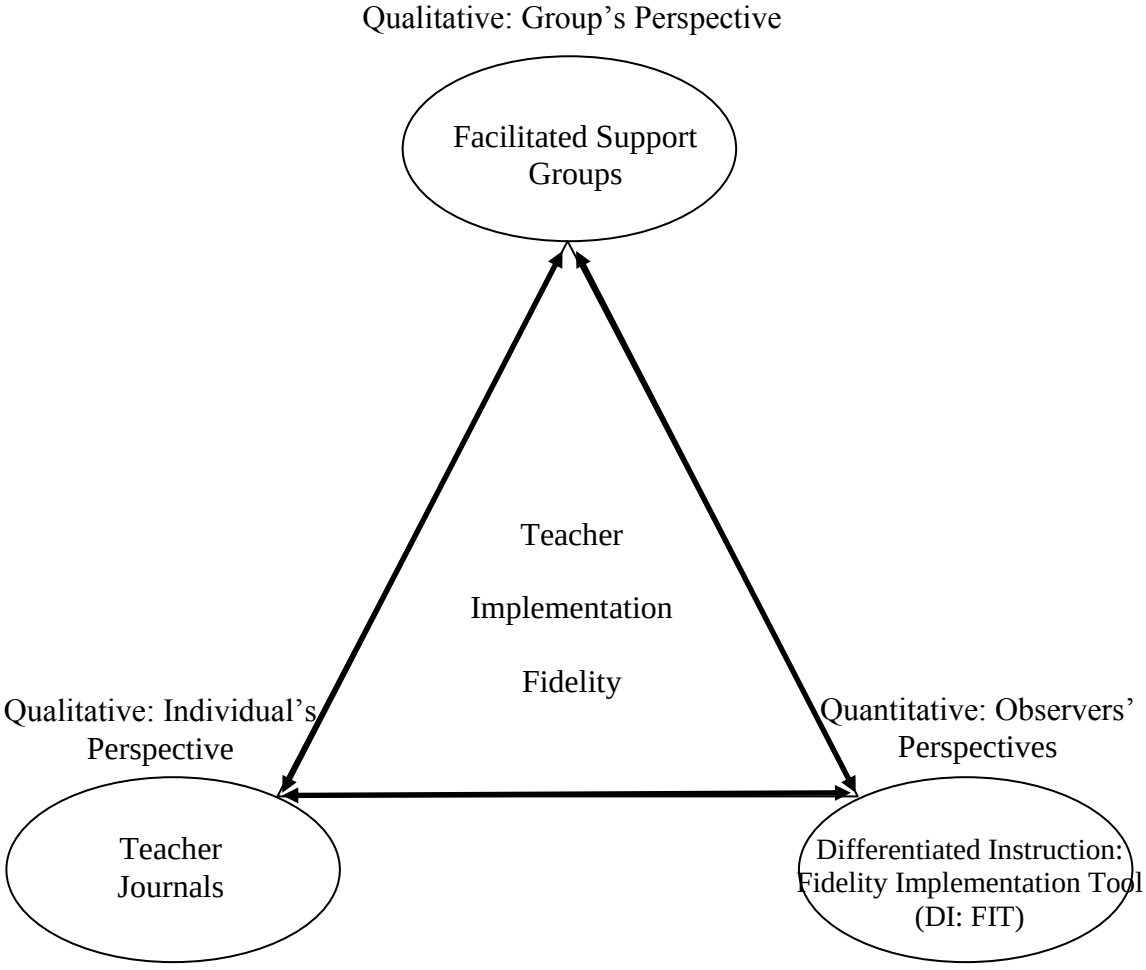


Figure 5. Supporting and assessing teacher implementation fidelity: Triangulation of teacher data.

Sample selection and assignment. During this second research phase, two matched urban, Title I middle schools were purposively selected to serve as research sites. First, permission was received from the district research and compliance office to use the school sites, and then the principal at each site was contacted to explain the study and gain permission. Once permission and support were obtained from each principal and the appropriate IRB permissions were obtained, the study was explained to both faculties. In School A, 28 teachers volunteered to participate and in School B 27, teachers volunteered. The participants were divided into a treatment group and control group based on currently existing interdisciplinary teams within each school because these teachers would be in daily contact with each other and often share strategies. This method of assignment was utilized because it would reduce the amount of cross contamination of the treatment. When determining which teachers would purposely be assigned to the treatment group, preference was given to the content area teams that had the greatest number of students with special needs included in their general education classrooms. The number of teachers selected was distributed among the three grade levels within each school and kept as balanced as possible. The remaining teachers were matched by grade level and content area within each site and assigned to serve as the control group (see Figure 6).

After much discussion and research, this type of assignment was determined to be the most effective because a school vs. school comparison would have produced data with poor reliability and validity. Assigning a treatment and a control group within each site was desirable because it minimized possible extraneous variables and the nesting

effect of individual school factors, such as school-wide reading programs, a strong administration, and extended learning programs.

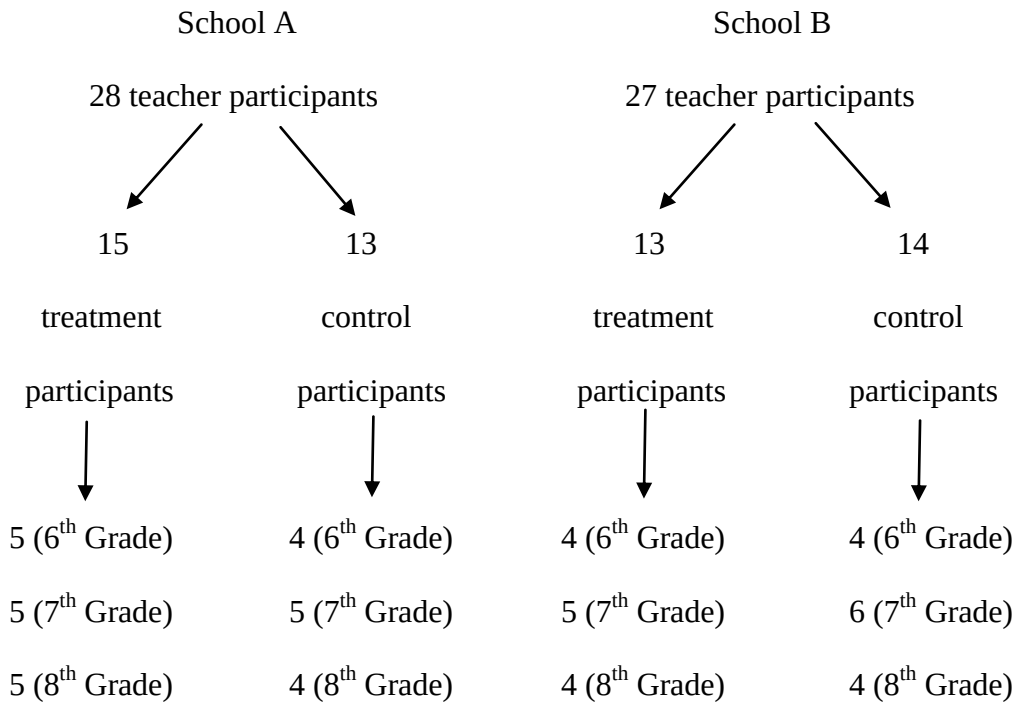


Figure 6. Design of teacher participants assignment by school and grade level (N = 55).

The demographics of each group were kept as balanced as possible. This is especially important because urban, Title 1 schools typically have a high teacher turnover and a large percentage of new teachers. The number of new teachers were dispersed among the treatment and control groups as much as possible and will be addressed in the results section.

Facilitated support group model. As part of the study, the teachers in the treatment and control group first attended the standard, district, seven-hour differentiated

instruction in-service workshop. During the next five months of the school year, each teacher in the treatment group attended five monthly, two-hour, facilitated support group sessions. The first four sessions were similar in format to the district's differentiated instruction book study in that each teacher has access to resources, support from the facilitator and the other members of the group, lesson and strategy support, and they could earn in-service points. The fifth and final meeting was a focus group which assisted the facilitator in obtaining qualitative feedback on the support group model.

The groups began with the typical formalities of introduction; and then after group norms were established, i.e., being on time, taking roll, respecting the opinion of others, and bringing the teacher reflective journal to each meeting so they could take notes and make connections to their classroom instruction, the format for the sessions was standardized. Each meeting began with teachers sharing their classroom differentiated instruction experiences and providing feedback to the group in a round robin fashion. Teachers were encouraged to share both successes and challenges so others could learn from them. Teachers also discussed future differentiated instruction lessons to get ideas from the group. The researcher acted as the group's facilitator and moderated the group's discussion to make sure the teachers stayed on topic and to ensure that all members had an opportunity to speak. Each month, the facilitator also made sure that the meeting space was reserved and arranged in a manner that would facilitate the group's discussion. The facilitator also provided the participants with snacks, refreshments, and an article that highlighted various differentiated classroom strategies in order to facilitate discussion and expand their teaching repertoire. The table portion of the session closed with questions from the participants in a round-robin style and a reminder of the next

session's date. The teachers were then allowed to peruse the school's differentiated instruction resource library and select a new differentiated instruction resource and/or strategy book (see Appendix E).

During the support group sessions, the facilitator's assistant recorded detailed minutes and teacher comments. In order to ensure the reliability of the data collected at the facilitated support group sessions, the minutes were e-mailed out on the following day for verification. All of the teachers in the treatment group were also provided with a *Facilitated Support Group Feedback Form* (see Appendix F) after each session. The form served two purposes, to verify that each member read the minutes and to obtain written feedback from each member regarding the accuracy of the minutes. If there were any changes, the minutes were amended and sent out again to ensure consensus by all of the members. This process was repeated until all members felt that the minutes were accurate and served as a member check. As a backup, the sessions were digitally recorded to insure the accuracy of the information and comments collected. If there was any disagreement, the tape was used for clarification. No teacher's name or identifying information was used in the final report to assure the anonymity of all participants.

This method was field tested during the pilot study and it was found that it provided the necessary information needed for the purposes of this study. At the final group meeting, the facilitator asked the participants to reflect and discuss what they liked and what they would like to change about the facilitated support group sessions and to suggest ways to improve upon the support group model/format for the upcoming year.

Teachers' resource library and reflective journals. The purpose of the differentiated instruction resource library for the teachers in the treatment groups was to

provide an additional opportunity for teachers to find new ideas and strategies and to encourage growth for the group, otherwise the group's idea pool could become stagnant. At the conclusion of each meeting, participants traded-in their book from the previous month and selected a new resource book to review for at least 30 minutes. They were encouraged to implement at least one new differentiated instruction strategy during the next three to four weeks in their classroom. In addition, teachers in the treatment group maintained implementation journals based on their differentiated instruction experiences. Although the teachers in the comparison groups have been exposed to the elements of differentiated instruction through standard district in-service opportunities, they did not receive any of the treatment interventions or additional supports (i.e., feedback from observations, participation in support groups, copies of minutes, use of reflective journals, or access to the reference library.)

Teacher observations. Prior to beginning the observations, each observer was pre-trained using tapes, and the process was practiced until each rating team reached an inter-rater reliability of .85 or better. In order to compare fidelity of the two groups, the principal investigator and research assistant observed and assessed the teachers in both the treatment and the comparison groups each nine-weeks to determine the degree of fidelity that each teacher demonstrated with respect to differentiated instruction using the *Differentiated Instruction: Fidelity Implementation Tool* (DI: FIT). Inter-rater reliability was re-checked each nine-weeks. If the correlation statistic between the trained observer and the principal investigator fell below .85 then the observing team participated in a re-training program with additional observations and follow-up discussions until a correlation of .85 or better was obtained. Additional training on the use of the DI: FIT

assessment tool was conducted if necessary. In this way, the researcher was able to empirically examine the difference in the differentiated instruction implementation scores of the two groups of teachers. This comparison was especially important because some teachers who have attended cooperative learning workshops may utilize some form of group work and/ or learning profile accommodations which does not make the lesson differentiated, but to the untrained observer it may look like differentiated instructional strategies.

Population and Sample

The targeted sample for this study was urban middle school teachers and students in Florida. Specifically, this research project was designed to provide strategies, support, and assessment tools for teachers who use a team approach and support the inclusion of diverse student populations, especially students with disabilities. The students selected for this study represent the diverse populations of students who typically live in the inner-city areas of large urban cities. Large percentages of these students are usually from low-income families and are primarily minorities. Further, the student population included general education students, "at-risk" students, and students with mild to moderate disabilities.

For the current research study, data were collected at two urban middle schools. The two schools were selected because they were closely matched based on 14 different criteria (see Table 5). In addition, these schools had been identified as "needing improvement" because they have not met the NCLB's adequate yearly progress (AYP) requirements for the past three years. The two middle schools selected met each of the following criteria:

- Teach students in sixth-through eighth-grades
- Student enrollment with approximately 500 students
- Located within the central portion of the school district, inner-city
- Located within 10 miles of each other
- More than 80% of the students receive free lunch, Title I
- More than 80% of the students are identified as minority
- More than 10% of the students receive exceptional student education services
- Less than 50% of the students met the state high standards in reading (3 or above on the FCAT)
- Less than 50% of the students met the state high standards in mathematics (3 or above on the FCAT)
- More than 40% of the students did not make gains in reading
- More than 30% of the students did not make gains in mathematics
- More than a third of the faculty has less than three years experience
- The school did not earn a grade of an 'A' during the last three years
- The school did not meet the federal requirements for Adequate Yearly Progress

Within each of the two school sites, three grade level teams of teacher participants were selected based on interest and assigned to the treatment group and three teaching teams of matched teachers were assigned to the comparison group. This method provided a total sample of 28 teacher participants in the treatment group and 27 in the control group. All of the grade level groups were relatively balanced with the largest having six members and the smallest group having four members. Further, there were a large

number of students in the sample. Because each of the six teaching teams provided instruction for a pool of approximately 80 students, approximately 480 students were part of the treatment group pool and 480 students were part of the comparison group pool, for a total estimate of 960 students. Some students were excluded because of excessive absences (>21 days during the study period) or missing data.

Table 5

Descriptive Statistics of the Student Population who Participated in the State Standardized Assessments at the Two School Sites During the 2006-2007 School Year (N = 1026)

School	Number of Students Enrolled in 2006-2007	% of Students on Free or Reduced Lunch	% of Minority Students	% of Students with Disabilities	% Meeting High Standards in Reading	% Meeting High Standards in Math	% Making Reading Gains	% Making Math Gains	% of Lowest 25% Making Learning Gains in Reading	School Grade		
										2004	2005	2006
A	519	85	86	11	42	49	53	66	59	C	B	C
B	507	92	87	12	30	34	53	56	68	C	C	C

Source. Data based on information reported by the FLDOE, retrieved August 16, 2007, from http://schoolgrades.fldoe.org/default.asp

Independent Variables

Experimental group, level one: Treatment group. The treatment group consisted of the middle school teacher participants who received a seven-hour differentiated instruction in-service workshop, which emphasized philosophy, underlying research, strategies, videos, and logistics of a differentiated classroom. The workshop was followed by five monthly facilitated support group sessions, on-going teacher support (in the form

of a mentor), and implementation observations using the DI: FIT with constructive feedback. There were a total of 28 teachers in the treatment group, 15 at School A and 13 at School B.

Experimental group, level two: Control group. The control group consisted of middle school teacher participants who did not participate in any of the differentiated support activities but still participated in all standard whole-school activities. The control group of teachers taught their classes without the benefit of the facilitated teacher support group. Each teacher in the comparison group was observed at least once per nine weeks using the DI: FIT. The student achievement scores and teacher DI: FIT scores will serve as a comparison with the student achievement and teacher fidelity score of the treatment group during the analysis phase of this study. There were a total of 27 teachers in the comparison group, 13 at School A and 14 at School B.

School site. This nominal assigned variable consisted of two urban middle schools, School site A and School site B. The students' academic achievement scores were analyzed first within each school site and then pooled between the two school sites.

Grade level. This ordinal assigned variable was limited to sixth, seventh and eighth-grade designation.

Dependent Variables

DI: FIT – fidelity observation score. Once each nine-weeks trained observers using the DI: FIT observation tool observed each teacher in the study. Each observation produced a discrete, ordinal observation score ranging from 0 to 20 and served as a numerical representation of the number of observable differentiated instruction

instructional strategies utilized by each teacher during the observation and follow-up conference.

FCAT Mathematics Predictor Test scores. In response to State of Florida's requirement for schools to assess and report the continuous progress of their students, especially the lower performing students, the district now requires all schools to assess and report reading and mathematics progress scores on all students. The FCAT Mathematics Predictor Tests were administered in January (Form B) and in May of 2007 (Form C). The FCAT Mathematics Predictor tests yield a discrete, interval percentile score ranging from 1 to 99. A pre-post comparison was calculated using the student scores from Form B and Form C, respectively.

Stanford Diagnostic Reading Test Scores. The Stanford Diagnostic Reading Test – Fourth Edition (SDRT 4) (Psychological Corporation, 1995) was automatically administered in September by district personnel to all middle school students. This score was utilized as a pre-test reading score. These same students were then re-tested in April to determine their growth. This score was used as the post-test reading score. The assessment report provided each student with a discrete, interval percentile score ranging from 1 to 99 and an approximate grade level equivalent. This benchmark test score is important and can be used in place of a student's FCAT score to determine if the student has met the district benchmark for promotion.

Instruments/ Measurement Tools

Differentiated Instruction: Fidelity Implementation Tool (DI: FIT). The DI: FIT assessment tool was developed and field tested during the first phase of this research project in 2005. This observation tool consists of 20 differentiated instruction indicators

that are dichotomously scored; two items have an "NA" option (see Appendix D). The indicators measure all aspects of the lesson including the teacher and students' behavior, collaboration, and lesson planning.

In the development of this instrument, a literature search was first conducted to see if any differentiated instruction fidelity observation tools were already developed and available. From this investigation, two observation tools developed by leaders in the field were located. The first, *The Differentiated Classroom Observation Form* (see Appendix G), was developed by Chapman and King (2005) and the second, the *Teacher/ Peer Reflection on Differentiation* (see Appendix H), was developed by Tomlinson and Allan (2000). Both of these evaluation tools allow the observer to mark on a scale the degree to which a teacher is implementing a particular strategy or demonstrating a behavior. Chapman and King (2005) further subdivided their tool into the following areas: "physical environment," 'teacher behaviors," "student engagement," "materials/ resources," and "instructional strategies" whereas Tomlinson and Allan (2000) subdivided theirs into "general," "content," "process," "product," and "instructional/ management strategies." Although the psychometric properties of these two observation tools were not available, both of these instruments provide the user with a great deal of worthwhile information regarding a teacher's use of differentiated instruction.

For the purposes of this study, a dichotomous observation tool was desired so multiple observers would be able to obtain a higher degree of agreement on the same observation. Because there is no middle ground, it requires the observer to select that the indicator was either evident or not evident. A higher inter-rater reliability will help increase the overall reliability of the observation data and reduce the chances of observer

differences. The dichotomous fidelity observation instrument entitled *Fidelity Instrument for Measuring the Use of Evidence-Based Academic Strategies in Special Education Classrooms (2005)* was used as a guide for layout and wording that would be specific and measurable.

Based on a review of the previously mentioned instruments and the field experiences of differentiated instruction district trainers, a new instrument was developed, the DI: FIT. Next, the DI: FIT was submitted to several professors in the College of Education at the University of South Florida. Specific feedback and suggestions were obtained regarding language, operational descriptions, and feasibility. Based on this expert feedback, several items were changed and specific numbers were added to make items more quantifiable and observable. During this revision, two of the items, "teacher as a facilitator" and "teacher promotes acceptance of differences," had an optional "NA" added because there may not be an opportunity to observe those two indicators in all lessons.

Later, the DI: FIT was submitted to Tomlinson at the University of Virginia. Again, specific feedback was requested concerning each item, the overall structure, and the validity item content. She suggested clarifying the terminology of tiered lessons and adding "2 out of 3" and "at least 2" to several of the items. These suggestions were also incorporated into the final version.

In order to assess teacher fidelity to the differentiated instruction model, a teacher who previously attended a differentiated instruction professional development workshop presented by the researcher was trained on how to use the instrument and specific evidence for each indicator was discussed. Then the researcher and the trained teacher

observed a differentiated lesson together, completing the instrument independently. After the observation, a follow-up interview with the teacher who taught the lesson was conducted by the two observers. Then the two observers compared their completed instruments. On the first comparison, there were three discrepancies, which were discussed until agreement was reached. A second classroom observation was scheduled and the process was repeated. This time the two observers only differed on one indicator. This item was discussed at length until both observers felt confident that they completely agreed on how to code this item during future observations.

During the pilot study, the two researchers then used the instrument to observe ten teachers participants. Because the teachers were several months into the pilot, the scores were high ranging from 16 to 19 out of a possible 20 points. The correlation statistic for the DI: FIT was calculated to be .86 ($p < .0013$) with a 95% confidence interval ranging from .67 to .96 [$CI_{95} = (.67, .96)$]. A correlation statistic of .85 or greater is considered good (Cohen, 1992).

FCAT Mathematics Predictor Tests. The FCAT Predictor Tests, Form B and Form C, were developed by the State of Florida to assess the Sunshine State Standards so the districts and schools could assess, monitor, and report the academic progress and skills of students as part of the continuous progress model. The items were modeled after the FCAT and many items are directly taken from previous test versions. The two tests are parallel test forms and each test consists of 25 items (24 multiple choice and 1 short response/ think, solve, and explain). The standardized mathematics assessment items measure the student's ability to respond to items that test number sense, number and operations, algebra, geometry, measurement, data analysis and probability, problem

solving, and reasoning as defined by the National Council of Teachers of Mathematics (2005).

According to the Florida Department of Education's (FLDOE) FCAT Assessment and *Accountability Briefing Book* (2005), the FCAT, a standards-based test, was developed with the intention of measuring students' achievement of skills and content described in the Sunshine State Standards. In order to ensure the content validity of FCAT, the Florida Department of Education (2004) implemented the following steps for all FCAT items:

> Educators and citizens judged the standards and skills acceptable.
>
> Item specifications were written.
>
> Test items were written according to the guidelines provided by the item specifications.
>
> The items were pilot tested using randomly selected groups of students at appropriate grade levels.
>
> All items were reviewed for cultural, ethnic, language, and gender bias and for issues of general concern to Florida citizens.
>
> Instructional specialists and practicing teachers reviewed the items.
>
> The items were field tested to determine their psychometric properties.
>
> The tests were carefully constructed with items that met specific psychometric standards.
>
> The constructed tests were equated to the base test to match both content coverage and test statistics. (p. 26)

In May 1996, the Florida Department of Education contracted CTB/ McGraw-Hill to develop the original form of the FCAT test for grades 4, 5, 8, and 10. In 1999, the Harcourt Educational Measurement Company was hired to develop the test for grades three through ten. In addition to the use of commercial testing companies and the establishment of the previously mentioned standards for test items, Florida DOE personnel collaborated with practicing Florida educators (e.g., teachers, curriculum specialists) in an effort to promote strong content validity across both the criterion-referenced and norm-referenced measures of the FCAT. Correlations between .70 and .81 were obtained for students tested in the aforementioned grades (Florida Department of Education, 2004).

Four kinds of reliability coefficients were used in the development of the FCAT: internal consistency, test-retest reliability, inter-rater reliability, and reliability of classifications. For any measure of reliability, the reliability coefficient can range from zero to one (0.0-1.00), with a zero score showing a lack of reliable results and a one reflecting extremely consistent results. The most commonly used measure of reliability with the FCAT is internal consistency, because it involves utilizing only one test administration per student. Internal consistency reliability is reported for the FCAT using Cronbach's Alpha and Item Response Theory (IRT) marginal reliability.

For the FCAT, the Cronbach's Alpha and the Kuder-Richardson 20 (KR-20) coefficients are based "on classical test theory" (Florida Department of Education, 2004, p. 25). The KR-20 formula is used with tests that contain items scored as either "correct" or "incorrect." FCAT reliability coefficients use Cronbach's Alpha for the FCAT component, which scores items between 0 and 4, and the KR-20 for the NRT comparison

part. The data on Cronbach's Alpha and IRT marginal reliabilities for the FCAT Mathematics SSS and NRT show strong reliability coefficients between .80 and .90.

Stanford Diagnostic Reading Test – Fourth Edition (SDRT 4) (Psychological Corporation, 1995). Karlsen and Gardner, two leaders in the field of assessment, updated and re-normed the SDRT in 1995. The resulting SDRT 4 was developed using the highest diagnostic standards. It can be administered to groups of students or individually. The test administration for the middle school level takes 85 minutes. All test items are in a multiple-choice format. The score report provides both criterion-referenced and norm-referenced scores. The test assesses each student's vocabulary, fluency, and comprehension skills. The diagnostic report provides sub-scores with reference to the student's skills in phonetic and structural analysis, vocabulary, literal and inferential comprehension of functional and recreational reading material, and reading rate. Reliability of greater than .85 was reported using internal-consistency measures. Test validity of greater than .85 was determined using the OLSAT-8.

Data Collection Procedures

Facilitated support groups and focus group. Over the course of a five-month period, each teacher attended monthly, 120-minute, facilitated support group sessions. During the teacher support sessions, the facilitator's assistant recorded detailed minutes and teacher comments. In order to ensure the reliability of the data collected at the facilitated support group sessions, the minutes were e-mailed out on the following day for verification and a feedback form was put in each teacher's box. Then, each participant was asked to respond if he/she felt the information was accurate, saw any changes, and/or had suggestions for the next meeting. If there were any changes, the minutes were

amended and sent out again to insure consensus by all of the members as a member check. As a backup, the sessions were recorded to insure accuracy of the information and comments collected. The facilitator also kept a reflection journal that was completed immediately after each session. In the reflective journal, the facilitator recorded specific events that might be important, connections that participants made during the meetings, notes on how to improve the process, and any general themes or emergent meaning that became apparent during the facilitated support group sessions. At the final group meeting, a focus group was conducted, the session was taped, and later transcribed. The facilitator provided a question/graphic organizer (see Appendix I) for the participants to capture what the teachers felt were the most and least valuable aspects of the facilitated support group sessions, suggested changes for future groups, and feedback on their overall differentiated instruction implementation experience.

Teacher journals. During the five-month period, each teacher maintained an implementation journal based on their experiences with the differentiated instruction philosophy and lessons. These journals helped to provide an alternative avenue for teachers to provide personal feedback on the study to the researcher. They were encouraged to write in a free response style in the books after concluding differentiated lessons, glue examples of student work, comment about the support group meetings, note strategies and lesson ideas for future use, and to reflect on how their students responded to differentiated lessons. The facilitator monitored the journals at the monthly meetings to ensure that the teachers were maintaining them. One teacher really liked expressing herself using the journal format, and her journal was used as an example for other participants. The journals were collected at the last teacher support group for analysis.

DI: FIT teacher observations. Once per nine-weeks, each teacher participant, from both the treatment and comparison groups, was observed using the DI: FIT observation tool. In addition to the researcher, two resource specialists, one at each school site, were trained to observe the teachers in the study. First, the observers attended a seven-hour training, taught by the researcher, so consistency of information could be maintained. Then, the observers practiced simultaneously observing and scoring a teacher who was not part of the study. Following the observation, the scoring of each indicator was compared and any discrepancies were discussed until everyone agreed. This process was repeated until the inter-rater reliability of the observers was greater than or equal to .85. Throughout the study, if the DI: FIT observation scores ever differed by more than a point then the observers repeated the inter-rater reliability process to insure consistency and the validity of the scoring process.

FCAT Mathematics Predictor Tests - Form B and C and Stanford Diagnostic Reading Test – Fourth Edition. Students' reading and mathematics achievement pre and post-data were collected on all students whose teachers were participants in the study as part of the county's routine standardized assessment procedure. All student assessment tests were supervised and administered following the state required procedures. All students were supervised to insure independence of the results. Once the student scores were obtained at each school, school personnel sorted the data by grade and by team and removed the student identifiers before providing the deidentified data to the researcher. Students who did not have achievement scores available from both the pre and post-tests were removed from the data set, because they were not present for the entire treatment period. In addition, a high degree of student absence could possibly limit the potential for

academic improvement due to the treatment. Therefore, because the state of Florida considers students who miss 21 or more days of school in one year excessive and a criteria for determining good schools, students who missed more than 21 days of school during the treatment period were also removed from the data set because they would have had limited exposure to the treatment.

Confidentiality

All written data, audiotapes, and videotapes were anonymously coded and stored in a secured file cabinet in the researcher's office. Only the researcher and major professor had access to the data. The researcher maintained physical possession of the data and ensured the safety of participants and confidentiality of the data. Data were safely stored after each observation and monthly facilitated support sessions. The signed informed consent forms (see Appendix J) will be stored for three years in a secured file cabinet in the researcher's office along with all data, tapes, and notes. After that time, these documents will be shredded or destroyed.

In accordance with the IRB 2006 requirements:

> The results of this study may be published. However, the data obtained from the participants will be combined with data from others in the publication. The published results will not include names or any other information that would personally identify the participants in any way. Furthermore, the privacy and research records of all participants will be kept confidential to the extent of the law. Authorized research personnel, employees of the Department of Health and Human Services, and the USF Institutional Review Board, and any other

individuals acting on behalf of USF, may inspect the records from this research project.

Data Analysis

Question 1:

What were the effects of differentiated instruction with teacher support during a five-month period on the academic achievement outcomes of urban, Title I, middle school students?

A quasi-experimental design was used to quantitatively compare the impact of differentiated instruction strategies on the reading and mathematics achievement scores of middle school students whose teachers were part of the intervention or comparison groups. This analysis was accomplished using the pretest-posttest data as measured by the FCAT Mathematics Predictor Tests Form B and C and Stanford Diagnostic Reading Test – Fourth Edition. The scores were then analyzed using an 2 (treatment/ comparison) x 3 (6^{th} grade/ 7^{th} grade/ 8^{th} grade) factorial Analysis of Variance (ANOVA), the level of significance was set at .05 using Statistical Analysis System (SAS). The data from students who were not present for the entire study and did not have both pre and post-test scores available were removed from the data set. Students who missed 21 or more days of school during the treatment period were also removed from the data set. Finally, the Cohen's (1977, 1988, 1992) *f* effect size was calculated on comparisons that were found to be statistically significant (see Table 6).

Question 2:

What were the statistical differences among teacher groups who participated in facilitated support groups and those who did not with respect to their implementation of differentiated instruction as measured by the Differentiated Instruction: Fidelity Implementation Tool (DI: FIT) observation tool?

In this quantitative analysis, the DI: FIT observation scores at Time 1 and Time 2 for all teacher participants separated by treatment, grade, and school and then basic descriptive statistics were run in order to examine the differences among all groups (see Table 6).

Question 3:

What was the relationship between the teachers' differentiated instruction implementation scores as measured by the DI: FIT and the students' achievement scores?

In order to answer this question, the mean teacher DI: FIT observation scores were analyzed by treatment group and school site using a 2x2 factorial Analysis of Variance (ANOVA). This analysis of the DI: FIT observation scores provided information as to whether on not a statistical difference existed among the teacher differentiated instruction implementation fidelity scores by treatment group and by school. The Statistical Analysis System (SAS) software was utilized for this analysis with the level of significance set at .05. Finally, the Cohen's (1977, 1988, 1992) *f* effect size was calculated on comparisons that were found to be statistically significant

Next, the teachers' DI: FIT fidelity implementation scores were compared to the reading and mathematics achievement scores of their students. Specifically, treatment and

control groups were analyzed to determine if a correlation existed between the teachers' DI: FIT observation scores and their students' mean academic change scores. This analysis was completed using a SAS correlation procedure. Correlations that were significant at the .05 level would suggest that a relationship exists between implementation fidelity and student achievement scores (see Table 6).

Question 4:

Using qualitative data and feedback provided by the teachers in the treatment groups, what were the teachers' perceptions of the facilitated support group model and their instructional growth?

This final question required the researcher to complete a qualitative analysis on the detailed minutes from the facilitated teacher support sessions, the teachers' implementation journals, the facilitator's reflective journal, and the feedback obtained at the final focus group session. Using an inductive analysis in conjunction with a document review, all data were reviewed, meaningful units were identified, units of data were coded, and then the data were categorized in order to identify basic themes. Then, the data were further reduced through a constant comparison, a consolidation of any redundant categories, and an analysis of emergent themes. These multiple methods of data collection allowed for method and data triangulation and increased credibility (see Table 6).

Table 6

Table of Research Questions, Data Collected, and Analyses Conducted

Research Question	Data	Analyses
Question 1:		
What were the effects of differentiated instruction with teacher support during a five-month period on the academic achievement outcomes of urban, Title I, middle school students?	♦ Student achievement mathematics and reading scores at Time 1 and Time 2 separated by treatment group, grade, and school	♦ Descriptive analysis of demographic and achievement data ♦ $T_2 - T_1$ = Change Score ♦ T-test analysis of T_1 scores between school sites and within sites to ensure that data are not statistically different before secondary analysis ♦ 2x3x1 Factorial ANOVA ♦ Effect score (Cohen's *f*)

(Table continues)

Table 6 (Continued)

Research Question	Data	Analyses
Question 2:		
What were the statistical differences among teacher groups who participated in facilitated support groups and those who did not with respect to their implementation of differentiated instruction as measured by the Differentiated Instruction: Fidelity Implementation Tool (DI: FIT) observation tool?	♦ DI: FIT observation scores at Time 1 and Time 2 for all teacher participants separated by treatment, grade, and school	♦ Descriptive analysis of demographic and observation data ♦ $(T_1 + T_2) / 2$ = Mean Observation Score
Question 3:		
What was the relationship between the teachers' differentiated instruction implementation scores as measured by the DI: FIT and the student achievement scores?	♦ Teachers' mean DI: FIT Score ♦ Students' mean change scores by teacher	♦ 2x2x1 Factorial ANOVA ♦ Effect score (Cohen's *f*) ♦ Correlation analysis ♦ R^2

(Table continues)

Table 6 (Continued)

Research Question	Data	Analyses
Question 4: Using qualitative data and feedback provided by the teachers in the treatment groups, what were the teachers' perceptions of the facilitated support group model and their instructional growth?	♦ Monthly minutes ♦ Individual Teacher Feedback Forms ♦ Teacher Journals ♦ Facilitator's Journal ♦ Focus Group minutes and transcripts	♦ Document review ♦ Inductive analysis ♦ Constant comparative analysis ♦ Identification of emergent themes

Chapter Four

Results

Overview

As previously stated, the goals of this study were (1) to investigate the effect of differentiated instruction on the mathematics and reading achievement of urban, middle school students; (2) to monitor teacher fidelity to the differentiated instruction model; (3) to assess the effect of facilitated support groups on teacher fidelity; and (4) to evaluate the relationship between the degree of teacher implementation of differentiated instruction and student achievement scores. However, prior to the analysis of the achievement data, a complete investigation of the demographics of the teacher and student samples at the two school sites had to be completed to ensure a like comparison of data. Further, the achievement data at Time 1 from the students had to be statistically compared to determine if the results could be compared within each site and/or combined between sites.

Once all of the teacher fidelity observations, support group meetings, and student pre and posttests were completed, then began the task of entering and analyzing thousands of achievement scores. The meticulous data entry process began by sorting scores by school, by grade, and by treatment (teacher teams). Next, any students who missed more than 21 days of school during the treatment period or had missing data due

to an absence or attrition, being withdrawn, were removed from the data set before any further analyses were completed. This resulted in the data from 46 students being removed from School A, leaving a total of 473 participants, and 74 students from School B, leaving a total of 433 participants. In order to ensure the accuracy of the data entry and increase reliability, the SAS reports were triple checked, twice by the researcher and once by the trained teacher observer who had been providing support throughout this study. The descriptive data analyses and ANOVAs were all completed using the 2007 version of the Statistical Analyses Software (SAS, Release 9.1). The results are presented in several sections: demographic statistics of participants, student achievement data analyses, teacher fidelity observation data analyses, interaction between fidelity and achievement, and support group analyses.

Demographics of Participants

Student demographics. It is important to insure that the student populations of the two-school sites were as similar as possible for future analyses. To assess the demographic composition of the student participants at each school a detailed frequency and percentage analysis was conducted with the assistance of school personnel. First, the data were analyzed by school and then pooled so a population comparison could be completed (see Table 7). The student demographics are remarkably similar, except that school site B has a heavier percentage (12.4%) of 6^{th} graders, 7.5% less 8^{th} graders, and 6.2% more students who are categorized as economically disadvantaged. Next, the student samples within each school site were further subdivided and analyzed by treatment group.

Table 7

Demographic Characteristics of Student Participants: Frequency and Percentage by School Site (N = 906)

	School A (n = 473)		School B (n = 433)		Total (n = 906)	
Variable	*f*	%	*f*	%	*f*	%
Grade						
6th	113	23.9	157	36.3	270	29.8
7th	167	35.3	132	30.5	299	33.0
8th	193	40.8	144	33.3	337	37.2
Gender						
Female	237	50.1	212	49.0	449	49.6
Male	236	49.9	221	51.0	457	50.4
Ethnicity						
Caucasian	65	13.7	58	13.4	123	13.6
African-American	303	64.1	295	68.1	598	66.0
Hispanic	97	20.5	79	18.2	176	19.4
Asian	6	1.3	1	.2	7	.8
Native American	2	.4	0	0	2	.2
Economically Disadvantaged	402	85.0	395	91.2	797	88
English Language Learners	35	7.4	36	8.3	71	7.8
Students with Disabilities	41	8.7	49	11.3	90	9.9

When the student participants were analyzed by group within each school site (see Table 8) the consistency of data continued. Again, there were relatively few differences among the samples. The percentages of student ethnicities, gender, students who were identified as economically disadvantaged, and students who were identified as English Language Learners were extremely similar across all four subgroups. Because the class rolls of each teach could not be manipulated, some small differences were expected to naturally exist. The only marginal differences were (1) within School B there were slight percentage differences by grade level (a difference of 16.9% in 6th grade and 10.3% in 7th grade) due to random assignment within classes and (2) within both sites, there were differences in the percentage of students with disabilities assigned to the groups. This was a direct result of the treatment assignments provision that teaching teams who support inclusion of students with disabilities in the general education environment were given preferential assignment into the treatment groups. Because the majority of students with disabilities are functioning below grade level, this added challenge of having a larger percentage of students with disabilities on their team meant that to show improvement when compared to the control groups they would have to increase their academic score significantly over their non-disabled peers. Thus is the hope that the differentiated instructional philosophy can help bridge the achievement gap for students with disabilities in the general education environment and increase access to the general education classroom.

Table 8

Demographic Characteristics of Student Participants: Frequency and Percentage by Group (N = 906)

	School A				School B			
	Treatment Group (n = 245)		Control Group (n = 228)		Treatment Group (n = 216)		Control Group (n = 217)	
Variable	*f*	%	*f*	%	*f*	%	*f*	%
Grade								
6th	66	26.9	47	20.6	60	27.8	97	44.7
7th	82	33.5	85	37.8	77	35.6	55	25.3
8th	97	39.6	96	42.1	79	36.6	65	30.0
Gender								
Female	121	49.4	116	50.9	105	48.6	107	49.3
Male	124	50.6	112	49.1	111	51.4	110	50.7
Ethnicity								
Caucasian	33	13.5	32	14.0	28	13.0	30	13.8
African-American	157	64.1	146	64.0	147	68.1	148	68.2
Hispanic	52	21.2	45	19.7	41	19.0	38	17.5
Asian	2	1.2	4	1.8	0	0	1	.5
Native American	1	.4	1	.4	0	0	0	0
Economically Disadvantaged	208	84.9	194	85.1	199	92.1	196	90.3
English Language Learners	15	6.1	20	8.8	17	7.9	19	8.8
Students with Disabilities	35	14.3	6	2.6	36	16.7	13	6.0

Teacher demographics. Since the study had the approval and support of both the school district and principal at each site, getting teachers to volunteer was relatively easy. Of all the possible interdisciplinary subject area teachers who were eligible for the study only one opted not to participate, which eased the commencement of the study and did not impact the study because she taught students who were cross-teamed and would not have been selected for assignment due to possible cross-contamination of treatment.

Next, in order to complete the demographic analysis of participants, basic demographic data were also collected on the teacher participants. After consents were signed, information regarding each teacher's grade level assignment, gender, ethnicity, certification status, and years of teaching experience was collected and is displayed in Table 9. The grade level assignment of teachers, grades 6 – 8, was purposely balanced and the resulting grade level percentages ranged from 28.6% to 42.9%. As expected, all groups were comprised of predominately female teachers, ranging closely from 71.4% to 86.7%. Conversely, the percentage of male teachers ranged from 13.3% to 28.6%. With regard to teacher ethnicities, the predominate category was Caucasian for all four groups (57.1% to 69.2%), followed by African-American teachers (30.8% to 38.5%). Certification data indicated a possible advantage for the control groups because they had only one or zero teachers that were uncertified, as opposed to the treatment groups who each had two teachers who were uncertified. Uncertified teachers either are teachers who are out of field, working on certification, or are in an alternative certification program. Likewise, the treatment groups had a slightly higher percentage of teachers with zero to three years experience. Thus, no significant differences existed that would warrant caution when interpreting the final data.

Table 9

Demographic Characteristics of Teacher Participants: Frequency and Percentage by Group (N = 55)

	School A				School B			
	Treatment Group (n = 15)		Control Group (n = 13)		Treatment Group (n = 13)		Control Group (n = 14)	
Variable	*f*	%	*f*	%	*f*	%	*f*	%
Grade								
6th	5	33.3	4	30.8	4	30.8	4	28.6
7th	5	33.3	5	38.5	5	30.8	6	42.9
8th	5	33.3	4	30.8	4	30.8	4	28.6
Gender								
Female	13	86.7	10	76.9	10	76.9	10	71.4
Male	2	13.3	3	23.1	3	23.1	4	28.6
Ethnicity								
Caucasian	10	66.7	9	69.2	8	61.5	8	57.1
African-American	5	33.3	4	30.8	5	38.5	5	35.7
Hispanic	0	0	0	0	0	0	1	7.1
Asian	0	0	0	0	0	0	0	0
Native American	0	0	0	0	0	0	0	0
Certified								
Yes	13	86.7	13	100.0	11	84.6	13	92.9
No	2	13.3	0	0	2	15.4	1	7.1
Teaching Experience								
0-3 years	9	60.0	6	46.1	5	38.5	4	28.6
4-9 years	3	20.0	4	30.1	5	38.5	7	50.0
>10 years	3	20.0	3	23.1	3	23.1	3	21.4

Student Achievement Data Analyses Overview

In the following sections, the summarized data and results of all student academic achievement analyses are presented in order to address the first research question:

> What were the effects of differentiated instruction with teacher support during a five-month period on the academic achievement outcomes of urban, Title I, middle school students?

Due to the large amount of data, reading and mathematics results will be separated into two sections and then further subdivided according to the data, time, and procedures that were performed.

Reading Achievement Analyses

Within site comparisons at Time 1. In order to assess the effect of the differentiated instruction and teacher support groups on reading achievement, all students were administered the Stanford Diagnostic Reading Test – Fourth Edition pretest and then comparative analyses on data were completed. First, the pretest results of the student achievement data at the beginning of the study, Time 1, were compared within each school site to demonstrate that the treatment and the control groups were not statistically different at the beginning of the study. The general descriptive statistics from both school sites at Time 1 portrays a relatively normal distribution of data with nothing remarkable to note (see Tables 10 and 11). Additionally, bar graphs of the mean, grade level reading test scores at Schools A and B by treatment group illustrates the closeness of the data at Time 1 (see Figures 7 and 8). The range of group means were relatively close with School A's mean scores ranging from 47.2 to 51.2 and School B's mean scores ranging from 41.7 to 44.2. Finally, an analysis of variance on the same data within each school

site demonstrated that the data were not statistically different at School A ($F(5, 466) = 1.01$, $p = .4085$) nor at School B ($F(5, 421) = 0.30$, $p = 0.9119$) (see Tables 12 and 13).

Table 10

Descriptive Statistics of Reading Scores for School A: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 1 (N = 472)

Group	N	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	65	47.2	13.0	.20	-.54	23	76
Grade 6 without Support	47	50.8	11.9	.42	.08	30	85
Grade 7 with Support	82	47.3	14.6	.58	-.35	21	81
Grade 7 without Support	85	49.2	15.2	.32	-.37	21	86
Grade 8 with Support	97	48.9	16.4	.03	-.66	16	86
Grade 8 without Support	96	51.2	14.2	-.25	-.56	20	79

Table 11

Descriptive Statistics of Reading Scores for School B: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 1 (N = 427)

Group	N	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	60	41.7	16.0	.44	-.17	13	82
Grade 6 Without Support	97	43.1	18.4	.56	-.15	11	91
Grade 7 with Support	77	42.1	12.0	.40	.31	11	76
Grade 7 Without Support	55	44.1	14.6	.19	-.52	13	74
Grade 8 with Support	76	42.2	11.6	.34	-.22	22	73
Grade 8 without Support	62	44.2	17.4	-.01	-.75	8	79

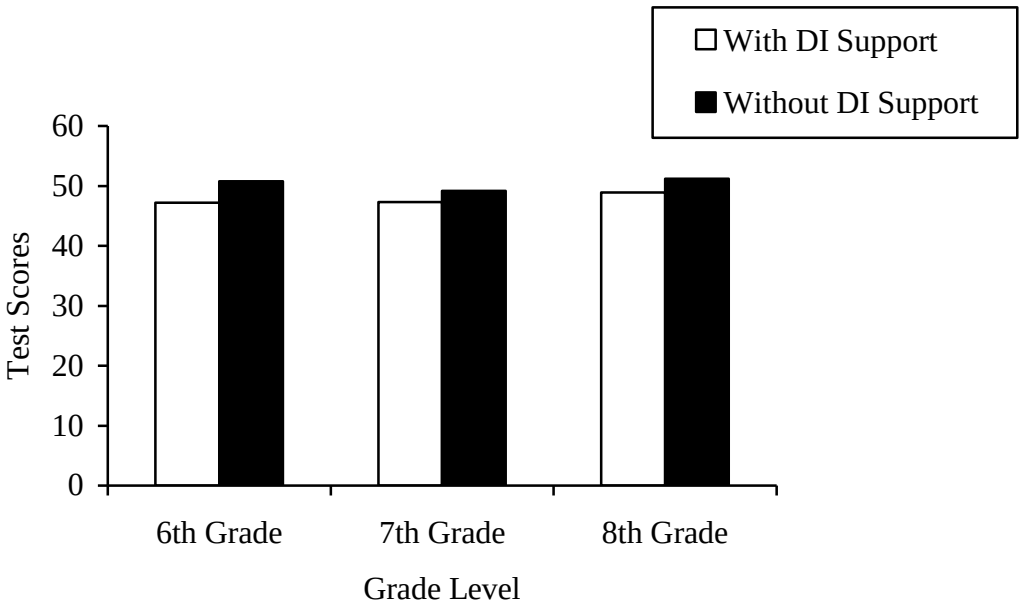


Figure 7. Mean grade level reading test scores at School A by treatment group at Time 1 (N = 472).

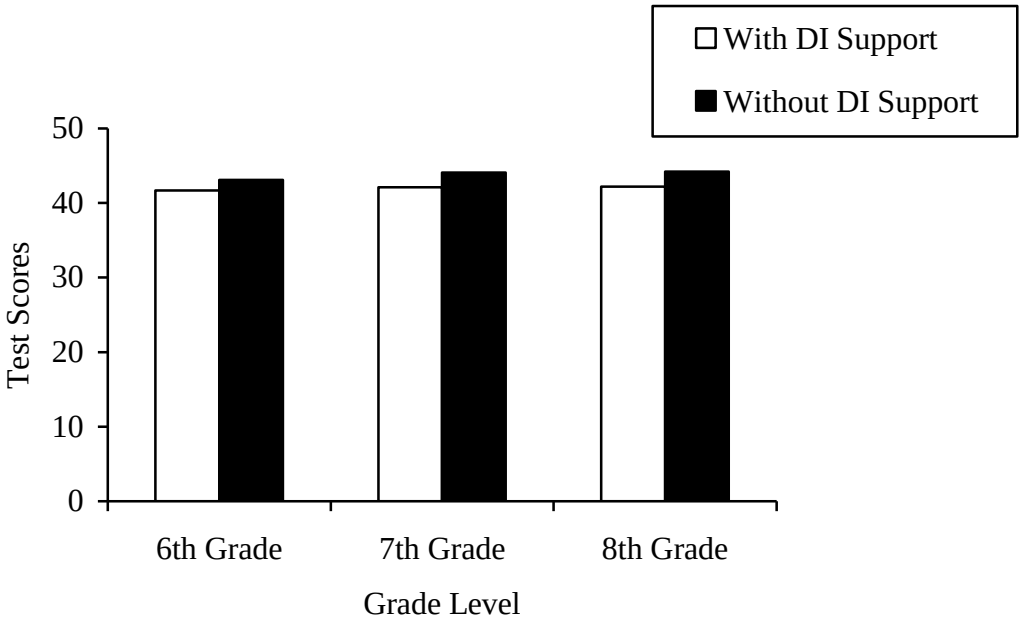


Figure 8. Mean grade level reading test scores at School B by treatment group at Time 1 (N = 427).

Table 12

Analysis of Variance Summary Table for School A's Within Site Comparison at Time 1 (N = 472)

Source	df	SS	MS	*F*	*p*
Model	5	1077.9	215.6	1.01	.4085
Grade	2	313.0	156.5	.74	.4794
Treatment	1	744.1	744.1	3.50	.0620
Grade x Treatment	2	46.2	23.1	.11	.8970
Within Group (Error)	466	99039.2	212.5		

Table 13

Analysis of Variance Summary Table for School B's Within Site Comparison at Time 1 (N = 427)

Source	df	SS	MS	*F*	*P*
Model	5	354.1	70.8	.30	.9119
Grade	2	58.8	29.4	.13	.8823
Treatment	1	325.2	325.2	1.38	.2400
Grade x Treatment	2	7.7	3.9	.02	.9837
Within Group (Error)	421	98911.8	234.9		

Between school comparisons at Time 1. For the next level of analysis, the reading pretest scores were compared between Schools A and B at Time 1 by treatment level. Although the means appear to be similar (see Table 14), a t-test at a significance level of .05 showed that the support groups ($t_{(455)} = 4.5$, $p < .0001$) and the control groups ($t_{(440)} = 4.5$, $p < .0001$) between school sites were statistically different. Because of this statistical level of difference, the next level of analysis at Time 2 was completed on each school separately.

Table 14

Descriptive Statistics of Student Reading Scores at Time 1 by Treatment Level Between School Sites (N = 899)

Variable	n	M	SD
Treatment Groups			
School A	244	47.9	14.9
School B	213	42.0	13.1
Control Groups			
School A	228	50.4	14.1
School B	214	43.6	17.1

Time 2 analyses. At the end of the school year, all students were again administered the Stanford Diagnostic Reading Test – Fourth Edition. The student reading achievement posttest results at Time 2 were then compared within each school site. The general descriptive statistics from both school sites at Time 2 portray a relatively normal distribution of data for most of the groups (see Tables 15 and 16), although a few of the

groups showed some minor deviation. For example, the sixth-grade group without support at School B showed a slight positive skewness (.83) and four of the 12 sub-groups were slightly platykurtic [School A: sixth-grade without support (-1.24), seventh-grade with support (-1.07), and eighth-grade without support (-1.04); School B: seventh-grade without support (-1.36)]. However, none of these conditions was significant enough to cause concern. Additionally, the bar graph of the mean, grade level reading test scores at Schools A and B by treatment group visually illustrate the improvement of the treatment groups' reading achievement scores over those of the control groups (see Figures 9 and 10).

Table 15

Descriptive Statistics of Reading Scores for School A: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 2 (N = 472)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	65	61.6	16.3	-.34	-0.26	24	91
Grade 6 without Support	47	55.6	17.9	-.07	-1.24	24	86
Grade 7 with Support	82	63.8	14.7	.05	-1.07	36	90
Grade 7 without Support	85	58.8	14.9	-.36	-.01	16	95
Grade 8 with Support	97	59.6	18.1	-.11	-.57	20	98
Grade 8 without Support	96	56.2	17.1	-.14	-1.04	22	92

Table 16

Descriptive Statistics of Reading Scores for School B: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 2 (N = 427)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	60	52.5	17.8	.01	-.74	15	85
Grade 6 Without Support	97	45.8	22.6	.83	-.53	14	96
Grade 7 with Support	77	58.9	13.3	.18	-.05	29	95
Grade 7 Without Support	55	54.2	19.2	-.27	-1.36	19	85
Grade 8 with Support	76	59.3	16.1	.19	-.75	30	94
Grade 8 Without Support	62	49.3	20.4	.14	-.69	12	92

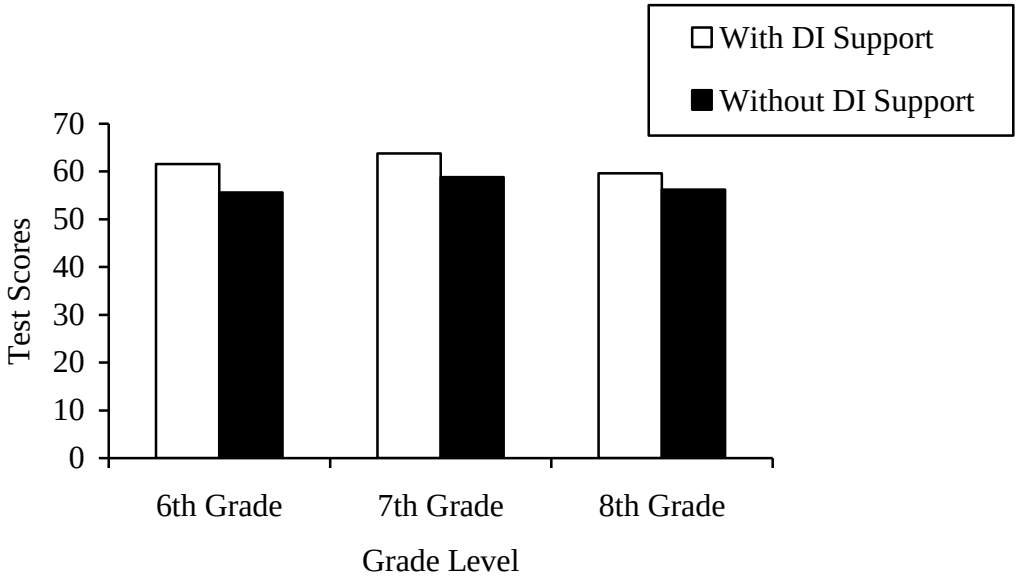


Figure 9. Mean grade level reading test scores at School A by treatment group at Time 2 (N = 472).

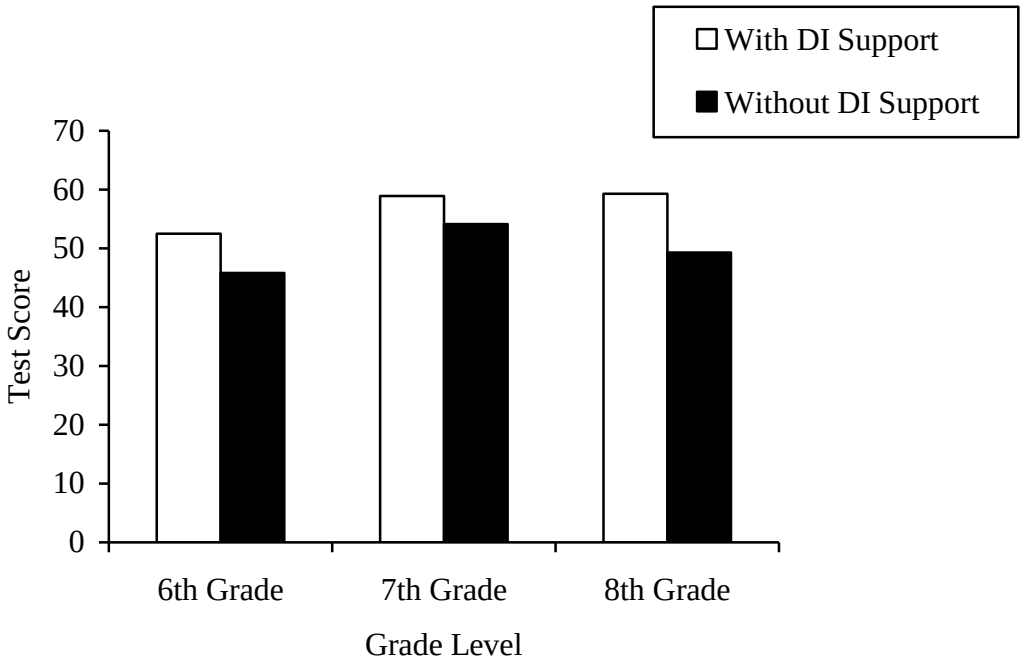


Figure 10. Mean grade level reading test scores at School B by treatment group at Time 2 (N = 427).

Change score analyses. In order to determine the true impact of the differentiated instruction teacher support model, a comparison now needed to be completed on the students' reading improvement from the pretest, Time 1, to the posttest, Time 2. The difference of these two scores will now be referred to as the change score (Time 2 – Time 1 = Change Score). The basic descriptive statistics for the change scores from School A and School B are presented in Tables 17 and 18, respectively. The data illustrate a relatively normal distribution for all twelve subgroups, except two groups that were slightly platykurtic, School A: seventh-grade without support (-.82) and School B: sixth-grade without support (-.83). These two groups are both within the range of reasonable distributions and will have very little effect when combined with other groups.

Table 17

Descriptive Statistics of Reading Change Scores for School A: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group (N = 472)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	65	14.26	15.61	.23	-.57	-14	46
Grade 6 without Support	47	4.81	11.07	.10	.08	-15	36
Grade 7 with Support	82	16.59	11.42	-.04	-.34	-10	38
Grade 7 without Support	85	9.39	15.78	.18	-.82	-19	43
Grade 8 with Support	97	10.78	13.87	.17	-.33	-18	40
Grade 8 without Support	96	4.96	11.80	-.16	-.02	-24	37

Table 18

Descriptive Statistics of Reading Change Scores for School B: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group (N = 427)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	60	10.85	12.62	.26	-0.08	-13	44
Grade 6 Without Support	97	2.77	15.52	.46	-0.83	-20	37
Grade 7 with Support	77	16.51	14.10	-.27	-0.27	-15	48
Grade 7 Without Support	55	9.80	15.44	.29	-0.30	-19	43
Grade 8 with Support	76	17.08	15.44	.13	-0.12	-17	53
Grade 8 Without Support	62	5.27	11.97	-.38	-0.64	-18	30

When the mean reading change scores were graphically compared, the treatment groups had increased their scores to a greater degree than the control groups at both sites (see Figures 11 and 12). In several grade level comparisons, the change for treatment group was as much as three to four times the growth of the control group. The least difference in reading scores between the treatment and control groups by grade level was 5.82 points at School A and 6.71 points at School B. The greatest difference in reading grade level scores was 9.45 points at School A and 11.81 points at School B.

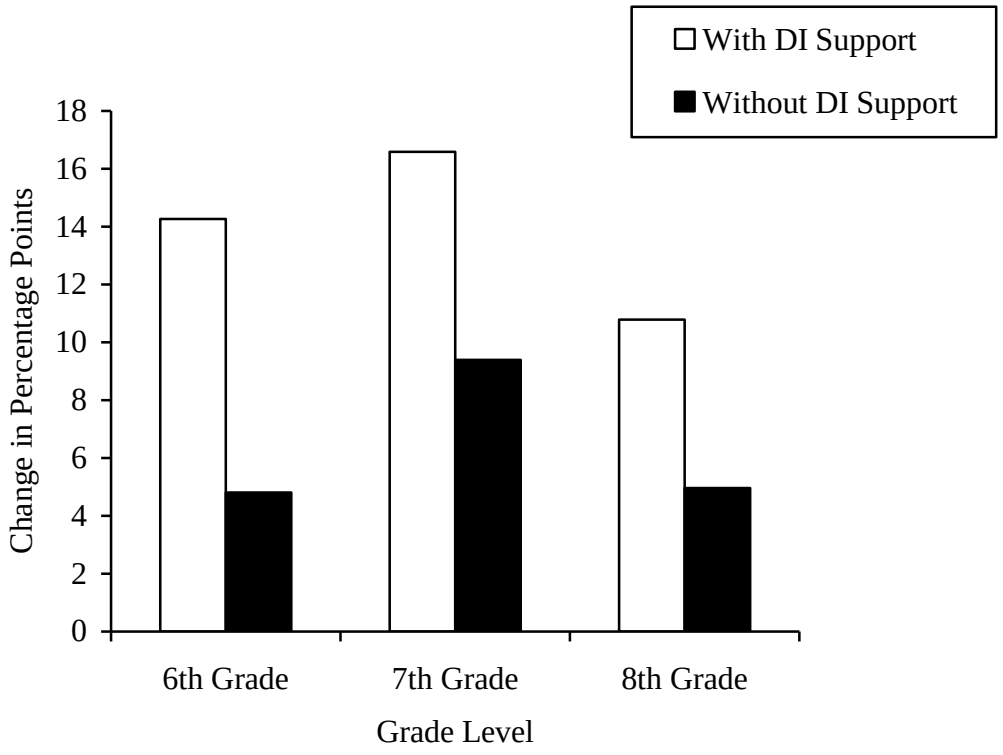


Figure 11. Mean change in School A's reading test scores by treatment level and grade level (N = 472).

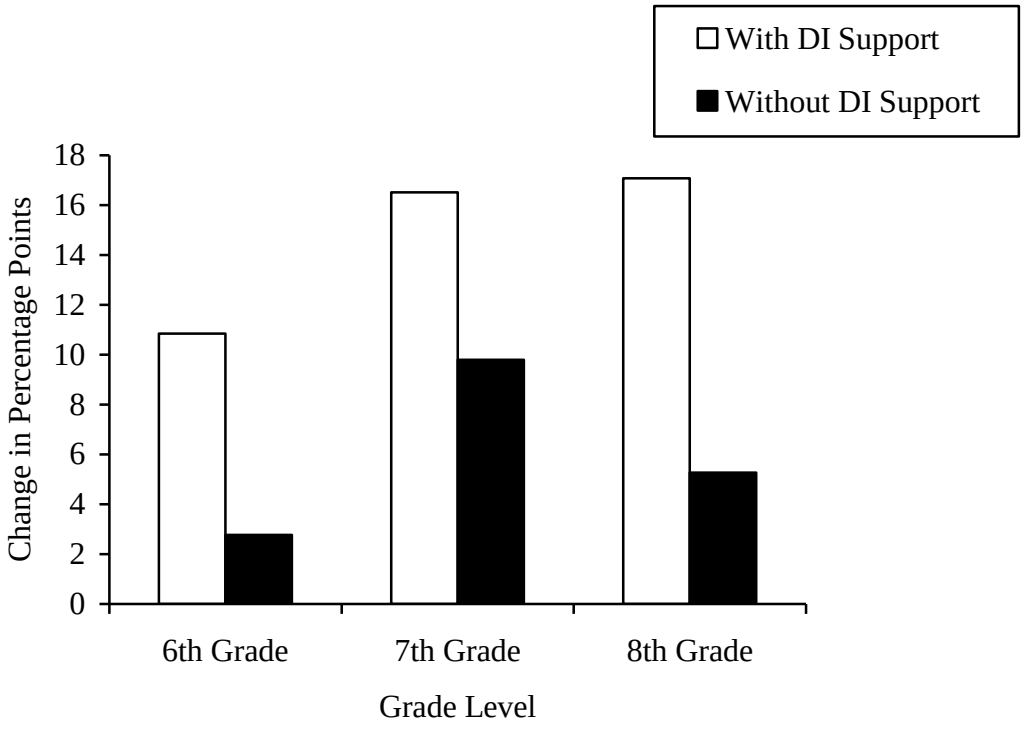


Figure 12. Mean change in School B's reading test scores by treatment level and grade level (N = 427).

In order to rule out chance, the final analysis on the students' reading achievement data was to conduct a 2 (treatment) x 3 (grade level) factorial ANOVA on the resulting change scores using an alpha level of .05 to test for each effect. Before proceeding with the analysis, the assumptions of independence, normality, and homogeneity of variances were investigated. First, because the students worked individually on their assessments and trained teachers proctored the assessment, it is therefore reasonable to assume that the assumption of independence has not been violated. Although the sample sizes are not exactly equal, they are relatively similar and the within group degrees of freedom was 466 for School A and 421 for School B making the sample sizes large enough to expect

robustness to violations of the normality assumption, therefore the normality assumption does not appear to be violated (Cohen, 1992). Finally, when considering the homogeneity of variances, the largest variance ratio was 1.84, less than 2.0, which means the equal variance assumption does not appear to be violated. Therefore, we would expect the ANOVA to be relatively robust to violations of the homogeneity of variance assumption. Based on this analysis of the assumptions, it was reasonable to proceed with the factorial ANOVA (Cohen, 1992).

The results of the two-way factorial ANOVA are presented in Tables 19 and 20. For School A, the model was statistically significant ($F(5,466) = 9.4$, $p < .0001$). Cohen's (1977, 1988, 1992) effect size *(f)* was calculated to be .32, which is generally considered a medium effect. The main effect of the use of support for teachers was also found to be statistically significant using the Type III Sum of Squares data due to the unequal group sizes, ($F(1,466) = 34.29$, $p < .0001$). Cohen's (1977, 1988, 1992) *f* was calculated to be .27, which is generally considered a medium effect size.

For School B, the model was also statistically significant ($F(5,421) = 13.09$, $p < .0001$). The Cohen's (1977, 1988, 1992) effect size *(f)* was again calculated and found to be .39, which is generally considered a large effect. The main effect of the use of differentiated instruction teacher support groups was also found to be statistically significant using the Type III Sum of Squares data, ($F(1,421) = 39.07$, $p < .0001$). Cohen's (1977, 1988, 1992) *f* was calculated to be .30, which is generally considered a medium effect size. It is notable that the reading achievement analyses from both schools indicated the effect of the differentiated instruction teacher support group was statistically significant.

Table 19

Analysis of Variance Summary Table for School A's Reading Change Scores (N = 472)

Source	df	SS	MS	*F*	*p*
Model	5	8507.7	1701.5	9.4	<.0001**
Grade	2	2384.8	1192.4	6.58	.0015*
Treatment	1	6209.7	6209.7	34.29	<.0001**
Grade x Treatment	2	229.4	114.7	.63	.5312
Within Group (Error)	466	84394.2	181.1		

*$p < .01$; **$p < .0001$

Table 20

Analysis of Variance Summary Table for School B's Reading Change Scores (N = 427)

Source	df	SS	MS	*F*	*p*
Model	5	13547.5	2709.5	13.09	<.0001**
Grade	2	2951.7	1475.9	7.13	.0009*
Treatment	1	8085.4	8085.4	39.07	<.0001**
Grade x Treatment	2	467.6	233.8	1.13	.3241
Within Group (Error)	421	87118.6	206.9		

*$p < .001$; **$p < .0001$

Pooled Reading Achievement Data

Because the results of the reading achievement data analyses were so similar between School A and School B, the data were pooled and the same analyses were re-run. The descriptive statistics of the combined school data portrayed a relatively normal distribution of scores with the groups' means ranging slightly from 44.5 to 48.5. The standard deviations, measures of skewness, and measure of kurtosis all fell within the normal ranges (see Table 21). Figure 13 is a graphic display of all of the groups' mean reading test scores at Time 1 by grade level.

Table 21

Descriptive Statistics of Combined Reading Scores: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 1 (N = 899)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	125	44.5	14.7	.22	-.35	13	82
Grade 6 without Support	144	45.6	16.9	.32	-.14	11	91
Grade 7 with Support	159	44.8	13.7	.60	.09	11	81
Grade 7 without Support	140	47.2	15.1	.30	-.37	13	86
Grade 8 with Support	173	46.0	14.9	.30	-.40	16	86
Grade 8 without Support	158	48.5	15.8	-.24	-.59	8	79

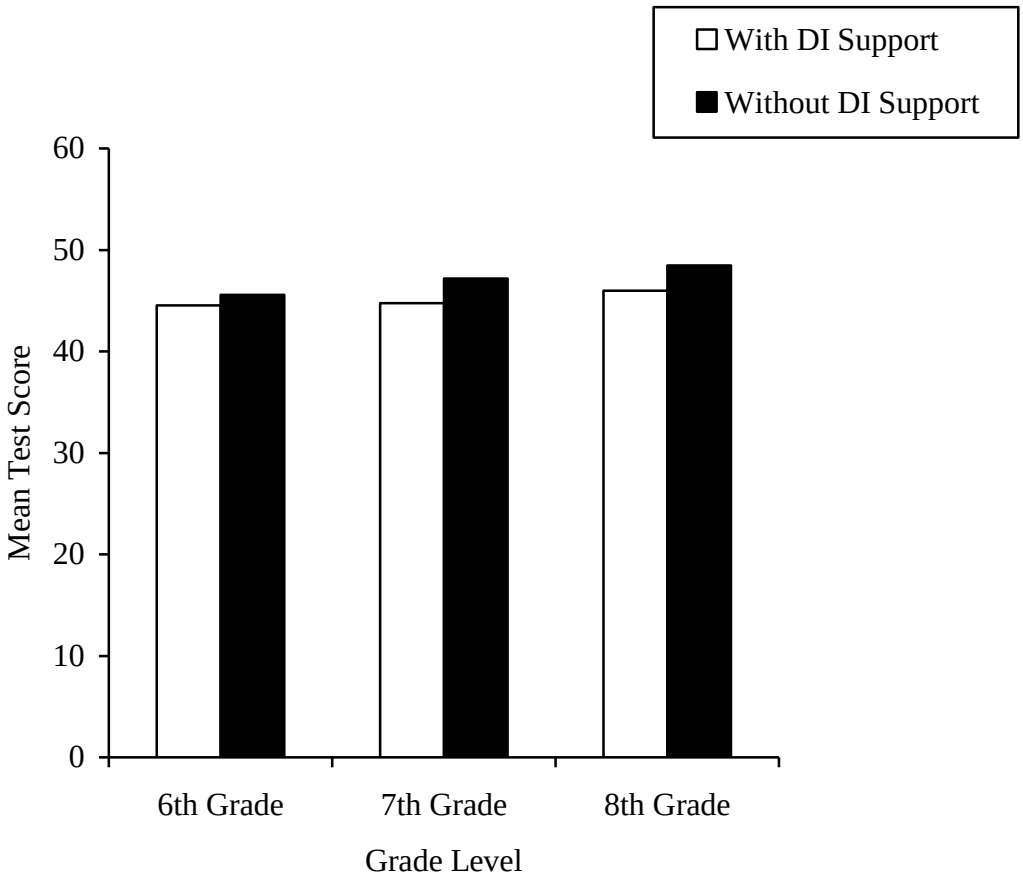


Figure 13. Mean combined schools' reading test scores by grade level at Time 1 (N = 899).

An analysis of variance procedure was then conducted on the pooled Time 1 data and the results showed that none of the comparisons were statistically significant (see Table 22). These results provide confirmation that the null would fail to be rejected and further analyses could be conducted because the groups were not statistically different at Time 1.

Table 22

Analysis of Variance Summary Table for Combined Reading Scores at Time 1 (N = 899)

Source	df	SS	MS	*F*	*P*
Model	5	1668.2	333.6	1.44	.2065
Grade	2	704.8	352.4	1.52	.2185
Treatment	1	866.5	866.5	3.76	.0532
Grade x Treatment	2	94.0	47.0	.20	.8162
Within Group (Error)	893	206548.8	231.3		

Next, the students' reading achievement scores from the two schools were pooled and compared at Time 2. The results of the descriptive statistics are presented in Table 23. The results portray a relatively normal distribution except that the sixth-grade group without support and the eighth-grade group without support had a negative kurtosis close to -1, meaning that they had slightly less outliers in their distribution. A bar graph of the groups' mean reading scores by grade level at Time 2 is illustrated in Figure 14.

Table 23

Descriptive Statistics of Combined Reading Scores: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 2 (N = 899)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 With Support	125	57.24	17.55	-.20	-.62	18	91
Grade 6 without Support	144	49.0	21.61	.50	-.94	15	96
Grade 7 With Support	159	61.4	14.19	.15	-.66	35	95
Grade 7 without Support	140	57.0	16.85	-.41	-.66	31	95
Grade 8 With Support	173	59.5	17.22	-.002	-.61	22	98
Grade 8 without Support	158	53.5	18.70	-.09	-.83	28	92

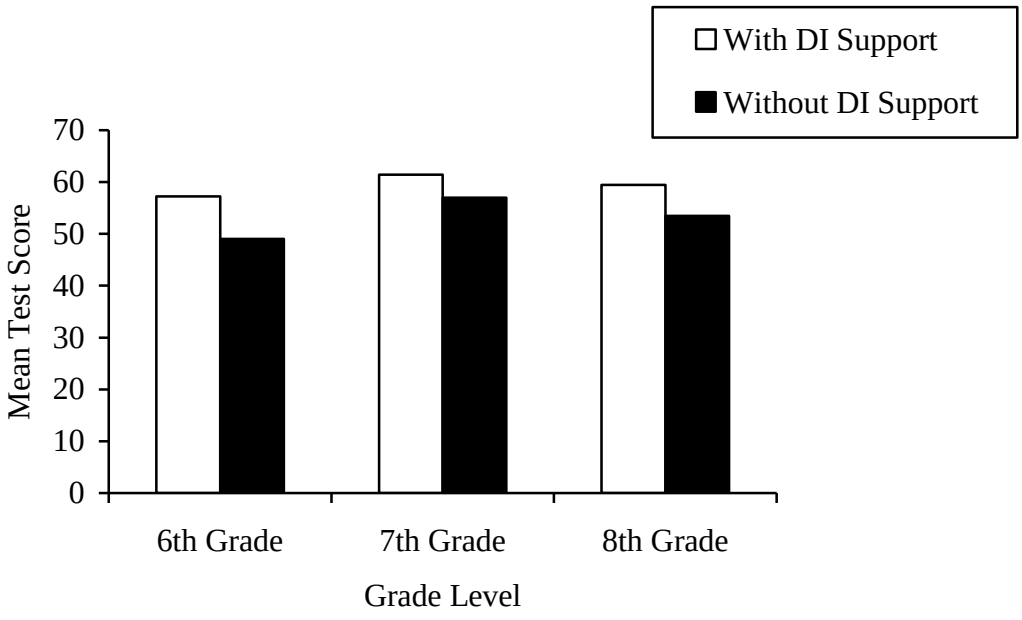


Figure 14. Mean combined schools' reading test scores by grade level at Time 2 (N = 899).

An analysis of variance conducted on the combined Time 2 reading data showed that the model, the effect by grade, and the main effect of treatment were statistically significant (see Table 24). Of particular importance to this study is the treatment effect ($F(1, 893) = 27.31$, $p < .0001$). The Cohen's (1977, 1988, 1992) effect size (*f)* for this difference was calculated to be .17, which is generally considered a small effect. The grade level effect was also small with a Cohen's *f* of .14 ($F(2, 893) = 8.25$, $p = .0003$).

Table 24

Analysis of Variance Summary Table for Combined Reading Scores at Time 2 (N =899)

Source	df	SS	MS	*F*	*p*
Model	5	15004.7	3000.9	9.5	<.0001**
Grade	2	5212.1	2606.1	8.25	.0003*
Treatment	1	8626.5	8626.6	27.31	<.0001**
Grade x Treatment	2	497.3	248.7	.79	.4555
Within Group (Error)	893	282109.3	315.9		

*$p < .001$; **$p < .0001$

Figure 15 illustrates the interaction of the groups when tracked over time from Time 1 to Time 2. In both schools, the mean reading score of the treatment participants started slightly below the control group at Time 1 and then by Time 2 their mean score was above the mean score of the control group students.

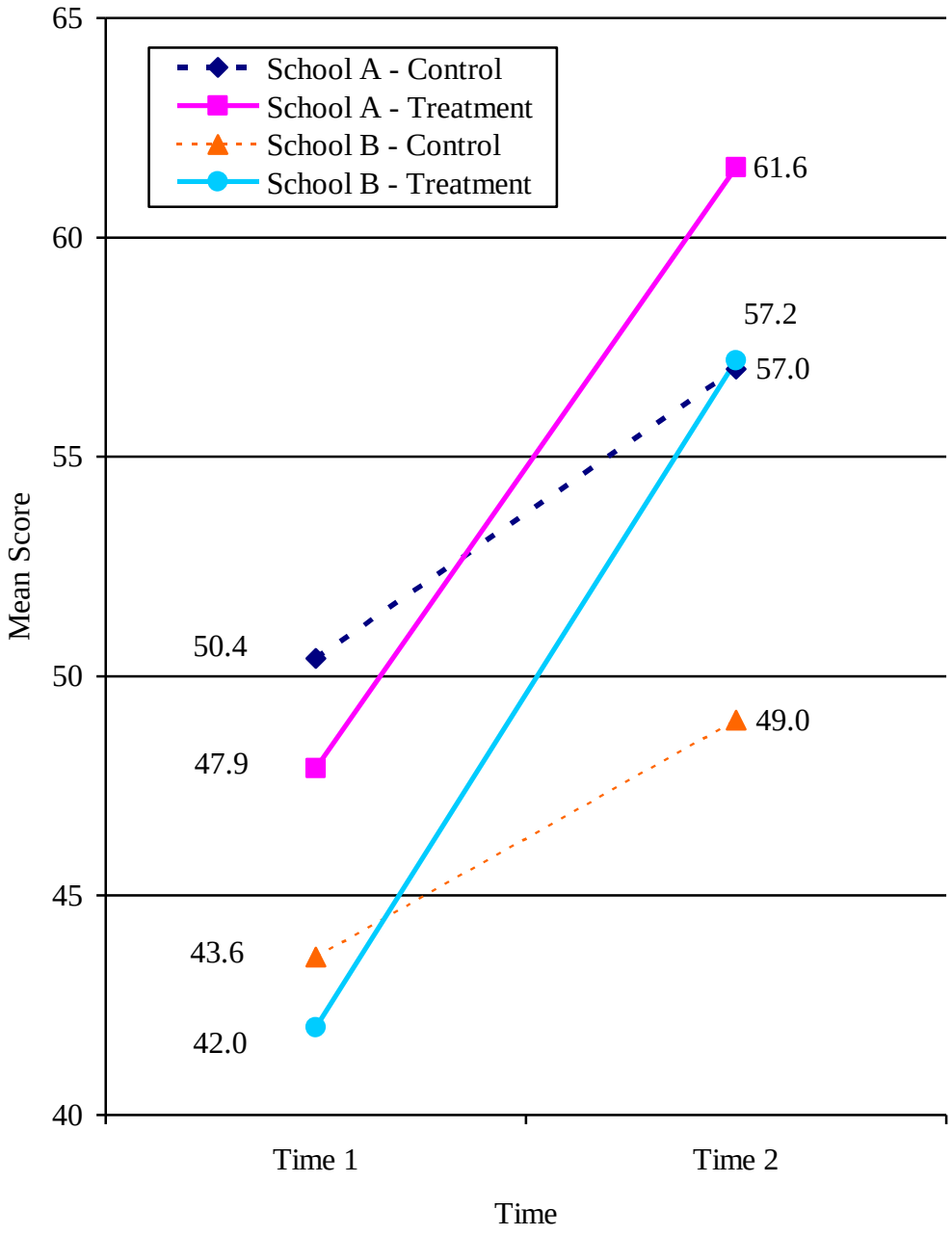


Figure 15. Line Graph of Mean Reading Scores at Time 1 and Time 2 (N = 899).

Next, the descriptive statistics of the combined schools' reading change scores by grade level and treatment were examined. The results are reported in Table 25 and reflect a relatively normal distribution of scores with respect to standard deviation, skewness, and kurtosis. Figure 16, presents a bar graph of the students' mean reading change scores by grade and treatment level.

Table 25

Descriptive Statistics of Reading Change Scores Combined Schools: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group (N = 899)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	125	12.62	14.30	.31	-.31	-14	46
Grade 6 without Support	144	3.44	14.22	.36	-.64	-20	37
Grade 7 with Support	159	16.55	12.75	-.19	-.20	-15	48
Grade 7 without Support	140	9.55	15.59	.21	-.65	-19	43
Grade 8 with Support	173	13.55	14.87	.20	-.19	-18	53
Grade 8 without Support	158	5.08	11.83	-.24	-.30	-24	37

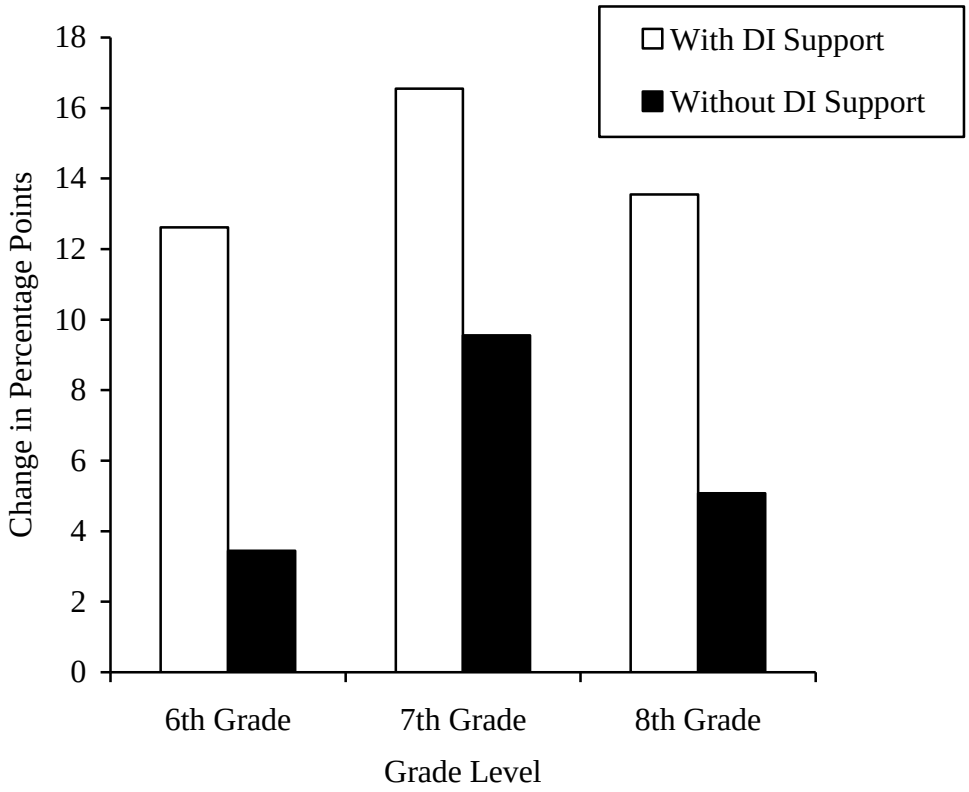


Figure 16. Mean change in combined schools' reading test scores by grade level (N = 899).

When the analysis of variance procedure was run on the combined schools reading change scores, again the model ($F(5, 893) = 20.42$, $p < .0001$), the effect by grade level ($F(2, 893) = 10.11$, $p < .0001$), and the effect of the treatment ($F(1, 893) = 77.16$, $p < .0001$), were all found to be statistically significant. The Cohen's (1977, 1988, 1992) effect size (*f*) was calculated to be .34 for the model (a medium effect), .29 for the treatment effect (a medium effect), and .15 for the grade level effect (a small effect).

Table 26

Analysis of Variance Summary Table for Combined Schools' Reading Change Scores (N = 899)

Source	df	SS	MS	*F*	*p*
Model	5	19862.82	3972.56	20.42	<.0001*
Grade	2	3933.56	1966.78	10.11	<.0001*
Treatment	1	15008.83	15008.83	77.16	<.0001*
Grade x Treatment	2	178.70	89.35	.46	.63
Within Group (Error)	893	173705.57	194.52		

*$p < .0001$

Mathematics Achievement Analyses

The following sections will describe the mathematics data that were collected during this study and the subsequent analyses. The student mathematics achievement data were submitted to the same rigorous analyses that were used on the reading achievement data.

Within site comparisons at Time 1. First, in order to assess the effect of the differentiated instruction and teacher support groups on mathematics achievement, all students were administered the FCAT Mathematics Achievement Predictor Test - Form B pretest. Next, like the reading results, the pretest results of the student mathematics achievement data at the beginning of the study, Time 1, were compared within each school site to determine if the treatment and the control groups were statistically different at the beginning of the study. The general descriptive statistics from both school sites at

Time 1 portrays a relatively normal distribution with only a few of the groups being slightly platykurtic and one group's data, seventh-grade with support at School B, being slightly positively skewed (see Tables 27 and 28). Additionally, bar graphs of the mean, grade level mathematics test scores at Schools A and B by treatment group visually illustrate the closeness of the data at Time 1 (see Figures 17 and 18). The range of group means were relatively close with School A's mean scores ranging from 52.7 to 58.1 and School B's mean scores ranging from 49.7 to 55.0. Finally, an analysis of variance on the Time 1 data within each school site demonstrated that the model was not statistically different at School A ($F(5, 463) = 1.48$, $p = .1954$) nor at School B ($F(5, 417) = 0.81$, $p = 0.5440$) (see Tables 29 and 30). Although, there was a statistical significance by grade level within School A $F(2,463) = 3.35$, $p = .0359$).

Table 27

Descriptive Statistics of Mathematics Scores for School A: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 1 (N = 469)

Group	N	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	66	56.9	22.2	-.09	-.91	11	97
Grade 6 without Support	45	55.6	17.1	.07	-.79	24	91
Grade 7 with Support	82	52.7	14.5	.44	-.46	28	87
Grade 7 without Support	84	53.0	13.5	.09	-.22	21	88
Grade 8 with Support	97	56.5	16.3	.23	-.70	22	98
Grade 8 without Support	95	58.1	17.3	-.74	-.36	19	88

Table 28

Descriptive Statistics of Mathematics Scores for School B: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 1 (N = 423)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	53	52.5	13.5	.23	-.49	29	80
Grade 6 without Support	94	55.0	18.7	.44	-1.00	28	96
Grade 7 with Support	77	49.7	15.0	1.07	.51	30	90
Grade 7 without Support	55	54.2	19.0	.01	-.94	20	92
Grade 8 with Support	79	54.4	19.6	.12	-1.02	22	92
Grade 8 without Support	65	53.3	16.6	-.12	-1.01	22	86

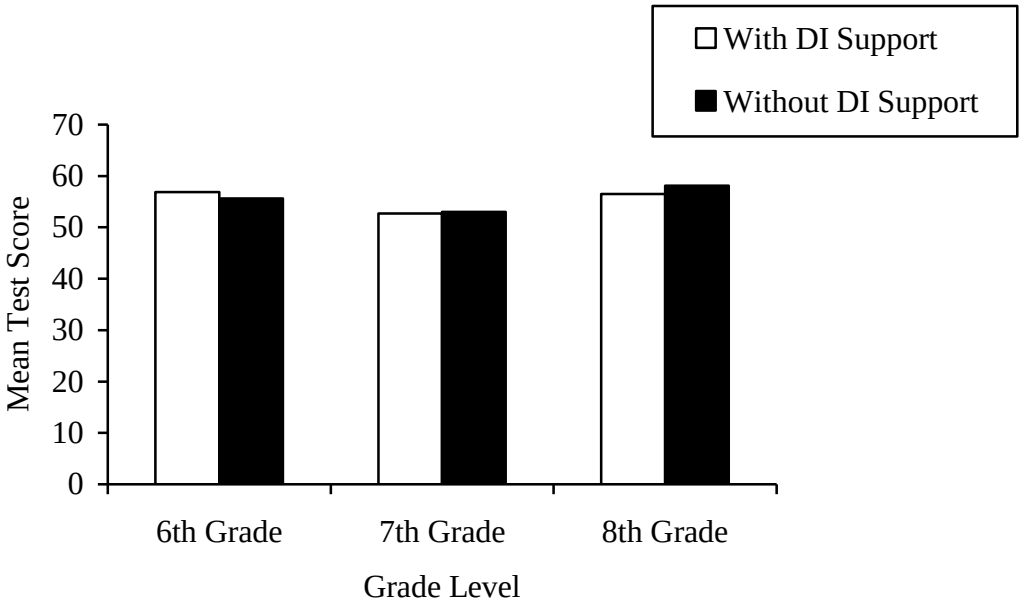


Figure 17. Mean grade level mathematics test scores at School A by treatment group at Time 1 (N = 469).

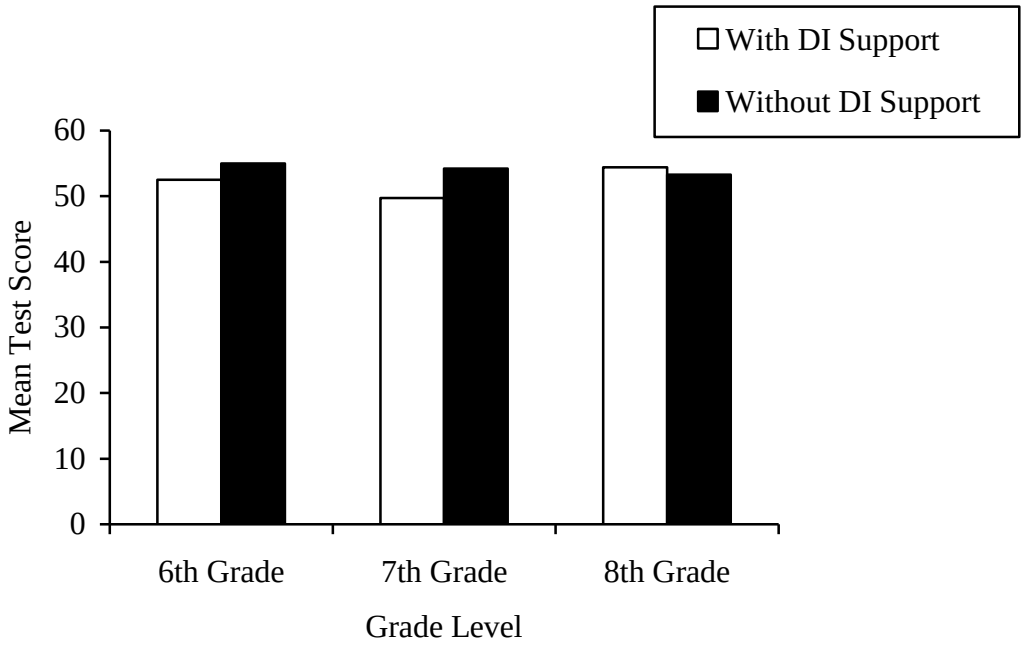


Figure 18. Mean grade level mathematics test scores at School B by treatment group at Time 1 (N = 423).

Table 29

Analysis of Variance Summary Table for School A's Within Site Comparison at Time 1 (N = 469)

Source	df	SS	MS	*F*	*p*
Model	5	2085.4	417.1	1.48	.1954
Grade	2	1891.5	945.8	3.35	.0359*
Treatment	1	5.4	5.4	.02	.8905
Grade x Treatment	2	148.3	74.2	.26	.7690
Within Group (Error)	463	130629.8	282.1		

*$p < .05$

Table 30

Analysis of Variance Summary Table for School B's Within Site Comparison at Time 1 (N = 423)

Source	df	SS	MS	*F*	*p*
Model	5	1229.3	245.9	.81	.5440
Grade	2	242.0	121.0	.40	.6720
Treatment	1	334.1	334.1	1.1	.2951
Grade x Treatment	2	484.6	242.3	.80	.4514
Within Group (Error)	417	126794.3	304.1		

Between school comparisons at Time 1. For the next level of analysis, the mathematics pretest scores were compared between Schools A and B at Time 1 by treatment level. The mean and standard deviation of the subgroups by treatment level and by school are presented in Table 31 and shows the similarity of the data from all four groups. A t-test at a significance level of .05 showed that the support groups ($t_{(455)} = 1.93$, $p = .06$) and the control groups ($t_{(440)} = .85$, $p = .39$) between school sites were not statistically different. This statistical level of difference meant that the null should fail to be rejected and that the student mathematics achievement data could be pooled between sites. Therefore, all future analyses performed on the student mathematics achievement scores will use combined school data to simplify the process and provide a larger sample.

Table 31

Descriptive Statistics of Student Mathematics Scores at Time 1 Pooled by Treatment Level Between School Sites (N = 892)

Variable	n	M	SD
Treatment Groups			
School A	245	55.3	17.6
School B	209	52.2	16.6
Control Groups			
School A	224	55.7	16.0
School B	214	54.3	18.1

Combined mathematics data at Time 1. The combined mathematics achievement data at Time 1 are presented in Table 32, which portray a similar distribution of means and standard deviations among groups. The data illustrate a relatively normal distribution for all subgroups, with the exception of a few groups that were slightly platykurtic, sixth-grade without support (-.95), eighth-grade with support (-.83), and eight-grade without support (-.77). Because these values were not less than -1.0, they were not considered a threat to the overall variation of data and allow for a continued analysis of combined data. Figure 19 visually depicts the closeness of the means among all groups.

Table 32

Descriptive Statistics of Combined School's Mathematics Scores: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 1 (N = 892)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	119	55.0	18.9	.11	-.51	11	97
Grade 6 Without Support	139	55.2	18.1	.33	-.95	24	96
Grade 7 with Support	159	51.2	14.7	.73	-.13	28	90
Grade 7 Without Support	139	53.4	15.8	.08	-.48	20	92
Grade 8 with Support	176	55.6	17.9	.13	-.83	22	98
Grade 8 Without Support	160	56.2	17.1	-.47	-.77	19	88

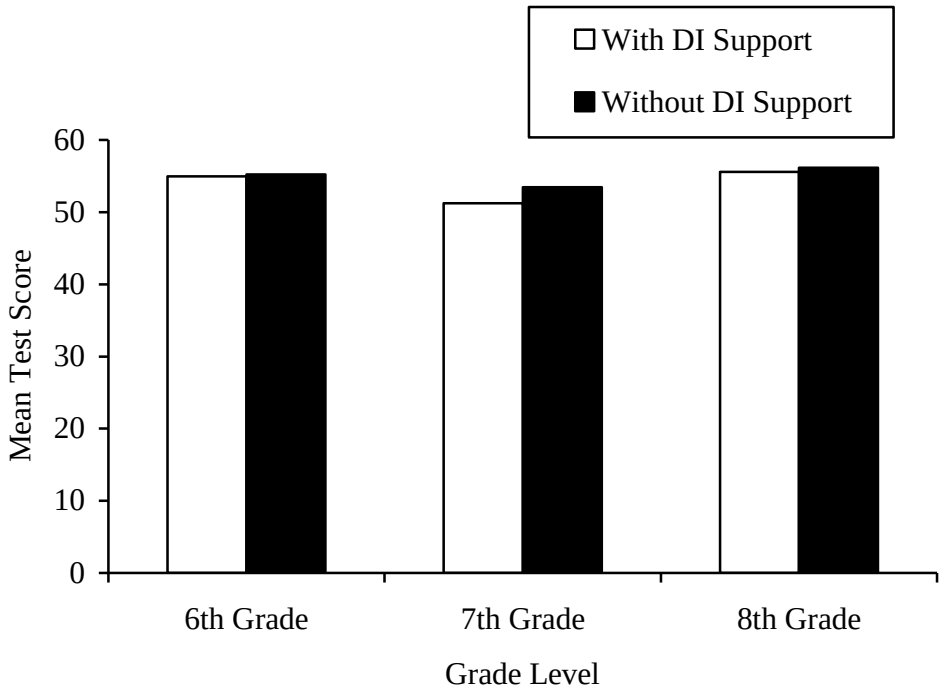


Figure 19. Mean combined schools' mathematics test scores by grade level at Time 1 (N = 892).

Time 2 analyses. At the end of the five month period, all students were administered the FCAT Mathematics Achievement Predictor Test – Form C. The student mathematics achievement posttest results at Time 2 were then compared. The general descriptive statistics from both school sites at Time 2 portray a relatively normal distribution of data for most of the groups (see Tables 33 and 34); although a few of the groups showed some minor deviation. For example, the eighth-grade group without support at School A showed a slight negative skewness (-.84) and 10 of the 12 sub-groups were slightly platykurtic which indicated less outliers and attests to the closeness of data. However, none of these conditions was significant enough to cause concern. When the data from both schools were combined, the similarity among the data sets

remained (see Table 35). Additionally, a bar graph of the combined mean, grade level mathematics test scores by treatment group visually illustrates the improvement of the treatment groups' mathematics achievement scores over those of the control groups, except in the eight-grade where the resulting mean scores are very close (see Figure 20).

Table 33

Descriptive Statistics of Mathematics Scores School A: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 2 (N = 469)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	66	59.3	20.80	.19	-1.19	23	96
Grade 6 without Support	45	56.3	21.3	.13	-1.01	15	96
Grade 7 with Support	82	69.4	13.3	-.13	-.09	33	97
Grade 7 without Support	84	63.9	17.2	-.06	-.72	26	98
Grade 8 with Support	97	68.6	15.4	-.03	-.80	39	99
Grade 8 without Support	95	68.9	18.6	-.84	-.08	18	97

Table 34

Descriptive Statistics of Mathematics Scores School B: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 2 (N = 423)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	53	59.6	14.7	.05	-1.07	33	87
Grade 6 without Support	94	50.3	20.0	.27	-.99	17	92
Grade 7 with Support	77	71.2	11.9	-.19	-.65	43	93
Grade 7 without Support	55	65.8	19.9	-.16	-1.17	32	98
Grade 8 with Support	79	64.2	18.9	-.7	-.92	26	98
Grade 8 without Support	65	62.1	16.0	-.24	-.98	29	92

Table 35

Descriptive Statistics of Combined School's Mathematics Scores: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group at Time 2 (N = 892)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	119	59.4	18.3	.15	-.98	23	96
Grade 6 without Support	139	52.2	20.6	.23	-.98	15	96
Grade 7 with Support	159	70.3	12.6	-.18	-.29	33	97
Grade 7 without Support	139	64.7	18.3	-.09	-.94	26	98
Grade 8 with Support	176	66.7	17.1	-.13	-.75	26	99
Grade 8 without Support	160	66.1	17.9	-.55	-.57	18	97

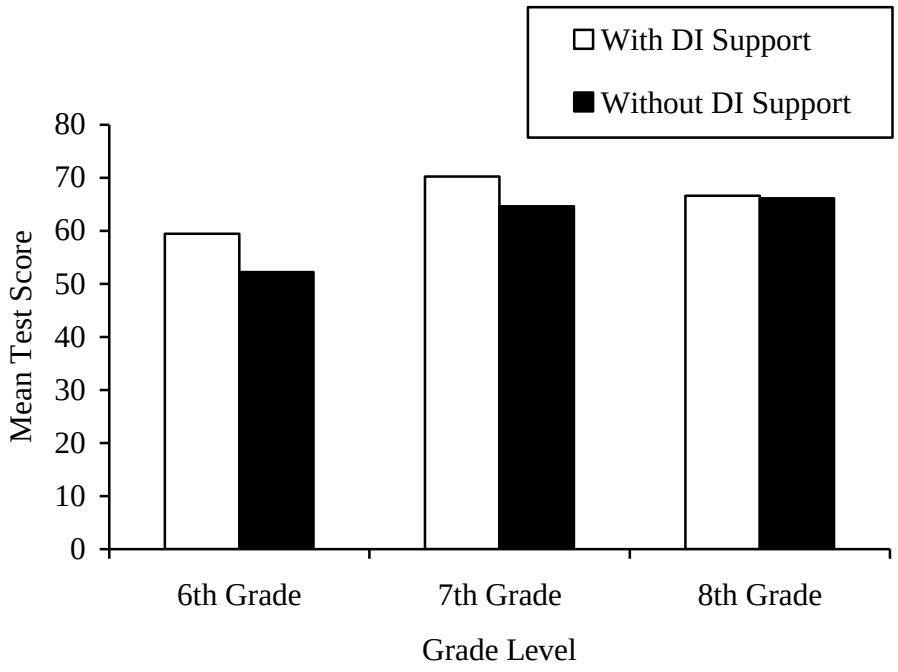


Figure 20. Mean combined schools' mathematics test scores by grade level at Time 2 (N = 892).

Change score analyses. In order to determine the resulting impact of the differentiated instruction teacher support model on student mathematics achievement, a comparison was then completed on the students' improvement from the pretest, Time 1, to the posttest, Time 2, referred to as the change score (Time 2 – Time 1 = Change Score). The descriptive statistics for the combined schools' change scores are presented in Table 36. The data illustrate a relatively normal distribution in seventh and eighth-grades. However, in sixth-grade, the change scores were low which was consistent with scores from other middle schools within the district. It is notable that the sixth-grade control groups' change score was negative, indicating that many students in that group actually scored lower on the posttest than the pretest.

Table 36

Descriptive Statistics of Mathematics Change Scores Combined Sites: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group (N = 892)

Group	n	M	SD	Skewness	Kurtosis	Minimum	Maximum
Grade 6 with Support	119	4.18	10.39	.14	-.37	-19	34
Grade 6 Without Support	139	-3.17	10.25	.49	-.25	-22	28
Grade 7 with Support	159	19.65	12.60	.27	-.56	-10	47
Grade 7 Without Support	139	11.20	10.93	-.09	.28	-18	43
Grade 8 with Support	176	11.10	9.44	.32	.28	-10	38
Grade 8 Without Support	160	9.87	9.31	-.25	.33	-19	34

When the pooled, mathematics change scores for all groups were graphically and numerically compared, the treatment groups had increased their scores to a greater degree than the control groups at all three-grade levels (see Figure 21). The seventh-grade had the largest difference of 8.45 percentage points and eighth-grade had the smallest difference of 1.23 points.

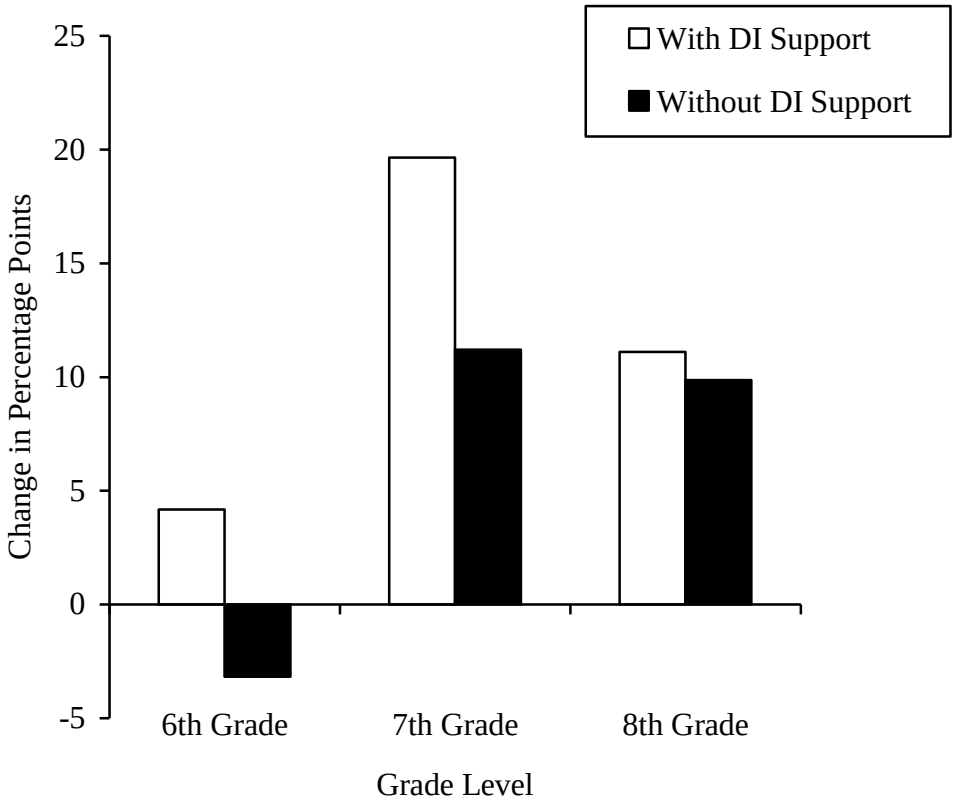


Figure 21. Mean change in combined schools' mathematics test scores by grade level (N = 892).

Figure 22 illustrates the interaction of the groups when tracked over time from Time 1 to Time 2. Within both schools, the mean mathematics score of the treatment participants started slightly below the control group at Time 1 and then by Time 2 their mean score was above the mean score of the control group students. In fact, School B had a significant difference of 7.7 points at Time 2.

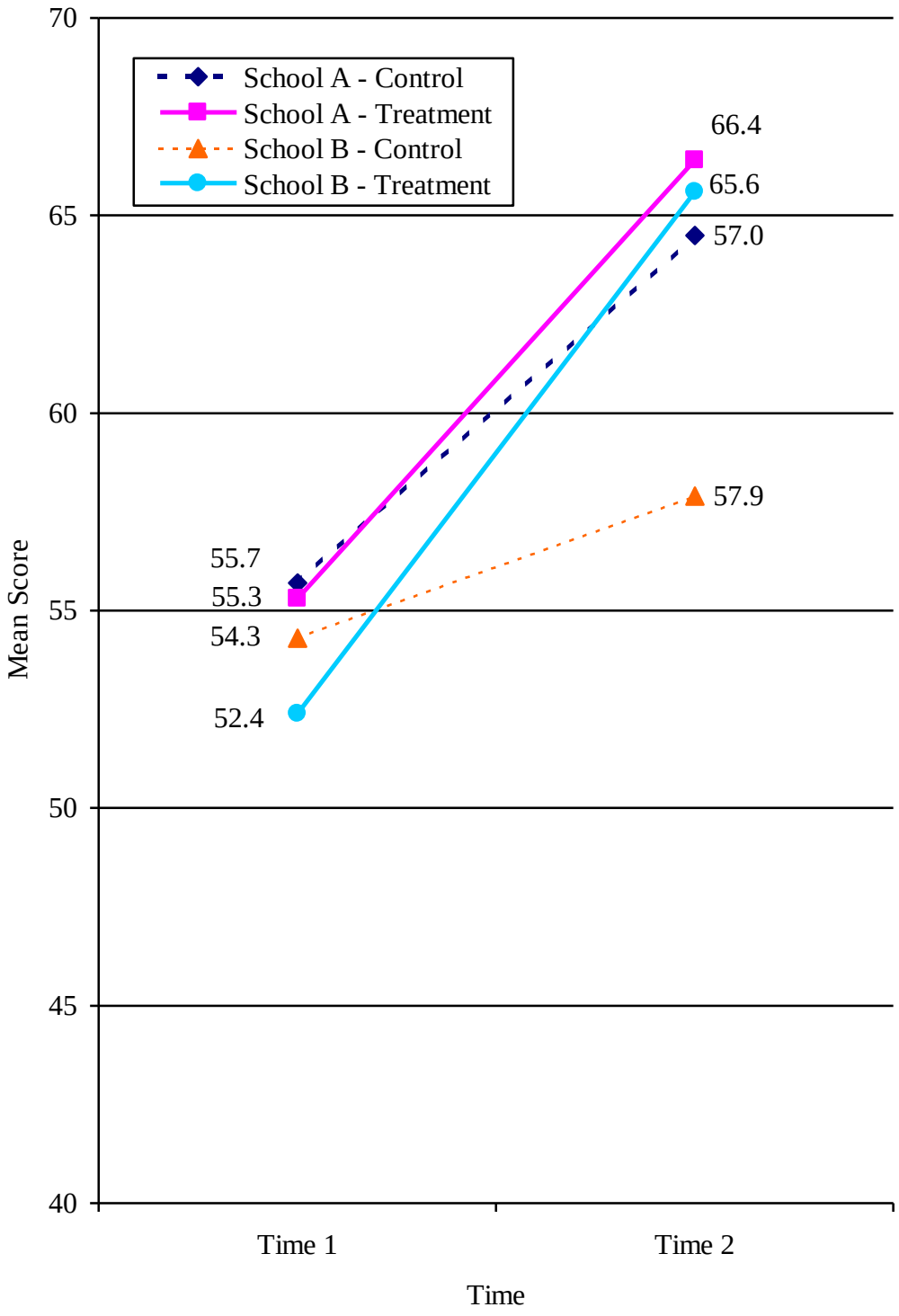


Figure 22. Line Graph of Mean Mathematics Scores at Time 1 and Time 2 (N = 892).

In order to rule out chance, the final analysis on the student mathematics achievement data was to conduct a 2 (treatment) X 3 (grade level) factorial ANOVA on the resulting change scores using an alpha level of .05 to test for each effect. Before proceeding with the analysis, the assumptions of independence, normality, and homogeneity of variances were investigated. First, because the students worked individually on their assessments and trained teachers proctored the assessment, it is therefore reasonable to assume that the assumption of independence has not been violated. Although the sample sizes are not exactly equal, they are relatively similar both within each grade level and among groups. The total number of student participants in the mathematics treatment group was 454 and the total in the control group was 438 making the sample sizes large enough to expect robustness to violations of the normality assumption, therefore the normality assumption does not appear to be violated. Finally, when considering the homogeneity of variances, the largest variance ratio was 1.48, less than 2.0, which means the equal variance assumption does not appear to be violated. Therefore, we would expect the ANOVA to be relatively robust to violations of the homogeneity of variance assumption. Based on this analysis of the assumptions, it was reasonable to proceed with the factorial ANOVA.

The results of the two-way factorial ANOVA are presented in Table 37. The overall model was statistically significant ($F(5,886) = 77.36$, $p < .0001$). Cohen's (1977, 1988, 1992) effect size *(f)* was calculated to be .66, which is generally considered a large effect. Due to grade level differences, the grade level effect of the use of teacher support groups was also found to be statistically significant using the Type III Sum of Squares data due to the unequal group sizes, ($F(2,886) = 141.72$, $p <.0001$). Cohen's (1977, 1988,

1992) *f* was calculated to be .56, which is again generally considered a large effect size. Of particular significance to this research project, the treatment effect was also statistically significant ($F(1,886) = 63.8$, $p < .0001$). The Cohen's (1977, 1988, 1992) effect size *(f)* was again calculated and found to be .38, which is generally considered a large effect. The interaction of the effect of both grade level and treatment level for the combined groups was significant ($F(1, 886) = 10.8$, $p < .0001$); although the Cohen's (1977, 1988, 1992) effect size (*f*) was .16, a small effect.

Table 37

Analysis of Variance Summary Table for Combined Schools' Mathematics Change Scores (N = 892)

Source	df	SS	MS	*F*	*p*
Model	5	42880.2	8576.0	77.36	<.0001*
Grade	2	31422.14	15711.07	141.72	<.0001*
Treatment	1	7072.13	7072.13	63.80	<.0001*
Grade x Treatment	2	2393.84	1196.92	10.80	<.0001*
Within Group (Error)	886	98219.06	110.86		

*$p < .0001$

The mathematics results were very encouraging. These data suggest that the use of differentiated instruction teacher support groups in urban, Title I schools was statistically significant with regard to improving student mathematics achievement scores, even though there was a grade level effect.

Teacher Fidelity Observation Data Analyses

In the following section, the summarized data and results of all teacher observation analyses are presented and evaluated in order to address the second research question:

> What were the statistical differences among teacher groups who participated in facilitated support groups and those who did not with respect to their implementation of differentiated instruction as measured by the Differentiated Instruction: Fidelity Implementation Tool (DI: FIT) observation tool?

First, all teachers were observed by trained observers to determine their degree of fidelity to the differentiated instruction model using the DI: FIT observation tool. Specifically, in order to confirm that the teachers in the treatment groups were applying the differentiated instruction philosophy to their classrooms and transferring the strategies learned in the differentiated instruction, facilitated teacher support groups, these teachers were observed once per nine-weeks and given specific feedback on ways to improve their instruction. The teachers in the control group were also observed during each nine-week period to assess whether or not they were differentiating their classroom instruction and, if they were, to what degree.

The basic descriptive statistics analysis of the teacher participants' mean DI: FIT observation scores revealed that all four groups had a similar sample size, standard deviation, and minimum scores (see Table 38). The groups were different with respect to the mean scores (see Figure 23). As expected, the participants who were a part of the differentiated instruction monthly support groups scored an average of 3.4 points higher on the 20-point scale. In addition, the treatment group at School A had a few more high

scores (skewness = -1.41, kurtosis = 2.10). The control group of participants at School B had a slightly more flat distribution of scores (kurtosis = -1.28). It is important to note that the differences of the combined scores by school site are very similar with the mean scores differing by only .68 points. These data again provide evidence that the schools were similarly matched.

Table 38

Descriptive Statistics of Teacher DI: FIT Observation Mean Scores: Mean, Standard Deviation, Skewness, Kurtosis, Minimum, and Maximum by Group and by School (N = 55)

Group	N	M	SD	Skewness	Kurtosis	Minimum	Maximum
Treatment Group at School A	15	17.5	2.29	-1.41	2.10	11.5	20.0
Control Group at School A	13	13.7	2.19	.26	-.87	10.5	17.5
Treatment Group at School B	13	16.5	2.11	-.63	-.51	12.5	19.5
Control Group at School B	14	13.7	2.04	-.35	-1.28	10.5	16.5
Combined Treatment Groups	28	17.1	2.22	-1.19	1.02	11.5	20.0
Combined Control Groups	27	13.7	2.07	-.29	-.81	10.5	17.5
School A: Treatment + Control	28	15.75	2.92	-.34	-1.05	10.5	20.0
School B: Treatment + Control	27	15.07	2.49	-.44	-.58	10.5	19.5

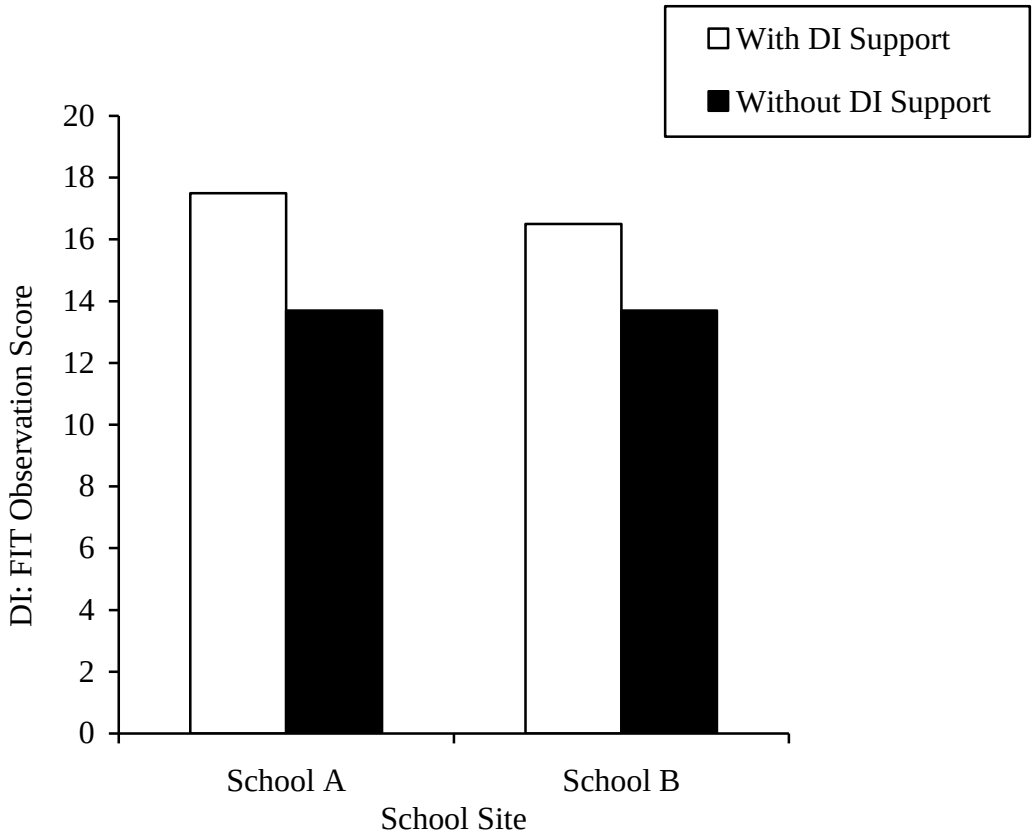


Figure 23. Mean Teacher DI: FIT Observation Scores by Treatment and by School (N = 55)

Relationship Between Fidelity and Achievement

In the following section, data for the analysis of the correlation between the teachers' DI: FIT observation scores (fidelity) and students' mean academic change scores are presented in order to address the third research question:

> What was the relationship between the teachers' differentiated instruction implementation scores as measured by the DI: FIT and the students' achievement change scores?

An analysis of variance on the same observations scores revealed the overall model was statistically significant ($F(3, 51) = 11.34$, $p < .0001$) (see Table 39). The Cohen's (1977, 1988, 1992) *f* effect size was calculated to be .78, which is considered a large effect size. The R^2 value reported by SAS was .40, which reflects the percentage of variance that was accounted for by the dependent variable. The use of the differentiated instruction teacher support groups, the treatment, was also statistically significant ($F(1, 51) = 31.86$, $p < .0001$), with a large Cohen's (1977, 1988, 1992) *f* effect size of .76. The effect of the school site and the interaction were not statistically significant. These data allowed for the secondary analyses of the correlation between the teachers' DI: FIT observation scores and students' mean academic change scores.

Table 39

Analysis of Variance Summary Table for Teacher DI: FIT Observation Scores by Treatment Level and by School (N=55)

Source	df	SS	MS	*F*	*p*
Model	3	159.0	53.0	11.34	<.0001*
Treatment	1	148.9	148.9	31.86	<.0001*
School	1	3.28	3.28	.70	.4064
Treatment x School	1	3.06	3.06	.65	.4223
Within Group (Error)	51	238.4	4.67		

*$p < .0001$

The analysis among the treatment, teachers' mean DI: FIT observation score, students' mean academic change score, and school variables shows that there was a very strong positive correlation between several of the variables (see Table 40). The correlation statistic between the teachers' DI: FIT observation score and the treatment was calculated to be .62 (p <.0001), a moderate correlation. The correlation between students' mean academic change score and the treatment was also a moderate correlation of .63. However, the correlation statistic between the teachers' mean DI: FIT observation score and the students' mean academic change score was calculated to be .79. From this statistic, R^2 was calculated and found to be .62, which is considered a moderate correlation and accounts for 62% of the variance. It is notable that again the correlation of the school site was not statistically significant and therefore had no effect on the other treatment correlations.

Table 40

Intercorrelations Between Teachers' DI: FIT Observation Score, Students' Mean Academic Change Score, Treatment, and School (N = 55)

Variable	Teachers' Mean DI: FIT Observation Score	Students' Mean Academic Change Score	School
Treatment	.62 (<.0001)*	.63 (<.0001)*	-.05 (.69)
Teachers' Mean DI: FIT Observation Score		.79 (<.0001)*	-.13 (.36)
Students' Mean Academic Change Score			.01 (.95)

Note. The values in the parentheses are *p*-values. *p < .0001

Support Group Analyses

In the following section, a qualitative analysis on the detailed minutes from the facilitated teacher support sessions, the teachers' implementation journals, the facilitator's reflective journal, and the feedback obtained at the final focus group session were examined in order to address the fourth research question:

> Using qualitative data and feedback provided by the teachers in the treatment groups, what were the teachers' perceptions of the facilitated support group model and their instructional growth?

In order to ensure that any instructional philosophy and teaching strategies learned are transferred to the classroom with fidelity, staff development personnel must make certain that the teachers have on-going support (Westling, Cooper-Duffy, Prohn, Ray, & Herzog, 2005). In this study a facilitated, teacher support group model was initiated to help provide the necessary support teachers would need to change their teaching practices and sustain the newly implemented skills and strategies with fidelity. The data from the DI: FIT teacher observations documented that the teachers who participated in the groups did, in fact, implement the components of differentiated instruction to a greater degree than the teachers in the control group.

Now, a closer analysis of the data gleaned from these meetings will help guide future meetings and to document the components that the teachers perceived as important. Using an inductive analysis in conjunction with a document review, all data were reviewed, meaningful units were identified, units of data were coded, and then the data were categorized in order to identify basic themes and views. Then, the data were further reduced through a constant comparison, a consolidation of any redundant

categories, and an analysis of emergent themes. The use of multiple methods of data collection allowed for method and data triangulation and increased credibility of the findings.

All of the teachers responded that they thought the support group model was beneficial and had positively impacted their teaching practices, their students' achievement, and the classroom learning community. One teacher even wrote in her journal, "I am so very grateful that I was invited to be part of the DI Support Team. It has made a huge impact on my teaching and my students." An overwhelming percentage of the 28 teacher in the treatment group, 96.4%, responded that they really liked the overall format of the facilitated, teacher support groups. Further, based on the results from the district in-service follow-up questions (see Appendix K and L), when teacher participants were asked if the content of the differentiated support group meetings was appropriate and built upon the knowledge and experiences of the intended participants, 90.5% of the teachers responded that they strongly agreed and 9.5% responded that they agreed; no participants responded that they were undecided, disagreed, or strongly disagreed.

During the final focus group session, 100% of the teachers responded that the length of each session, 2 hours, was adequate. However, 89.3% of the teachers requested that the groups meet more often during the beginning of the year. Specifically, 71.4% requested that the sessions begin during pre-planning so the teachers could begin to create differentiated lesson plans and create their flexible grouping student cards with available information, like students' names and the previous year's FCAT achievement scores. Then after the year began, 89.3% expressed an interest in meeting twice a month for the first three months.

On-going support of the teachers is important before, during, and after support group meetings. In this study, this was accomplished in several ways. First, the use of the differentiated instruction, teacher resource library was appreciated by all of the support group, teacher participants, and 92.9% of the teachers agreed that they found ideas they could implement and/or encouraged the use of new differentiated instruction components into their lesson plans. One teacher wrote, "The books helped me to research more ideas to make my classes better." The school district was also supportive and provided all of the members of the support group with access to a DI conference area on their email desktop. This provided the teachers with an opportunity to share lesson plans and internet links, ask questions, post concerns, receive support, and view the ideas and growth of others. In addition, it served as a daily reminder for teachers to strive toward increased implementation of the differentiated instruction principles. One hundred percent of the participants responded that they appreciated the DI conference area and expressed a desire for it to be continued into the next year.

Several improvements were suggested by the teachers during the final focus group to improve the process for the upcoming year. All of the teachers, 100%, supported the idea to make the support groups open to all faculty members. The majority of the teachers, 92.9%, wrote on their *DI Support Group Feedback* sheets that they would like to be able to observe other teachers' differentiate instruction lessons. As a group, they came up with several suggestions to accomplish this. First, they suggested having teachers volunteer to be "demonstration classrooms" so whenever they were doing a differentiated lesson they were particularly proud of, they could invite the rest of the faculty to "visit" during their conference period. They also suggested developing peer

partner relationships so they would have the input of other classroom teachers and not just the trained observers.

A theme that was expressed through the journals, the feedback sheets, and the focus groups was that the teachers genuinely enjoyed the sheer process of sharing ideas and collaborating about teaching through a professional forum. One teacher said, "It was nice to bounce ideas off each other and help others to work through their problems." Another also shared, "The discussion was very helpful; it gave me ideas and inspiration for my classroom. I felt like my students were learning more." A different teacher wrote, "The open forum lent itself for honest sharing and for offering of professional support and encouragement among peers. Teachers emerged as leaders among peers." A fourth teacher wrote, "The sharing process helped to confirm what I was doing." Several participants also referenced the benefit of the vertical and horizontal articulation among subject area teachers and across disciplines as being a benefit. They stated that they just do not have time during the day to sit down and talk with other teachers about what they are doing in their classroom.

When asked which elements of the study encouraged them to sustain their instructional fidelity throughout the study, there was a variety of responses. The majority of teachers referenced the support group meetings, the fact that they knew an observer would be coming, and they did not want to disappoint the group. One teacher wrote, "The meetings kept my instructional focus and renewed my zeal to learn and implement new DI lessons. With other district in-services, I lost interest and went back to the teaching strategies I used previously." Another shared, "The support group holds me more accountable and reminds me of its [differentiated instruction] importance." A fourth

teacher wrote, “The support group meetings helped me keep in focus the role DI can play in better instructing and evaluating students.”

It is important to remember that the 50% of the teachers in the treatment group were new teachers (0-3 years of teaching experience) and 14.3% were not certified in the subject they were teaching, which is very representative of the teachers who commonly teach at Urban, Title I schools. In one of the teacher’s journals, he wrote, “I did not know very much about differentiating instruction at the beginning, but I feel more confident now.” Another new teacher shared in the focus group meeting that, “It [the support group meeting] allows me to ‘talk the talk’ with others in the group and in talking with them – the creative juices continue to flow.” A third teacher wrote in their journal, “In the beginning of the year, I was leery of DI, but I really feel now that my students can really get behind their work when they feel it is geared specifically for them.”

The participants’ responses were very encouraging and provided feedback on ways the process worked well and ways it can be improved. For the future, the support group model holds promise as an important tool districts and in-service personnel can use to help sustain implementation with fidelity of any instructional philosophy or program.

Summary

In this chapter, numerous data sources have been provided to assure that the treatment and control groups were as closely matched as possible. Next, data on teacher fidelity and student achievement was presented so the impact of the differentiated instruction support group model could be substantiated. Finally, the support group model itself has been closely scrutinized. Of importance is the fact that although there was a school level effect with respect to the reading achievement and a grade level effect with

respect to mathematics, students, as a whole, whose teachers were participants in the differentiated instruction support groups out performed their peers. Even though the difference was slight, in most cases, the change in academic achievement was statistically significant and a few points difference is all that many students need to meet state and district level benchmarks and standards.

Chapter Five

Discussion

Purpose

From the onset, this multifaceted study had a myriad of purposes. First, it was the desire of the researcher to contribute to the research and knowledge base of differentiated instruction, facilitated support groups, and implementation fidelity of evidence-based practices. Further, this study utilized and collected additional data on a new fidelity assessment tool for teachers, the DI: FIT, and evaluated teachers' perceptions of the facilitated support group model. Finally, the study examined the relationship between implementation of differentiated instruction with fidelity and the academic outcomes of urban, middle school adolescents.

Method

This study utilized a quasi-experimental, mixed-method design to investigate multiple components over a five-month period. For this study, two matched urban, Title I, middle schools were selected as implementation sites; and within each of the sites, teams of teachers who taught the same body of students were purposefully assigned to the treatment or control group at each grade level. The demographics of both students and teachers were compared to assure that all comparisons were similar samples. The combined sample contained 55 teachers and 906 students. A triangulation of data from

the facilitated support group meetings, the teachers' individual journals, and the fidelity observations were utilized to analyze the teachers' fidelity with respect to differentiated instruction, correlation with changes in student achievement, and teachers' perceptions of the support group model and instructional change.

All teacher participants in the treatment group attended a two-hour, monthly, differentiated instruction support group meeting, had access to a differentiated instruction resource library, and received feedback and instructional suggestions following all observations, i.e., reinforcement on which differentiated components were observed, possible flexible grouping strategies, and ideas for future lessons. The final session of the year was a focus group so the researcher could obtain feedback from the teachers regarding which components of the support group they felt were the most valuable, which components need to be continued, and what components needed to be added.

Reading achievement was assessed using the Stanford Diagnostic Reading Test – Fourth Edition and mathematics achievement was assessed using the district's FCAT Mathematics Predictor Tests. All student achievement results were then analyzed using basic descriptive statistical procedures in addition to an analysis of variance procedure. An effect size was also calculated on all statistically significant findings. All teacher participants were exposed to the differentiated instruction model and then observed each nine-week period to assess the degree of implementation fidelity. After all of the observations had been completed, the mean teacher fidelity observation scores were statistically compared using an analysis of variance by treatment and school level. Next, a correlation between the teachers' degree of fidelity with respect to the observed differentiated instruction components and the students' mean academic change score was

used to determine the relationship between fidelity of implementation and student achievement.

Results

The results of this study are very encouraging. Both the reading and mathematics achievement change scores and the difference in the teacher fidelity observation scores, DI: FIT, by treatment group were statistically significant. A clear relationship also existed between the teachers' mean implementation fidelity scores and the student achievement scores. In addition, the teachers who participated in the support group meetings clearly felt that the support was beneficial.

The overall mean difference in students' reading achievement change scores between the treatment and control groups was 8.38 percentage points, which was statistically significant ($F(1, 893) = 77.16$, $p < .0001$, $N = 899$). The Cohen's (1977, 1988, 1992) *f* effect size for the treatment comparison was calculated to have a medium effect size of .29. A Cohen's *f* of .1 is considered small, .25 medium, and .4 large (Cohen 1977, 1988, 1992.) The reading achievement results were similar across all three grade levels and the difference between schools was not statistically significant.

The overall mean difference in the students' mathematics achievement change scores was 6.12 percentage points. This comparison by treatment level was also statistically significant ($F(1, 886) = 63.80$, $p < .0001$, $N = 892$) and had a Cohen's (1977, 1988, 1992) *f* effect size of .38, a medium to large effect. With reference to the mathematical achievement data, grade level differences were found. The grade level effect was statistically significant ($F(2,886) = 141.72$, $p < .0001$, Cohen's $f = .56$) as was the interaction effect of grade level by treatment ($F(1,886) = 10.8$, $p < .0001$, Cohen's $f =$

.16). The overall mathematics achievement data were effected by the low performance of the sixth-grade students. Sixth-grade is usually a difficult transition year for students and the fact that three of the four sixth-grade mathematics teachers were non-tenured, explains some of the possible reasons for the lack of mathematics growth in this grade. The fact that the majority of the sixth grade teachers were not tenured is a common scenario in middle schools. This is often a result of the more senior teachers selecting to teach the older, more mature students leaving the administration with the task of having to fill the sixth-grade positions with new, inexperienced teachers. Although, the sixth-grade students experienced limited growth in this study, their change scores were consistent with the mean change scores of other sixth-graders in the school district during the same period. In the final data collection, there was also little difference between the treatment and control group's mathematics achievement data in the eighth-grade. This lack of difference in achievement results is possibly due to the high differentiated instruction implementation scores of the mathematics teachers as evidenced by their DI: FIT observations. Unlike other general education teachers in the control group, the eighth-grade mathematics teachers implemented many of the differentiated instruction components without the help of the support group. Their high degree of implementation on their own is testimony to the fact that some teachers will implement evidence-based strategies on their own; however, in this study they represented only 2 out of 27 teachers, 7.4%. Even with the closeness of data in eighth-grade, the overall treatment effect was still statistically significant. In urban schools, where large percentages of students are academically below grade level, these small gains can help "close the gap" in

performance and provide some students with the few points they may need to pass their benchmarks in order to be promoted to the next grade.

The teacher DI: FIT observation data also supported the conclusion that a statistically significant difference existed between the teachers who participated in the differentiated instruction support groups and the teachers who did not. While the overall mean difference by school was only .68 points, the overall mean difference by treatment level was 3.4 points out of a possible 20 points. The implications are that the average teacher who participated in the support group utilized approximately 3 to 4 more differentiated instruction components in their classroom than the teachers who did not participate. The indicators that were often omitted by the control group were: #4 Lesson is differentiated by content, product, or process; #5 Lesson is differentiated according to students' readiness, interests, or learning profiles; #9 Teacher uses anchor activities; #10 Teacher acts as a facilitator; #13 Teacher uses flexible and purposeful grouping; and #20 Teacher and students collaborate in the learning process.

For research replication purposes, both of the co-observers were easily trained within a few hours. After a discussion of fidelity indicators and a practice coding session, a high degree of inter-rater reliability was obtained. In the future, more teachers and administrators will need to be trained on the correct use of the DI: FIT instrument.

The analysis of variance on the teacher mean observation scores by treatment level revealed a statistically significant relationship ($F(1, 51) = 31.86$, $p < .0001$), with a Cohen's (1977, 1988, 1992) *f* effect size of .76, which is considered to be a large effect. Additionally, when the teachers' mean observation scores were correlated to the academic change scores of their students, the resulting correlation value of .79 suggested

that there was a moderate correlation between these two variables. This moderate correlation value accounts for 62% of the variance in the model ($R^2 = .62$).

When teachers in the support group were provided an opportunity to share their perceptions of the differentiated instruction support group model, they provided some keen insight to common implementation difficulties and suggestions for future support group models. Further, many of the teachers shared that if they had not been a part of the support group model, they would have abandoned the differentiated instruction philosophy early in the treatment period and opted for a more traditional approach to teaching because of the amount of time that it took to create project and lesson options and to change their pedagogy. All 28 of the teachers in the treatment group felt that the coaching and teacher-to-teacher sharing aspects of the support group model were key components in helping them maintain their instructional fidelity. The teacher participants also cited group accountability as a big motivator.

Based on previous research, initial implementation, fidelity, and sustainability are common problem areas where district and school implementation projects experience their biggest challenges (Webster-Stratton, 2003). Moreover, Joyce and Showers (2002), two notable researchers in the field of professional development, have conducted and reviewed hundreds of studies on this topic and they caution purveyors that the most effective intervention will not produce desired effects if it is not implemented with fidelity. Their well-known meta-analysis on training and coaching teachers indicated that the key components to implementation fidelity by practitioners were practice and feedback in training and on-going coaching in the classroom (Joyce & Showers, 2002).

Limitations

The results of this study are only applicable to urban, Title I, middle school teachers and students. The results are also limited to the tests that were utilized in this study. Further replications will need to be conducted with additional achievement tests to see if the results can be generalized. To compensate for some of the possible external threats to validity, a treatment and control group was selected within each school. All effects reported in this study have limited generalizability due to the specific demographics of the population studied and the sample of participants utilized. Although special care was used to ensure that student and teacher participant groups were closely matched and the study had a large number of participants, it was still a convenience sample; therefore the results must be viewed with caution.

Significance

In spite of these limitations, this study provided data on the use of a viable model that enabled two urban schools to implement evidence-based practices successfully. The need for such a model is well documented in the research (Brandt, 1998; Greenwood & Abbott, 2001; Greenberg, Weissberg, O'Brien, Zins, Fredericks, Resnik, et al., 2003; Joyce & Showers, 2002; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Little, 1993; Richardson, 1997; Showers, Joyce & Bennett, 1987; Spencer & Logan, 2003; U.S. Department of Education, NCLB, 2002). Even though there has been a wealth of research on effective teaching components, there continues to be a gap between research and practice in the field of education. In order to implement evidence-based practices in schools, policy makers, trainers, coaches, and practitioners need a clear model that will help teachers implement and sustain their instructional fidelity. The teacher support group

model utilized in this study demonstrates the potential of this type of model in urban, Title I middle schools. This type of classroom environment holds many challenges for educators, because many of the children have low scores and come from impoverish backgrounds. Since the model utilized in this study did provide statistically significant academic improvements in both reading and mathematics, it may possibly work for other populations of students in other regions.

The current study also provides additional data on effective professional development and instructional implementation practices that are sorely needed in the field of education. The widely cited research of Guskey (1986), Joyce and Showers (2002), and the U.S. Department of Education's National Center for Education Statistics (2000) document the limited transferability of instructional knowledge learned through in-service opportunities to the classroom. Each year, school districts spend a great deal of money paying for trainers, supplies, and participant salaries and then very little of the knowledge and/or skills become part of the teachers' instructional repertoire (Joyce & Showers, 2002). Based on the DI: FIT teacher observation data, the majority of the teachers who participated in the on-going support group during the five-month period were able to implement many of the differentiated instructional strategies with fidelity. They were also able to sustain their instructional enthusiasm through the support of their colleagues at the monthly teacher support group meetings. The data obtained from the many facets of this study's model provides support for some possible implementation solutions that are missing in many professional development programs. Unquestionably, districts need to find a more cost-efficient way to bridge the research to implementation gap; and although the teacher sample size in this study was small, the student data

supported the effectiveness of the teacher support model. The model utilized could easily be adjusted and the lessons learned can provide future guidelines for other districts wishing to expand their professional development programs.

Furthermore, this study has demonstrated that when administrators, trainers, coachers, and teachers utilize a support group model in addition to fidelity observations with feedback, student achievement can be affected. Currently, many educators find themselves struggling to teach their increasingly diverse classes. The differentiated instruction philosophy combines many evidence-based practices for teaching within one educational philosophy, which provides opportunities for increased social interaction, appropriate learning strategies, helpful feedback, and a positive learning environment (Brandt, 1998; Chapman & King, 2005; Hornsby & Diket, 1999; Tomlinson, 2005). Teachers can no longer afford not to differentiate their instruction and administrators can no longer afford to take a passive role in supporting their teachers' endeavors. The differentiated instruction model's full potential can only be actualized through school personnel's collaborative efforts, and thus, *all* students will be afforded opportunities to learn.

Implications for Practice

The two schools utilized for this study were purposefully selected because they characterize urban, Title I schools that typically have increased challenges due to large percentages of students who are functioning academically below grade level, increased discipline problems, high percentages of students from impoverished backgrounds, high teacher turn-over, very diverse student populations, and high percentages of teachers who are either not certified or lack classroom experience. In the current study, the

differentiated instruction, support group model was a useful vehicle for raising the reading and mathematics achievement scores of students at both school sites and raising the teachers' implementation fidelity to differentiated instruction. The overall program was very cost-efficient. After the initial costs of the district in-service and the minimal costs for a reference library, the only continual cost was teacher time, which can be done either after school or during school planning times with the support of administration. If all of the components of the current model were implemented, it would require a person at each site assume the role of facilitator or organizer, the commitment of the participants, the training of participants and observers, the adoption of a fidelity observation tool, and coordinated meeting times for teams to plan and share ideas.

In order to implement evidence-based practices with fidelity, district supervisors, administrators, in-service trainers, and peer coaches need to develop, utilize, and support on-going, facilitated support groups at individual school sites so teachers can work with other professionals and problem-solve site-based solutions to the inevitable challenges of implementation. Once a district or faculty adopts a particular evidence-based practice, the next step is the initial implementation efforts coupled with monitoring fidelity, the development of positive peer-coaching partnerships, and the alignment of district policy to support these innovative implementation sites. After the initial implementation, the next step for schools, such as the two in the current study, is school-wide implementation and sustainability of the evidence-based practice, which requires a long-term commitment of time, effort, and training funds. Then over time, an evidence-based practice can evolve into the standard instructional practice of a school, recognized and supported by all members of the staff.

Implications for Research

The first goal achieved by this research study was to add to the current knowledge and research base of studies that are available for purveyors and practitioners regarding the implementation of differentiated instruction, fidelity assessment, support group models, and implementation science. In order to bridge the research to implementation gap, researchers and practitioners must have two-way conversations about what works and what does not work. Researchers need to listen to needs of classroom teachers and the implementation problems that occur in today's diverse classrooms. These enhanced partnerships and practices between teachers will help establish professional learning communities at actual implementation sites, which will further support the implementation of evidence-based practices (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005).

Several possible research applications can be linked to the current study. First, response-to-intervention researchers are currently investigating promising models that promote the use of evidence-based instruction in the general education environment. In the current study, the differentiated instruction support group model has shown promise for helping to remediate the academic deficits of students who are "at-risk." Further studies should be conducted in which the current model is utilized as a first tier of primary prevention in the general education classroom to help remediate student deficits quickly and reduce the number of unnecessary evaluations for special education services. Second, in order to expand the current study, the next research steps could include: replication of the same model and method in other regions and with other student populations; further data collection on the use of the DI: FIT; the examining of sub-

populations within the sample to see if the achievement varied by populations, i.e., students with disabilities; use of the model with random-controlled trials; and a detailed analysis of differentiated instruction components to determine which components had the greatest impact on academic achievement, i.e. flexible grouping, tiered lessons, use of on-going assessment to guide instruction.

The rigorous search for effective core intervention components will require much time because the same model and methods will need to be replicated many times with different combinations of components to see which components of the particular evidence-based practice, in this case differentiated instruction, hold the most promise for producing the desired effect. Unfortunately, this type of study will require the support of a foundation of federal grant money to carry out the detailed analysis over time. To achieve this, Fixsen, Naoom, Blasé, Friedman, and Wallace (2005) recommend that, "federal and state governments need to invest in the development and use of implementation strategies and methods that are grounded in research and elaborated through accumulated experience" (p. 73).

Equally important is the question of how to maintain teachers' implementation of evidence-based practices over time. The problem of sustainability is a complex one that will involve many layers of research. There is the question of how much each teacher's initial perceptions and/or technical teaching ability effects their participation and fidelity. In the current study, there was a high percentage of non-tenured teachers and yet they were able to help their students make academic gains. In order to maintain implementation fidelity and achieve sustainability, what incentives will districts and administrators have to make available to classroom teachers and the designated school

facilitator? Will teachers continue to implement after the first year, or will they need an incentive?

References

Affholder, L. P. (2003). Differentiated instruction in inclusive elementary classrooms (Doctoral dissertation, University of Kansas, 2003).

ASCD website: About the Authors. (2004). Retrieved November 20, 2004 from the Association for Supervision and Curriculum Development Website: http://www.ascd.org/ publications/books

Baumgartner, T., Lipowski, M. B., & Rush, C. (2003). *Increasing reading achievement of primary and middle school students through differentiated instruction*. Master of Arts Research Project. University of Skylight Professional Development Field-Based Master's Program. (ERIC Document No. ED479203).

Blozowich, D. G. (2001). Differentiated instruction in heterogeneously-grouped sixth-grade classrooms (Doctoral dissertations, Immaculata College, 2001).

Boyd, K. (2001). A case study: Climate, best practices, administrative support, and staff attitudes in a middle school inclusion program (Doctoral dissertation, University of Central Florida, 2001).

Brandt, J. (1998). *Powerful learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Chapman, C., & King, R. (2005). Eleven practical ways to guide teachers toward differentiation. *Journal of Staff Development, 26*(4), 20-25.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised edition). New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.

Cohen. J. (1992). A power primer. *Psychology Bulletin, 112*(1), 155-159.

Cook, T. D., & Campbell, D. T. (1975). The design and conduct of quasi-experiments and true experiments in field settings. In M. D. Dunette and J. P. Campbell, *Handbook of Industrial and Organizational Research*, 223-325. Chicago, IL: Rand McNally.

Council for Educational Policy, Research, and Improvement (CEPRI). (2005). In-service education: The challenge of determining cost and effectiveness. Retrieved July 14, 2006, from http://www.cepri.state.fl.us/pdf/2005%20Exec%20Summary%20and%20Full%20Inservice%20Report.pdf

Davis, L. D. (2003). Exploring the relationship between study groups, transfer of professional development, and adult learning styles (Doctoral dissertation, Lehigh University, 2003).

Fixsen, D., Naoom, S., Blasé, K., Friedman, R., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Fleming, L. C., & Baker, P. H. (2002). Differentiating in the classroom: A study of student teachers (Report No. SP041709). Ashland, OH: Ashland University. (ERIC Document Reproduction Service No. ED479480).

Florida Department of Education. (2005). *Assessment and accountability briefing book: FCAT, school accountability, and teacher certification tests.* Retrieved March 30, 2006, from http://firn.edu/doe/sas/fcat/pdf/fcataabb.pdf

Florida Department of Education (2007). *Florida School Grades: Adequate Yearly Progress Report.* Retrieved August 16, 2007, from http://schoolgrades.fldoe.org/default.asp

Forness, S. (2001). Special education and related services: What have we learned from meta-analysis? *Exceptionality, 9*(4), 185-197.

Fullan, M. (2001). *Leading in a culture of change.* San Francisco, CA: Jossey-Bass.

Good, J., & Weaver, A. (2001). *Interpreting teachers' stories to inform education practices: A review of the literature.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association: Little Rock, AR. (ERIC Document No. ED461637).

Grant, S. G. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York state-testing program. *Education Policy Analysis Archives, 8*(14), 1-27. Retrieved October 5, 2005, from Education Policy Analysis Archives electronic journal: http:// epaa.asu.edu /epaa/v8n14.html

Greenberg, M., Weissberg, R., O'Brien, M., Zins, J., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*, 466-474.

Greenwood, C. R., & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education, 24*(4), 276-289.

Gregory, G. H. (2003). *Differentiated instructional strategies in practice: Training, implementation, and supervision*. Thousand Oaks, VA: Corwin Press.

Guskey, T. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*(5), 5-12.

Hale, M.S. (1999). Collaboration in a study group as a vehicle for professional development (Doctoral dissertation, University of North Carolina, 1999).

Heacox, D. (2002). *Differentiating instruction in the regular classroom: How to reach and teach all learners, Grades 3-12*. Minneapolis, MN: Free Spirit Inc.

Hobson, B. (2004). An analysis of differentiated strategies utilized by middle school teachers in a heterogeneously grouped classroom (Doctoral dissertation, Capella University, 2004).

Hodge, P. (1997). An analysis of the impact of a prescribed staff development program in differentiated instruction on student achievement and the attitudes of teachers and parents toward that instruction (Doctoral dissertation, University of Alabama, 1997).

Hornsby, J., & Diket, R. (1999). Nurturing environment: A community of learners' classroom. *The Delta Kappa Gamma Bulletin, 65*(2), 5-10.

Joyce. B. & Showers, B. (2002). *Student achievement through staff development* (3rd Ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Klingner, J., Ahwee, S., Pilonieta, P., & Menendez, R. (2003). Barriers and facilitators in scaling up research-based practices. *Exceptional Children, 69*(4), 411-429.

Kutash, K., & Duchnowski, A. (2005). *Fidelity instrument for measuring the use of evidence-based academic strategies in special education classrooms*. Instrument developed as part of Office of Special Education Programs (U.S.D.O.E.) Grant #H324T000019.

Lawrence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards-based learning that benefit the whole class. *American Secondary Education, 32*(3), 34-62.

Lester, J. (2003). Planning effective secondary professional development programs. *American Secondary Education, 32*(1), 49-61.

Little, J. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 15*(2), 129-151.

Martin, B. (2000). Exploring the effects of collegial study groups on the transfer of professional development training to classroom implementation (Doctoral dissertation, Leigh University, 2000).

McAdamis, S. (2001). Teachers tailor their instruction to meet a variety of students needs. *Journal of Staff Development*, *22*(2), 1-5. Retrieved November 27, 2004, from the National Staff Development Council Website: www.nsdc.org/library/ publications/ jsd/mcadamis222.cfm

Mokrue, K., Elias, M. J., & Bry, B. H. (2005). Dosage effect and the efficacy of a video-based teamwork-building series with urban elementary school children. *Journal of Applied School Psychology, 21*(1), 67-97.

Murphy, C., & Lick, D. (2001). *Whole faculty study groups: A powerful way to change schools* (2nd ed.). Thousand Oaks, CA: Corwin Press.

The National Academy of Sciences (NAS). (2002). *The national academy of sciences report: Minorities in special and gifted education.* Washington, DC: National Academy Press.

National Council of Teachers of Mathematics. (2005). Overview of principles and standards for school mathematics. Retrieved July 1, 2006, from http://www.nctm.org/standards/overview.htm

Osborne, B. (1993, March 25). *Understanding change in a time of change*. Paper presented at the Annual Conference on Creating the Quality School: Oklahoma City, OK. (ERIC Document No. ED358544).

Pfaff, M. E. (1999). The effects on teacher efficacy of school-based collaborative activities structured s professional study groups (Doctoral dissertation, University of Maryland, 1999).

President's Commission on Excellence in Special Education (PCESE). (2002). *A New Era: Revitalizing Special Education for Children and Their Families.* Washington, DC: U.S. Department of Education Office of Special Education and Rehabilitative Services. Retrieved May 24, 2005, from http://www.ed.gov/inits/commissionsboards/_whspecialeducation/reports.html

Richardson, J. (1997, March). Dealing with resisters biggest challenge for staff developers. *The Developer*. Retrieved on November 27, 2004 from the National Staff Development Council website: www.nsdc.org/library/publications/developer/dev3-97rich.cfm

Seachrest, L., West, S. G., Phillips, M. A., Redner, R., & Yeaton, W. (1979). Some neglected problems in evaluation research: Strength and integrity of treatments. *Evaluation Studies Review Annual, 4,* 15-35.

Showers, B., Joyce, B., & Bennett, C. (1987). Synthesis of research on staff development: A framework for future study and the state-of-the art analysis. *Educational Leadership, 45*(3), 77-87.

Sparks, D. (2001). Why change is so challenging for schools: An interview with Peter Senge. *Journal of Staff Development*, *22*(3), 1-6. Retrieved on November 27, 2004 from the National Staff Development Council website: www.nsdc.org/ library/publications/ jsd/senge223.cfm

Spencer, S., & Logan, K. (2003). Bridging the gap: A school based staff development model that bridges the gap from research to practice. *Teacher Education and Special Education, 26*(1), 51-62.

Tomlinson, C. A. (1995). Deciding to differentiate instruction in middle school: One school's journey. *Gifted Child Quarterly, 39,* 77-87.

Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners.* Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2000). Reconcilable differences: Standards-based teaching and differentiation. *Educational Leadership, 58*(1), 6-11.

Tomlinson, C. A. (2001). *How to Differentiate Instruction in Mixed-Ability Classrooms* (2nd Ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2003a). Deciding to teach them all. *Educational Leadership, 61*(2), 6-11.

Tomlinson, C. A. (2003b). *Fulfilling the promise of the differentiated classroom: Strategies and tools for responsive teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2005, July 7). You can't rush differentiated instruction. Brief from 2005 ASCD Summer Conference. Retrieved October 9, 2005, from ASCD website: http://2005ascdsummerconference.blogspot.com/2005/07/ you-cant-rush-differentiated. html

Tomlinson, C. A., & Allan, S. D. (2000). *Leadership for differentiating schools & classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A., Tomchin, E. M., Callahan, C. M., Adams, C. M., Pizzat-Tinnin, P., Cunningham, C. M., et al. (1994). Practices of pre-service teachers related to gifted and other academically divers learners. *Gifted Child Quarterly, 38,* 106-114.

U.S. Department of Education. (2002). *The Elementary and Secondary Education Act* (No Child Left Behind Act of 2001). Retrieved October 9, 2005, from http://www.ed.gov/_policy/elsec/leg/esea02/index.html

U. S. Department of Education, Consortium for Policy Research in Education (CPRE). (1995). *Professional development today.* Washington, DC: Institute of Educational Sciences. Retrieved July 14, 2006, from http://www.ed.gov/ pubs/CPRE/t61/index.html

U. S. Department of Education, National Center for Educational Statistics. (2000). *Teacher preparation and professional development: 2000.* Washington, DC: Institute of Educational Sciences. Retrieved July 14, 2006, from http://nces.ed.gov/surveys/frss/publications/2001088/

U.S. Department of Education, National Center for Education Statistics. (2004). *Digest of education statistics tables and figures.* Retrieved October 14, 2005, from http://nces. ed.gov/programs/digest/d04/lt2.asp#c2_1

U.S. Department of Education, Office of Post Secondary Education (OPSE). (2005). *The secretary's fourth annual report on teacher quality: A highly qualified teacher in every classroom*. Retrieved October 14, 2005, from http://www.ed.gov/about/ reports/annual/ teachprep/2005Title2-Report.doc

Vanfleet, K. P. (2002). An analysis of differentiated instruction applied to teaching academics in secondary schools as a means to facilitate the inclusion of students with mild to moderate disabilities in the general education classroom (Doctoral dissertation, The University of Alabama, 2002).

Webster-Stratton, C. (2003). Quality training, supervision, on-going monitoring, and agency support: Key ingredients to implementing The Incredible Years programs with fidelity. Retrieved October 9, 2005, from The Incredible Years website: http://www.incredible years.com/research/ article-quality-key-ingredients-fidelity-04.pdf

Wertheim, C., & Leyser, Y. (2002). Efficacy beliefs, background variables, and differentiated instruction of Israeli prospective teachers. *The Journal of Educational Research*, *96*(1), 54-63.

Westling, D. L., Cooper-Duffy, K., Prohn, K., Ray, M., & Herzog, M. J. (2005, May/June). Building a teacher support program. *Teaching Exceptional Children, 37*(5), 8-13.

Willis, S., & Mann, L. (2000, Winter). Differentiating instruction: Finding manageable ways to meet individual needs. *Curriculum Update.* Retrieved November 10, 2004 from Association for Supervision and Curriculum Development (ASCD) website: http://www.ascd.org/cms/objectlib/ascdframeset/index.cfm?publication

Appendices

Appendix A

Strategies for Differentiating Instruction

1. *Stations:* The use of stations involves setting up different spots in the classroom where students work on various tasks simultaneously. These stations invite flexible grouping because not all students need to go to all stations all the time.
2. *Compacting*: This strategy encourages teachers to assess students before beginning a unit of study or development of a skill. Students who do well on the pre-assessment do not continue work on what they already know.
3. *Agendas:* Agendas are personalized lists of tasks that a student must complete in a specified time, usually two to three weeks.
4. *Complex Instruction:* This strategy uses challenging materials, open-ended tasks, and small instructional groups. Teachers move among the groups as they work, asking students questions and probing their thinking.
5. *Orbital Studies:* These independent investigations, generally lasting three to six weeks, revolve around some facet of the curriculum. Students select their own topics, and they work with guidance and coaching from the teacher.
6. *Entry Points:* This strategy from Howard Gardner proposes student exploration of a given topic through as many as five avenues: narrational (presenting a story), logical-quantitative (using numbers or deduction), foundational (examining philosophy and vocabulary), aesthetic (focusing on sensory features), and experiential (hands-on).
7. *Problem-Based Learning:* This strategy places students in the active role of solving problems in much the same way adult professionals perform their jobs.
8. *Choice Boards:* With this strategy, work assignments are written on cards that are placed in hanging pockets. By asking a student to select a card from a particular row of pockets, the teacher targets work toward student needs yet allows student choice.
9. *4MAT:* Teachers who use 4MAT plan instruction for each of four learning preferences over the course of several days on a given topic. Some lessons focus on mastery, some on understanding, some on personal involvement, and some on synthesis.

Note. From *The Differentiated Classroom: Responding to the Needs of All Learners,* by Carol Ann Tomlinson, 1999, Alexandria, VA: Association for Supervision and Curriculum Development, p. 75 – 93. Reprinted with permission from ASCD, February 2006.

Appendix B

Comparison of the Traditional vs. Differentiated Classroom

Traditional Classroom	Differentiated Classroom
• Student differences are masked or acted upon when problematic	• Student differences are studied as a basis for planning
• Assessment is most common at the end of learning to see "who got it"	• Assessment is on-going and diagnostic in order to make instruction more responsive to learner needs
• A relatively narrow sense of intelligence prevails	• Focus on multiple forms of intelligence is evident
• A single definition of excellence exists	• Excellence is defined in large measure by individual growth from a starting point
• Student interest is infrequently tapped	• Students are frequently guided in making interest-based learning choices
• Relatively few learning profile options are taken into account	• Many learning profile options are provided
• Whole-class instruction dominates	• Many instructional arrangements are used
• Coverage of texts and curriculum guides drives instruction	• Student readiness, interest, and learning profile shape instruction
• Mastery of facts and skills out-of-context are the focus of learning	• Use of essential skills to make sense of and understand key concepts and principles is the focus of learning
• Single option assignments are the norm	• Multi-option assignments are frequently used
• Time is relatively inflexible	• Time is used flexibly according to student need
• A single text prevails	• Multiple materials are provided
• Single interpretations of ideas and events may be sought	• Multiple perspectives on ideas and events are sought
• The teacher directs student behavior	• The teacher facilitates students' skills at becoming more self-reliant learners
• The teacher solves problems	• Students help other students solve problems
• The teacher provides whole-class standards for grading	• Students work with the teacher to establish both whole-class and individual learning goals
• A single form of assessment is often used	• Students are assessed in multiple ways

Note. From *The Differentiated Classroom: Responding to the Needs of All Learners*, by Carol Ann Tomlinson, 1999, Alexandria, VA: Association for Supervision and Curriculum Development, p. 16. Reprinted with permission from ASCD, February 2006.

Appendix C

Qualities of a Supportive Classroom Environment for Differentiation

---

A supportive classroom environment is vital to your success in differentiating instruction.

Such an environment:

- Promotes acceptance of differences
- Affirms that all students have learning strengths
- Acknowledged that students learn at different rates and in different ways
- Recognizes that for work to be fair, it must sometimes be different
- Acknowledges that success means different things to different people
- Allows students to work with various people for various purposes
- Recognizes that the key to motivation is interest, and that all students have different interests
- Promotes personal responsibility for learning
- Builds feelings of personal competence and confidence in learning
- Values effort and "personal best"
- Nurtures skills of independence
- Supports and celebrates student success in challenging work
- Encourages exploration of each student's interests, strengths, and learning preferences
- Nurtures the creative spirit in all students
- Honors everyone's work

Note. From *Differentiating Instruction in the Regular Classroom: How to Reach and Teach All Learners, Grades 3-12*, by Diane Heacox, 2002, Minneapolis, MN: Free Spirit Inc, p. 12-13. Reprinted with permission from ASCD, February 2006.

Appendix D

Differentiated Instruction: Fidelity Implementation Tool (DI: FIT)

Teacher: _____ School: _____ Observer: _____

Date: _____ Class: _____ Lesson: _____

#	DI Strategy	Evaluator Task	Score "1" if:	Circle One		
1	Teacher ensures students understand the purpose of the lesson	Observe if agenda and/ or objectives are posted or ask a student to explain the objective of the lesson.	Goal(s) is (are) visible, two out of three students can verbalize the objective of the lesson	1		0
2	Teacher creates respectful assignments	Observe the lesson	The observed lesson is designed with at least one clear example of cultural, ethnic, and linguistic sensitivity	1		0
3	Teacher creates respectful assignments	Observe the lesson	Assignments are designed so all students are working toward the same goal and/ or understanding.	1		0
4	Lesson is differentiated by content, product, or process	Observe the lesson	The observed lesson is differentiated by content, product, or process	1		0
5	Lesson is differentiated according to students' readiness, interests, or learning profiles	Observe the lesson	The observed lesson is designed to address readiness, interests, or learning profiles	1		0
6	Visible use of supports	Observe the lesson	The teacher employs clear examples of supports e.g., organizers, peers, manipulatives, technology	1		0
7	Class functions as a community	Observe the lesson	There is evidence in the classroom of at least two of the following: positive reinforcement by peers, cooperative learning activities, teacher-student collaboration, student-student collaboration, peer support	1		0
8	Students demonstrate genuine interest in learning	Observe the lesson	Eighty percent or more of students appear to be engaged and interested in the lesson, e.g., asking questions, participating, interacting with others	1		0
9	Teacher uses anchor activities	Observe the lesson	The teacher has established activities and routines for students who are finished	1		0
10	Teacher acts as a facilitator	Observe the lesson	The teacher guides learning and nurtures student independence	1	0	NA
11	Teacher promotes acceptance of differences	Observe the lesson	The teacher models and nurtures students' acceptance of differences if there is an opportunity	1	0	NA

Appendix D (Continued)

#	DI Strategy	Evaluator Task	Score "1" if:	Circle One	
12	All assignments provide a slight challenge for learners	Ask the teacher to justify how he/she decided and designed the difficulty of the assignments for each child and/ or group	Teacher can verbalize criteria and reasons for how the assignments were created and assigned. Answer either mentions or alludes that students will be challenged	1	0
13	Lesson is centered on key concepts or essential learning	Ask the teacher to explain the purpose of the lesson and how he/she decided on what to include or not include	Teacher can either verbalize or provide recent* lesson plans that specify the standards and/ or justify lesson importance	1	0
14	Teacher uses flexible and purposeful grouping	Ask the teacher how he/ she selects student groups and to give examples	Teacher can give multiple, recent* examples of diverse and flexible student groupings	1	0
15	Use of learning stations and/ or independent study	Ask the teacher to describe or provide a recent* DI lesson plan that employed stations or independent study	Teacher can provide recent* examples of student learning stations and/ or independent study	1	0
16	Students are given meaningful, learning choices	Ask the teacher to describe or provide a recent* DI lesson plan that employed student choice options	Teacher can provide recent* examples of meaningful, learning choices provided to students	1	0
17	Teacher provides a variety of product options	Ask the teacher to describe or provide a recent* DI lesson plan that employed product options	Teacher can provide recent* examples of a variety of product options provided to students	1	0
18	Teacher uses a variety of assessment tools	Ask the teacher to describe and provide examples of the types of assessment used	Teacher can provide recent* proof of multiple forms of assessment used, e.g., portfolios, rubrics, traditional tests	1	0
19	Teacher utilizes on-going assessment	Ask the teacher to explain how he/she uses assessment to guide instruction	Teacher can provide specific examples of how he/ she used recent* assessments to guide instruction	1	0
20	Teacher and students collaborate in the learning process	Ask the teacher how he/she involves the students in the learning process	Teacher provides students with the opportunity to be a stakeholder in the learning process, i.e., the teacher honors students' ideas	1	0
	* recent (within one month)		Total score (20 possible)		

Deborah Hellman
January 2006

Appendix E

List of Book Titles Included in Each School's Reference Library

---

Chapman, C., & King, R. (2003). *Differentiated instructional strategies for reading in the content area*. Thousand Oaks, CA: Corwin Press.

Coil, C., & Merritt, D. (2001). *Solving the assessment puzzle: Piece by piece*. Marion, IL: Pieces of Learning.

Cramer, K., Twyman, S., & Winholtz, W. (1998). *61 cooperative learning activities for science classes*. Portland, ME: J. Weston Walch.

Dodge, J. (2006). *Differentiation in action: A complete resource with research-supported strategies to help you plan and organize differentiated instruction and achieve success with all learners.* New York, NY: Scholastic, Inc.

Drapeau, P. (2004). *Differentiated instruction: Making it work : A practical guide to planning, managing, and implementing differentiated instruction to meet the needs of all learners.* New York, NY: Scholastic, Inc.

Forsten, C. (2003). *Differentiating textbooks: Strategies to improve student comprehension & motivation.* Peterborough, NH: Crystal Springs Books.

Forsten, C., Grant, J., & Hollas, B. (2002). *Differentiated instruction: Different strategies for different learners*. Peterborough, NH: Crystal Springs Books.

Heacox, D. (2002). *Differentiating instruction in the regular classroom: How to reach and teach all learners, Grades 3-12*. Minneapolis, MN: Free Spirit Inc.

Hollas, B. (2005). *Differentiating instruction in a whole-group setting*. Peterborough, NH: Crystal Springs Books.

Paterson, K. (2005). *Differentiated learning: Language and literacy projects that address diverse backgrounds and cultures*. Markham, Ontario: Pembroke Publishers.

Tilton, L. (2003). *The teacher's toolbox for differentiating instruction: 700 strategies, tips, tools, and techniques*. Shorewood, MN: Covington Cove Publications.

Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners.* Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd Ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2003). *Fulfilling the promise of the differentiated classroom: Strategies and tools for responsive teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. & Allan, S. D. (2000). *Leadership for differentiating schools & classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.

Wormeli, R. (2006). Fair Isn't Always Equal: Assessing & Grading In the Differentiated Classroom. Portland, ME: Stenhouse Publishers.

Appendix F

Facilitated Support Group Feedback Form

Name: _____ Date: _____

I have read the minutes of this month's support group session. Yes ___ No ___

I agree that the minutes are accurate. Yes ___ No ___

Corrections needed:

_____

_____

_____

_____

Suggestions for next month:

_____

_____

_____

Appendix G

The Differentiated Classroom Observation Form

Check the appropriate box next to each item. Use the comment box to provide ideas for improvement in specific areas. If the form is completed during multiple observations, use tally marks. Review the results with the teacher as soon as possible to identify specific areas for improvement and to praise strengths.

Teacher: Grade Level/Subject Area: Observer: Date:

	Evidence of Implementation			
	Often	Sometimes	Little or no	Comments
PHYSICAL ENVIRONMENT Presents an inviting, relaxed environment for learning				
Provides comfortable desks and work areas				
Contains individual, designated personal spaces for extra books and other items				
Is designed for quick and easy groupings of tables and chairs				
Is arranged for teacher and student movement during work sessions				
Provides work areas for individual needs, including knowledge/ability levels				
Reflects current content or skills through student displays and artifacts				
TEACHER BEHAVIORS Works with total groups, individuals, and small groups				
Monitors individuals and small groups				
Uses a variety of ongoing assessment tools such as checklists, surveys, and anecdotal records				
Applies assessment information to guide instruction				
Addresses academic, emotional, social, and physical student needs				
Provides time for students to actively process information				
Gives specific feedback to individuals and/or small groups				
STUDENT ENGAGEMENT Exhibits on-task behavior while working alone				
Works effectively in small groups				
Works on their individual knowledge or ability levels				
Uses materials/ resources on the student's own level of success				
Feels respected and emotionally safe				
Uses self-discipline				
MATERIALS/ RESOURCES Includes a variety of reading levels related to the subject or topic				
Are accessible to students				
Supports the standards and topic				
Are age-appropriate				
Are up-to-date				
Are available in an adequate number for the class size				
Include appropriate reference sources and materials				
INSTRUCTIONAL STRATEGIES Uses a variety of assessment tools before, during, and after learning				
Uses a variety of instructional strategies and activities to teach standards				
Meets the divers needs of learners				
Engages students in various flexible grouping designs				
Uses centers and/or stations for individual and small group instruction				
Engages students with projects and/ or problem-solving activities				
Presents students with choices in learning activities				

Note. From 11 Practical Ways to Guide Teachers toward Differentiation (and an evaluation tool), by C. Chapman and R. King, 2005, *Journal of Staff Development, 26*(4), p. 24. Used with permission of the National Staff Development Council, www.nsdc.org, 2007. All rights reserved.

Appendix H

Teacher/Peer Reflection on Differentiation

The following scale may be useful

(1) little or no evidence
(2) to some degree
(3) demonstrates competence
(4) demonstrates proficiency
(5) demonstrates exemplary performance

GENERAL					
Pre-assesses students to determine level of understanding.	1	2	3	4	5
Assesses student interests.	1	2	3	4	5
Identifies students' learning profiles.	1	2	3	4	5
Develops a student-centered classroom.	1	2	3	4	5
Ensures respectful assignments for all learners.	1	2	3	4	5
Consistently uses flexible grouping.	1	2	3	4	5
Varies the pace of learning for varying learner needs.	1	2	3	4	5
Utilizes active learning.	1	2	3	4	5
Demonstrates escalating expectations.	1	2	3	4	5
Students' grades reflect individual growth and progress.	1	2	3	4	5
CONTENT					
Differentiates using major concepts and generalizations.	1	2	3	4	5
Uses a variety of materials other than the standard text.	1	2	3	4	5
Various support mechanisms (e.g., reading buddies, organizers, study guides).	1	2	3	4	5
PROCESS					
Activities necessitate that students *do* something with their knowledge (apply and extend major concepts and generalizations as opposed to just repeating it back).	1	2	3	4	5
Uses higher-level tasks for all learners (e.g., application, elaboration, providing evidence, synthesis) to provide appropriate challenge.	1	2	3	4	5

PROCESS (continued)					
Uses tiered activities.	1	2	3	4	5
Activities involve all learners in both critical and creative thinking.	1	2	3	4	5
Varies tasks along continuum of the equalizer.	1	2	3	4	5
Varies tasks by students interests.	1	2	3	4	5
Varies tasks by learner profile.	1	2	3	4	5
PRODUCT					
Provides opportunities for student products to be based upon the solving of real and relevant problems.	1	2	3	4	5
Allows for a wide range of product alternatives (e.g., oral, visual, kinesthetic, musical, spatial, creative, practical).	1	2	3	4	5
Product assignments differ based on individual (or group) readiness, learning needs, and interest.	1	2	3	4	5
Teacher supports students in using a wide range of varied resources.	1	2	3	4	5
Product assignment necessitates that students conduct research.	1	2	3	4	5
Product assignment balances structure and choice.	1	2	3	4	5
Encourages students to use different avenues of exploration and a variety of media.	1	2	3	4	5
Works with individual students (or groups) to determine what form the product will take.	1	2	3	4	5
Necessitates that students apply key understandings and skills of the subject to their own interest areas.	1	2	3	4	5
Works with individual students to apply key understandings and skills of the discipline by which the product will be judged.	1	2	3	4	5
Uses both formative and summative evaluation.	1	2	3	4	5

INSTRUCTIONAL/MANAGEMENT STRATEGIES					
Uses compacting.	1	2	3	4	5
Uses student learning contracts.	1	2	3	4	5
Uses independent study.	1	2	3	4	5
Uses interest centers/groups.	1	2	3	4	5
Uses learning centers/groups.	1	2	3	4	5
Uses various instructional strategies to differentiate (e.g. organizers, cubing, etc.).	1	2	3	4	5
Uses high-level cooperative strategies (e.g., complex instruction, group investigation).	1	2	3	4	5
Other _____	1	2	3	4	5
Other _____	1	2	3	4	5

Note. From *Leadership for Differentiating Schools & Classrooms*, by C. A. Tomlinson and S. D. Allan, 2000, Alexandria, VA: Association for Supervision and Curriculum Development, pp. 144-146. Reprinted with permission from ASCD, February 2006.

Appendix I

DI Support Group Feedback

Was the number of meetings sufficient? Too many? Too few? _____

_____

_____

Was the length of the support group meetings effective? Too short? Too long?

_____

_____

_____

Was the discourse, professional discussion with colleagues, helpful or not? How? Be specific. _____

_____

_____

_____

_____

Was the DI conference area on your e-mail desktop helpful? Explain.

_____

_____

_____

_____

Has participation in the on-going support group helped to sustain your implementation of differentiated instruction? _____

_____

_____

_____

_____

Appendix I (Continued)

What attributes of the DI Support Group Meetings do you think were helpful and would like to see continued? What areas or qualities of the DI Support Group Meetings do you think we need to change or improve on?

Helpful or Positive Aspects to Keep	Suggested Changes or Improvements

Appendix J

Informed Consent to Participate in Research
Information to Consider Before Taking Part in this Research Study

---

Researchers at the University of South Florida (USF) study many topics. To do this, we need the help of people who agree to take part in a research study. This form tells you about this research study. We are asking you to take part in a research study that is called:

Implementing Differentiated Instruction in Urban, Title I Schools: Effects of Facilitated Support Groups and Program Fidelity on Student Achievement

The person who is in charge of this research study is Deborah Hellman. Other research personnel who you may be involved with include: (*names removed for security*).
The research will be done at School A and School B (*names removed for security*).

Purpose of the study
The purpose of this research study is to collect and analyze data on the impact of differentiated instruction, facilitated support groups, and implementation fidelity on student achievement.

Study Procedures
You have been selected as one of 52 possible participants for this study. If you are selected as one of the approximately 26 teachers in the treatment group, you will participate in a series of five facilitated support groups that will last approximately 120 minutes each. You will have access to a library of support materials and will serve as part of the teacher support/ study group. At the conclusion of each meeting, you will select one or more resources to review for 30 minutes and then select at least one strategy to implement in the next three to four weeks in your classroom. At the next meeting, you will share their experience and provide feedback to the group concerning the strategy implemented. During these sessions, the conversations will be digitally recorded to ensure accuracy of the information and comments collected, no names will be used in the final reports. Once per nine-weeks, the principal investigator and/ or research assistant will observe you to determine the degree of fidelity that you demonstrate with respect to differentiated instruction. If you are selected as one of the approximately 26 teachers in the control group, you will carry out the normal requirements of your teaching position and you will be observed once per nine weeks to determine the degree of fidelity that you demonstrate with respect to differentiated instruction.

At the end of the school year, student performance data will be collected and analyzed from your students and the students of the other participants in the study. The effect of teacher fidelity to the differentiated instruction model will also be analyzed. Data collection and analysis will last from January 2007 until December 2007.

Alternatives
You have the alternative to choose not to participate in this research study.

Benefits
The potential benefits to you are: by taking part in this research study, you may increase our overall knowledge of differentiated instruction strategies. If you are selected as a participant in the treatment group you will be provided with direct support services, strategies, and access to resources. The principal investigator will also be available as a mentor throughout the process. If you are selected as part of the control group, you will help to add to the body of evidence of implementation of evidence based practices.

Risks or Discomfort
There are no known risks to those who take part in this study**.**

Compensation
We will not pay you for the time you volunteer while being in this study.

Confidentiality
We must keep your study records confidential. The data obtained from you will be combined with data from others in the publication. The published results will not include your name or any other information that would personally identify you in any way. All written data, audiotapes, and videotapes will be anonymously coded and stored in a secured file cabinet in the researcher's office. The signed consent forms will be stored for three years in a secured file cabinet in the PI's office along with all data, tapes, and notes. After that time, they will be shredded. However, certain people may need to see your study records. By law, anyone who looks at your records must keep them completely confidential. The only people who will be allowed to see these records are:

- The research team, including the Principal Investigator, study coordinator, and all other research staff.
- Certain government and university people who need to know more about the study. For example, individuals who provide oversight on this study may need to look at your records. This is done to make sure that we are doing the study in the right way. They also need to make sure that we are protecting your rights and your safety.) These include:

- The University of South Florida Institutional Review Board (IRB) and the staff that work for the IRB. Other individuals who work for USF that provide other kinds of oversight may also need to look at your records.
- The Florida Department of Health, people from the Food and Drug Administration (FDA), and people from the Department of Health and Human Services (DHHS).
- The school district's Research and Evaluation Office may look at the study records to make sure the study was done correctly.

- We may publish what we learn from this study. If we do, we will not let anyone know your name. We will not publish anything else that would let people know who you are.

Voluntary Participation / Withdrawal

You should only take part in this study if you want to volunteer. You should not feel that there is any pressure to take part in the study, to please the investigator or the research staff. You are free to participate in this research or withdraw at any time. There will be no penalty or loss of benefits you are entitled to receive if you stop taking part in this study. Your decision to participate or not to participate will not affect your or job status.

If you have any questions, concerns or complaints about this study, call Deborah Hellman at 493-3302.

If you have questions about your rights, general questions, complaints, or issues as a person taking part in this study, call the Division of Research Integrity and Compliance of the University of South Florida at (813) 974-9343.

If you experience an adverse event or unanticipated problem call Deborah Hellman at 493-3302.

Consent to Take Part in this Research Study

It is up to you to decide whether you want to take part in this study. If you want to take part, please sign the form, if the following statements are true.

I freely give my consent to take part in this study. I understand that by signing this form I am agreeing to take part in research. I have received a copy of this form to take with me.

_____ _____

Signature of Person Taking Part in Study Date

_____

Printed Name of Person Taking Part in Study

Statement of Person Obtaining Informed Consent

I have carefully explained to the person taking part in the study what he or she can expect.
I hereby certify that when this person signs this form, to the best of my knowledge, he or she understands:

- What the study is about.
- What procedures/interventions/investigational drugs or devices will be used.
- What the potential benefits might be.
- What the known risks might be.

I also certify that he or she does not have any problems that could make it hard to understand what it means to take part in this research. This person speaks the language that was used to explain this research.

This person reads well enough to understand this form or, if not, this person is able to hear and understand when the form is read to him or her.

This person does not have a medical/psychological problem that would compromise comprehension and therefore makes it hard to understand what is being explained and can, therefore, give informed consent.

This person is not taking drugs that may cloud their judgment or make it hard to understand what is being explained and can, therefore, give informed consent.

_____ _____

Signature of Person Obtaining Informed Consent Date

_____

Printed Name of Person Obtaining Informed Consent

Appendix K

District In-service Evaluation Summary: School A

District Level Questions

1. Training content was appropriate and built upon knowledge / experiences of intended participants.

A. Strongly Agree	91%
B. Agree	9%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

2. Content of training included information relevant and useful in my position.

A. Strongly Agree	82%
B. Agree	18%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

3. Training activities, assignments, and / or materials were related to course objectives.

A. Strongly Agree	82%
B. Agree	18%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

4. Trainer demonstrated knowledge and positive attitude toward content.

A. Strongly Agree	92%
B. Agree	8%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

5. Training environment was appropriate for the course.

A. Strongly Agree	90%
B. Agree	10%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

6. What is the Primary Purpose of this course?

A-Add-on Endorsement	10%
B-Alternate Certification	0%
C-Florida Educators Certificate Renewal	36%
D-Other Professional Certificate/License renewal	0%
E-Professional Skill Building	54%

7. Is this training aligned with your Individual Professional Development Plan (IPDP)?

A – Yes	100%
B – No	0%
C - N/A (Instructional Support Only)	0%

Appendix L

District In-service Evaluation Summary: School B

District Level Questions

1. Training content was appropriate and built upon knowledge / experiences of intended participants.

A. Strongly Agree	90%
B. Agree	10%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

2. Content of training included information relevant and useful in my position.

A. Strongly Agree	80%
B. Agree	20%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

3. Training activities, assignments, and / or materials were related to course objectives.

A. Strongly Agree	90%
B. Agree	10%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

4. Trainer demonstrated knowledge and positive attitude toward content.

A. Strongly Agree	90%
B. Agree	10%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

5. Training environment was appropriate for the course.

A. Strongly Agree	70%
B. Agree	30%
C. Undecided	0%
D. Disagree	0%
E. Strongly Disagree	0%

6. What is the Primary Purpose of this course?

A-Add-on Endorsement	0%
B-Alternate Certification	0%
C-Florida Educators Certificate Renewal	20%
D-Other Professional Certificate/License renewal	0%
E-Professional Skill Building	80%

7. Is this training aligned with your Individual Professional Development Plan (IPDP)?

A – Yes	90%
B – No	10%
C - N/A (Instructional Support Only)	0%

About the Author

Deborah Hellman received her bachelor's degree at the University of Florida and her master's degree in special education at the University of South Florida (USF). She has been a special education and inclusion co-teacher for 27 years and is Nationally Board Certified. Mrs. Hellman is currently a district trainer for differentiated instruction, collaborative teaching, and teaming. In addition, she actively mentors other teachers and pre-service teachers as an ESE Mentor, Model Classroom Teacher, and Professional Practice Partner with USF. She has received many honors and scholarships including Teacher of the Year, CEC Teacher of the Year, the Landis M. Stetler ESE Leadership Scholarship, USF's Predoctoral Fellowship, the Laurie Ann Richardson Masters Scholarship, and the Cathy Lynn Richardson Doctoral Scholarship. She began the doctoral program in the fall of 2004 and has been a graduate teaching instructor for EEX4070, Integrating the Exceptional Child for the past year.