2009

The Effects of Being Placed in Special Education Classes Versus General Education Classes and Teacher Certification on Students' High-Stakes Testing Scores

Bethany Ann Porter
Louisiana State University and Agricultural and Mechanical College

THE EFFECTS OF BEING PLACED IN SPECIAL EDUCATION CLASSES VERSUS GENERAL EDUCATION CLASSES AND TEACHER CERTIFICATION ON STUDENTS' HIGH-STAKES TESTING SCORES

A Dissertation

Submitted to the Graduate Faculty of the
Lousiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

By
Bethany Ann Porter
B.S., Jacksonville State University, 2003
M.A., Louisiana State University, 2006
December, 2009

ACKNOWLEDGEMENTS

I would like to extend my deepest appreciation and gratitude to Dr. George H. Noell for being an excellent mentor and friend. Without his patience, guidance, and support, this project would not have been possible. I would also like to thank Dr. Frank Gresham, Dr. Jeffrey Tiger, Dr. Mary Lou Kelley, and Dr. Cassandra Chaney for being members of my dissertation committee. I would also like to extend my appreciation to Maria Patt and Amanda Dahir for their help with the data, their patience, and support in this process. Finally, I would like to thank my husband, Luke Gleason, for his tireless support, patience, and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The introduction of new laws such as the amendments to the Individuals with Disabilities Education Act (IDEA, 1997) and the No Child Left Behind Act have changed education for special education students in the United States. Students with disabilities are now more frequently being held to the same standards as those students without disabilities. These federal laws are designed to help close the achievement gap among all students regardless of race, gender, poverty, or disability status. Special education students are now required to participate in statewide high-stakes testing programs alongside their nondisabled peers. Another movement involves including students with disabilities in the general education classroom for their instruction rather than being segregated. This movement corresponds with the "least restrictive environment" that has been a part of IDEA since its inception, but whose implementation in practice has not been consistent. Research has provided evidence for the social benefits of inclusion, but little evidence exists for the academic benefits. Special education students may benefit academically from being included in the general education classroom, but variables that affect their performance need to be investigated. Evidence exists for the positive effects of certain teacher qualifications with nondisabled students, but again, little research has looked at these effects with disabled students. The purpose of the current study was to investigate the affect on disabled students' high-stakes test performance when they are included in the general education classroom. Also, it examined certain teacher qualifications that may affect special education students' high-stakes test scores. A value-added model was used to examine these variables.

INTRODUCTION

Standardized assessments are a common aspect of school systems today. "A standardized assessment is defined as a large-scale, externally developed and mandated, uniformly administered and scored evaluation of student learning" (Wang, Beckett, & Brown, 2006, p.306). Standardized assessment practices are not new to schools. In fact, the history of standardized assessment practices starts with the Committee of Ten in the late 1800's and has continued on all the way up to the most recent No Child Left Behind Act (NCLB) of 2001 (Wang et al.). Each successive movement has placed an increasing emphasis on standardized assessment as a reform catalyst and quality control mechanism (Horn, 2002; Linn, 2000). Standardized assessments are used to hold schools accountable for student achievement on performance standards.

Not only are these assessments being used for accountability, but many states also use these assessments as criteria for graduation and/or grade promotion. Roughly half of the states in the U.S. currently require students to pass an exam to receive a high school diploma and 7 require a passing score to be promoted to the next grade (National Education Association, 2001). These numbers are expected to rise in the future. Because these tests carry such "high-stakes" they are considered high-stakes tests. High-stakes tests place high standards on students and teachers alike and with the amendments of the *Individuals with Disabilities Act* (IDEA 97) all students must be included in state and district wide assessments. This amendment was geared towards increasing participation of special education students in the general education standards-based reform and accountability programs (Schulte, Villwock, Whichard, & Stallings, 2001). The goal of these acts, that places such high value on assessments, is to close the achievement gap among all children regardless of race, class, or disability (Wang et al., 2006).

New issues arise now that schools are required to include more and more students with disabilities in their statewide assessments. The *Individuals with Disabilities Education Act* has an amendment of "least restrictive environment" (LRE). With this amendment more individuals with disabilities are given the opportunity to participate in the general education classroom. This amendment brought about movements and debates regarding mainstreaming and inclusion of special education students (Junkala & Mooney, 1986). One of the newest debates regarding mainstreaming and inclusion relates to including disabled students in accountability assessments. Now that the schools are being held accountable for students with disabilities in their assessments, what is the best method of making sure these students meet the standards?

Many variables play a part in a student's success on high-stakes tests. Some of these variables include a student's prior achievement and student demographics. Not only do the student's characteristics play a part, but also teacher characteristics affect student outcomes on high-stakes tests. It has been suggested in previous research that teachers do affect student achievement to a large degree even when student characteristics are taken into account (Sanders & Rivers, 1996; Noell & Burns, 2006). It has long been a belief that teaching matters. Many debates as to just how important and how much of an effect teacher characteristics and teaching practices have on student's success have appeared over the years though (Noell & Burns). One of these teacher characteristics that has been examined is teacher certification (Wayne & Youngs, 2003; Croninger, Rice, Rathbun, & Nishio, 2007). Are children identified as being disabled being instructed by teachers who are prepared to instruct these children? Now that students with disabilities are being included more and more in the general education classroom, it is important to know whether the teachers are adequately prepared to

instruct them. Also are these teachers competent in addressing issues such as accommodations for students and administering alternative assessments?

Not only should education policy makers be concerned with whether teachers are certified to teach special education students, but also whether teachers are certified to teach in the content area. If students are being tested on skills of English language arts and mathematics, then does it matter if the teacher instructing them in this area is not certified to do so? Now that children, teachers, and schools are being held to higher standards these issues need to be addressed to determine how they can meet these standards.

Value-added models (VAM) are an emerging method for examining the contribution of educational factors to student achievement (Tekwe et al., 2004; Lockwood & McCaffrey, 2007; Ballou, Sanders, & Wright, 2004). VAMs have increased in popularity over the last few years. VAMs in education are statistical techniques in which student performance on standardized assessments can be used to investigate effects of schools and teachers. These models typically closely resemble the structure of education using nested multilayered models in which the nesting structure of students within classrooms is preserved and may allow researchers to examine important questions in education using models that are statistically sound for the natural structure of schools (Noell, 2006; Noell, Porter, & Pratt, 2007).

Special education is a substantial element of education in the United States serving hundreds of thousands of students. Research is needed to identify variables that may affect the performance of these disabled students. Not only are disabled students being included in high-stakes testing, but there is also a movement for inclusion. This movement needs to be examined to determine whether or not it is academically beneficial

for disabled students. Other variables such as teacher qualifications and characteristics need to be examined to determine their effect on disabled students' performance on high-stakes tests. Overall, the main goal is to improve education for students with disabilities and investigating these questions may help the cause.

REVIEW OF THE LITERATURE

High-Stakes Testing

Education in the United States is a topic of discussion and debate. Many scholars believe that U.S. students are not receiving an adequate education, especially children with disabilities (Schulte, Villwock, Whichard, & Stallings, 2001; Moody, Vaughn,Hughes, & Fisher, 2000; Schulte, Osborne, & Erchul, 1998). Results reported by the National Assessment of Educational Progress (NAEP) frequently raise concerns for scholars, researchers, and politicians. Even though results seem to be improving for American students, some scholars doubt the validity of these tests (Lewis, 2005). Education is also a consistent feature of politics within the United States. This is evident with the passing of new amendments, such as the No Child Left Behind Act of 2001 and new requirements as outlined by the *Individuals With Disabilities Act Amendments of 1997*. These acts have given these standardized tests the status of high-stakes tests because there are now major consequences for students, teachers, and schools based on test results and all students must participate in testing.

Consequences attached to these high-stakes tests can be sanctions or rewards. Consequences can be centered on students, teachers, or schools, and they can be low, moderate, or high (Heubert & Hauser, 1999). For students, tests results may keep them from graduating or being promoted to the next grade level. Teachers may receive monetary rewards for good student performance on the tests. Schools may also receive funds or have governmental funds withdrawn based on student performance. Currently, standardized tests are used for accountability within the school. These tests are administered to students at specific grade levels. Test standards are set to measure student outcomes and when students do not meet these standards there are consequences

for all involved, particularly the teachers and schools (Casbarro, 2005). It is argued by some that these tests are more about evaluating the schools than they are with evaluating student progress (Casbarro).

Evaluating schools based on student outcomes on standardized tests suggests that these tests measure how well the teacher and school have prepared the student. By holding the school accountable for student performance on standardized tests the assumption is adopted that if the student performs well then the school must be successful. This assumption may suggest to some that teachers are the primary source responsible for student outcomes and that all students come to school equally ready to learn (Casbarro, 2005). The creators of the NCLB Act appear to take on this view because of the expectation that all students and schools are to meet the same standards regardless of student characteristics. Is this really a fair assumption? Take for example the teacher who is instructing an entire classroom of English language learners; should this teacher be held entirely responsible for the outcome of his/her students? Because such high consequences for the teachers and schools exist in some systems, it opens the door for cheating and turning classrooms into test preparation rooms rather than educational environments.

A study conducted by Herman & Golan (1993) found that teachers feel pressure to improve student test scores. Because of this pressure, teachers reported that they adjust their plan for instruction based on tests and spend more time on test preparation activities than instructional activities. The teachers in this study also reported that testing creates substantial tension for both teachers and students. This study supports the assumption that high-stakes aligned with test results can change the environment of a classroom from a learning facility to a test preparation facility. This may negate in part

one of the main goals of accountability: improvement of instruction and student learning (Lane, 2004).

Other debates have developed out of the standardized assessments movement. One of these debates is concerned with whether or not assessments should drive school reform (Wang, Beckett, & Brown, 2006). Many of the most recent educational reforms have made claims that the U.S. is being undermined by poor educational outcomes (Sykes, 1995). On one side of the debate, scholars believe that this tragedy of education in the U.S. is a "manufactured crisis" which has been predicated on false data (Berliner & Biddle, 1995). These scholars have focused on student improvements in areas such as a decrease in dropout rates and increases in SAT scores (Wang et al., 2006). However, those on the other side of the debate place great emphasis on national rankings and believe that this assessment driven reform will have positive effects on curriculum and instruction and improve U.S. rankings in the national arena of education.

Another debate regarding high-stakes testing is about standards-based assessment. Standards-based assessment refers to measuring student outcomes based on set standards rather than by comparison to other norm groups. One concern with standards-based assessment is whether or not students should be expected to meet the same set of standards regardless of socioeconomic status, race, or disability (Wang et al., 2006). Those in favor of standards-based assessment believe that traditional testing does not allow for a common goal. With standards-based assessment, ambiguity in the goal is potentially removed and a common goal is set for all (Schiller, 2000). Those opposed to standards-based assessment feel that these tests create "one size fits all" criteria. Many students are not going to fit these criteria, but will still be expected to meet the standards to be successful.

Yet another highly debated issue is about assessment-centered accountability that has been put in place. Assessment-centered accountability states that students, teachers, and schools are all held accountable for student outcomes on standardized tests (Wang et al., 2006). Proponents of this form of accountability believe that those involved in student education should be held accountable for the children's performance. Having an assessment-centered accountability in place makes this possible. Also, the school's performance can be measured and monitored this way. Some arguments against assessment-centered accountability are that just one test should not be used to measure student learning, test scores are strongly affected by socio-economic status, one test can not be expected to serve a multitude of purposes, and that this type of accountability does not allow for different kinds of instruction for different kinds of students (Koretz, 1995; Popham, 1999).

At present high-stakes testing is a part of today's educational system regardless of the controversies surrounding it. Educators, scholars, and politicians need to work together to create a cohesive system so that positive results will come of the new standards. The main goal of large-scale assessments is to promote student learning for all students.

High-Stakes Testing and Students with Disabilities

The amendments of IDEA 97 and NCLB have initiated reforms have already been extensively implemented. Now more than ever, *all* students are being included in standardized assessments. However, little is known about the effects of including students with disabilities in these same reform efforts (McDonnell, McLaughlin, & Morison, 1997; Schulte, Villwock, Whichard, & Stallings, 2001). Because students with

disabilities have been excluded for many years the effects of not including them are well known.

When students are excluded from standardized assessments because of disabilities there are concerns of inappropriate referrals to special education and rates of retention in grades prior to testing (Thurlow & Johnson, 2000). Students who are not included in assessments may also be instructed differently. Some schools ask that students not involved in testing stay at home for the day of testing or that those students go on a field trip (Thurlow & Johnson). This creates issues that these students are missing instruction simply because they are not included in the assessments. Another unintended consequence of excluding students from assessments is that they suffer from lowered expectations and access to the general education curriculum. However, even though there are complications with excluding students from assessments there are also intended and unintended consequences from including students in assessments.

"Test results, either favorable or unfavorable, are designed to have an effect on the content in focus as a curriculum, instructional strategies, intervention strategies to improve the learning of all students, professional development support for teachers and administrators, the use of assessment results, and the use and nature of test preparation materials" (Thurlow & Johnson, 2000, p. 307). These are just some of the intended consequences of including all students in assessments and using test results to examine performance. Conversely, there are also some unintended consequences to including students with disabilities in standardized testing. Some of the unintended consequences are increased referrals to special education, lowered expectations, narrowing the curriculum and instruction simply to test content, teaching to tests, limiting the range of

program options, and using test results for graduation/promotion decisions (Thurlow & Johnson).

Even though students with disabilities are required to take these same assessments, most are allowed accommodations or alternate tests. An accommodation refers to any changes made to the test that still allow for the measure of the student's abilities rather than their disability (Washburn-Moses, 2003). Accommodations can come in various forms ranging from the way the test is presented to the timing/scheduling of the test. Appropriate accommodations should be chosen with the purpose of allowing the student to achieve valid scores, rather than optimal scores (Fuchs and Fuchs, 1999). For example, students with high-incidence disabilities (e.g. learning disorders and speech disorders) can be part of the standard test administration while other students with disabilities may receive accommodations on tests, while a smaller group may receive a modified standard test or take an alternate test to ensure that the test results lead to valid decisions (Hollenbeck, Tindal, & Almond, 1998). Accommodations are important for both students with disabilities and the schools, yet controversy surrounds the use of certain accommodations. States and districts have different policies regarding the use of accommodations and they can be very complex (Thurlow, House, Scott, & Ysseldyke, 2000).

Alternate assessments are measures that are given to students unable to take state and district assessments. Alternate assessments are fairly new and about only 2% of the total student population take an alternate assessment (Thurlow & Johnson, 2000). Alternate assessments are still being made by most states and districts. These assessments can range from being different versions of the paper-and-pencil test to portfolio tests (Thurlow & Johnson).

Statutes increasingly require that schools include nearly all students in state and district wide assessments, but are they? Previous studies have found that participation rates vary markedly in large scale testing programs (Elliott, Erickson, Thurlow, & Shriner, 2000; Heubert & Hauser, 1999; Zlatos, 1994). Schulte, Villwock, Whichard, and Stallings (2001) found that the number of students who participated in a testing program in one district in North Carolina increased by 11% over their five-years of data collection. It is interesting to note that this same study also found an increase in the number of students with learning disabilities who scored at or above the grade level proficiency standards in reading. These results are promising and align with the intent of IDEA 97 to improve outcomes for students with disabilities on large scale assessments. However, it is important to keep in mind that this study was conducted with only one school district and is very limited in its generalizability.

Since schools are being held accountable for assessment outcomes for all students it is important that each student receive an equal opportunity for success. The majority of learning disabled students are included in the general education classroom (U.S. Department of Education, 2000), however, many other disabled students never attend general education classrooms, even though all students who participate in high-stakes testing are expected to have access to the general education curriculum (Washburn-Moses, 2003). This is important because many general education teachers are "teaching to the test." These teachers are only teaching material relevant to the test and also teaching highly specific test techniques. Students with disabilities may benefit and perform better on high-stakes tests if allowed to participate in the general education.

Mainstreaming and Inclusion

IDEA 97 requires that students must be taught in the "least restrictive environments" (LRE). The least restrictive environment is usually considered to be the general education classroom because this is where they would be placed if they did not have a disability. "The LRE concept is based on a 'Cascade of special education services' paradigm presented by Reynolds (1962) and Deno (1970)" (Junkala & Mooney, 1986, p. 218). Based on this paradigm, students are only kept out of the regular classroom in proportion to the support services that are required by their educational program (Junkala & Mooney). Two ways of including students in the LRE are mainstreaming and inclusion.

Mainstreaming and inclusion are often used interchangeably, but they are two different systems. Mainstreaming refers to keeping disabled students in the regular education classroom for certain classes and then separating them for others (Rogers, 1993). Under this system, generally, the student must be able to keep up with the work assigned to him or her with appropriate supports to stay in the classroom. Mainstreaming was popularized in part by a classic article written by Dunn in 1968. In this article, Dunn pointed out problems with separating disabled students from the general education classroom. This article also laid the groundwork for key components for the Education for All Handicapped Children Act (renamed the Individuals with Disabilities Education Act), which was passed only 7 years after the publication of Dunn's article (McLeskey, 2004).

On the other hand, inclusion refers to allowing the student to remain in the regular education classroom or the classroom they would have attended had they not been disabled as much as possible. The supports are brought to the child with the hope that the

student can attain benefits from remaining in the regular classroom (Rogers, 1993). Inclusion appears to be promising and there has been a large push for it in the United States since the early 1990's. An important aspect of inclusion is that students are not segregated from their nondisabled peers. Many classrooms are intended to be inclusive, but then simply end up replicating special education services in the general education classroom. In these contexts disabled students are commonly segregated within general education. "The ultimate goal of inclusion is to make an increasingly wider range of differences ordinary in a general education classroom" (McLeskey & Waldron, 2007, p. 163). McLeskey & Waldron (2007) discuss four ways in which this goal can be achieved.

The first way is to help create a classroom in which varying behaviors are more tolerable and become a routine part of the regular education classroom using supports. This will help the school community accept a larger variety of differences and become a common aspect of the school (McLeskey & Waldron, 2007). In any general education classroom students display a broad range of academic and social skills, which are considered typical and acceptable. Many teachers will arrange their classroom to accommodate these different ranges of student levels. However, there are some students who do not fall within this acceptable range and may require special services. So for inclusion to be successful, teachers need to expand their level of tolerance so that these students who fall outside of the acceptable range can remain in the classroom. This may also require changes in practices and resources available in general education to meet the needs of increasingly diverse students. Another issue is to keep classroom supports natural and unobtrusive. Evidence has indicated that these types of supports work best and appear to be used continually because they fit naturally into the flow of the general education classroom (McLeskey & Waldron, 2007; Gersten, Chard, & Baker, 2000;

Klinger, Arguelles, Hughes, & Vaughn, 2001). Natural and unobtrusive supports will also have less of a negative effect on those students with disabilities.

A third issue to consider when trying to maintain an inclusive classroom is that the rhythm of the classroom stays as typical as possible. That is, the school day for the disabled student should be as similar as possible to other students. Many problems arise for a student when moving from classroom to classroom fragments his or her school day. This also helps to ensure that the student's differences become ordinary and that they become part of the learning and social community (McLeskey & Waldron, 2007). Finally, all students must become part of the learning and social community of the classroom. Students with disabilities must be considered regular and included in the regular education classroom just as those students without disabilities. Many teachers will need to challenge their assumptions about disabled students to make this possible.

Successful inclusion can result in many benefits for students with disabilities (Begeny & Martens, 2007; Karagiannis, Stainback & Stainback, 1996). Some of the intended benefits of inclusion include improvements in academics, social skills, and being more apt to live in the community with little assistance (Begeny & Martens, 2007). Other benefits include improving teacher skills, helping students develop more positive attitudes towards others with disabilities, and establishing social principles based on equality (Begeny & Martens; Karagiannis et al., 1996).

With the potential for numerous benefits that may result from inclusion it may be hard to believe that there are those opposed to inclusion. However, there are some scholars and policy makers who argue against inclusion. Some of these arguments include believing that general education is not ready or prepared for inclusion and that full inclusion cannot be accomplished because it is too complex for the general education

classroom. Along with this argument is the belief that students with disabilities require intensive interventions that cannot successfully be provided in the general education classroom. Finally a strong point made by those who resist inclusion is that there is not enough empirical evidence to validate the effectiveness of inclusion (Begeny & Martens, 2007). There may not be much evidence for the support of inclusion in the United States, but many supporters point to Italy as a model. Italy initiated inclusion, albeit with some resistance, starting as early as the 1960's (Begeny & Marten). Current research and reports in Italy suggest that inclusion in the schools is now widely supported (Balboni & Pedrabissi, 2000). Some United Stated educators have tried to use the Italian model to highlight the benefits of inclusion, such as the positive attitudes that educators and teachers have towards inclusion (Begeny & Martens, 2007).

Even though there is resistance, there are schools in the United States that are promoting inclusion. One study conducted by Idol (2006) investigated a few of these schools and the outcomes of inclusion. This study evaluated the effects of including students with disabilities in the general education classroom. Idol (2006) looked at administrative support and attitudes toward inclusion, teacher attitudes toward inclusion, and also the academic impacts of inclusion. Overall, the general findings suggested that the administrators at the schools involved in the study were supporting their teachers and were positive about their inclusion program. Most teachers supported inclusion and thought that the impact of having disabled students with nondisabled students in the classroom was favorable. It was also found, that with these schools, including disabled students in statewide testing did not appear to be deleterious to the test performance of the general education students. This study suggests that inclusion can be accomplished in the United States with positive outcomes.

Some may ask why should schools turn to inclusion when we already have special education programs in place? One way to answer this question is how Brucker (1994) stated it, "As much as it may hurt to admit it, we have been generally unsuccessful in our current mode of service delivery, although we have had some individual successes. The operation may have been a success, but the patient died!" (Brucker, 1994, p. 582). Here are some harsh statistics that reflect the failure of our current system of special education as reported by Brucker (1994):

> "The 14th Annual Report to Congress (U.S. Department of Education, 1992) indicated that only 57% of these students graduate with a diploma or certificate of graduation. A 1989 study by Gartner and Lipsky of 26 cities revealed that less than 5% of all students, including students with learning disabilities, leave special education once they are identified. The National Longitudinal Transition Study results, as cited in NASBE (1992), revealed that only 49% of these students ages 15 to 20 are employed 2 years after graduation, and only 13.4% are living independently."

These are not the outcome that the framers and providers of special education would have hoped for. As concluded by the National Association of State Boards of Education (1992) Study Group on Special Education these results are mainly due to the unnecessary segregation and labeling of children with disabilities. Also, the ineffective practice of mainstreaming has had negative effects on students both academically and socially (Brucker, 1994). More research is needed in order to determine the full benefits of inclusion over the current special education programs in place. Some of the social benefits of inclusion for disabled students are known and empirically supported (Begeny

& Martens, 2007; Karagiannis et al., 1996), but investigating the academic benefits of inclusion will be important for future academic success of disabled students.

Teacher Effects

One major factor to consider when evaluating student performance on high-stakes tests, disabled or not, is the teacher. Many scholars would argue that teaching does matter (McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Noell & Burns, 2006). Now that students with disabilities are also being included in high-stakes testing the same pressure is put on the teachers instructing these students.

Studies examining teacher effectiveness are on the rise. Recent research on teacher effectiveness has shown a direct relationship between its quality and student learning (Darling-Hammond & Young, 2002; Ding & Sherman, 2006). When studying teacher effectiveness it is important to distinguish between teacher effectiveness and teacher effects. Teacher effects are those factors that are quantified in research studies. Odden, Borman, and Fermanich (2004) identified some of these teacher effects from reviewing the literature. Specifically, they identified teacher effects associated with student achievement, and these include: (1) years of teaching, (2) major of undergraduate study, particularly for mathematics and science, (3) ACT or SAT test scores, (4) course work or degree obtained, (5) quality of high school, (6) earning of a license, and (7) verbal ability (Ding & Sherman, 2006).

Teacher effectiveness is more difficult to define. Typically, teacher effectiveness is operationalized in terms of teacher effects. For example, many standards classify a teacher as being effective if they are fully licensed or in terms of specific teaching practices. However, there is danger in equating teacher effectiveness with teacher effects. Not all teachers who are fully licensed will be effective, just as some teachers who are

not fully licensed may be extremely effective. Research investigating teacher effectiveness must clearly articulate what is meant by teacher effectiveness just as teacher effects are operationally defined (Ding & Sherman).

There is no doubt that not all teachers have the same effectiveness. Evidence on teacher effectiveness is mixed (Nye, Konstantopoulos, Hedges, 2004). One must be cautious when interpreting findings from studies of teacher effectiveness. It is important to keep in mind that student background characteristics must be accounted for in the analyses because students are not randomly assigned to schools (Nye et al.). Teachers are assigned the students that are placed in their classrooms, not who they choose to teach. Some teachers receive classrooms full of gifted students while others receive classrooms of low SES students who are disabled. These factors need to be taken into account when looking at how effective a teacher is.

Studies conducted by Darling-Hammond (2000) concluded that teacher effects on student performance outweighs student characteristics such as poverty, language background, and minority status. These conclusions were drawn from using surveys from a 50-state survey of policies, by examining state case studies, using the 1993-1994 Schools and Staffing Survey (SASS), and information drawn from the National Assessment of Education Progress (NAEP) (Darling-Hammond). The authors of the study used the information to evaluate how teacher qualifications and other school characteristics affect student achievement. Similar conclusions were drawn in a study conducted by Rivkin, Hanushek, & Kain (1998) about teacher effectiveness. In fact, in their study, they claim that teacher quality is the most important predictor of student achievement. The authors of this study used panel data containing student, teacher, and school data from the state of Texas with many observations. A value-added model was

used and analyses run to investigate teacher effects. These studies draw on the importance of determining what makes a teacher highly qualified or effective.

So, the question of critical importance becomes what makes an effective teacher? Do teachers need to be certified in the content area they are teaching, do they need to be certified to teach special education if instructing students with disabilities, or do they need simply to be instructed in teaching practices? According to the United States Department of Education (USDOE), teacher preparation is of little importance for enhancing student achievement (Boe, Shin, & Cook, 2007). According to a report from the USDOE (2002), student achievement was not improved by teachers attending traditional schools of education or teacher certification. The USDOE stated that "the best available research shows that solid verbal ability and content knowledge are what matters most" when it comes to highly qualified teachers (2002, p. 9).

However, when it comes to the USDOE's interpretation of the NCLB definition of a highly qualified teacher (HQT) the focus is on content knowledge, which is emphasized through teacher preparation and professional development (Boe et al., 2007). This view is contradictory to the view that preparation does not matter. In fact, the USDOE concluded that schools are failing to produce highly qualified teachers that are needed by the NCLB Act (2002). They also go on to state that this failure is due in part to the "burdensome requirements" of a "shocking number of education courses" (p. 31). The USDOE policy argues in part that the best way to produce highly qualified teachers is to have a "fast track" alternative teacher preparation program that leads to certification and shortens the coursework that is taken (Boe et al.).

The USDOE's view of a HQT is at least partially self-contradictory. While the USDOE focuses on only the content knowledge requirement of an HQT, NCLB also

requires that teachers have a full state certification, a high level of content knowledge, and must earn at least a bachelor's degree (Boe et al., 2007). The requirement of having a full state certification typically entails extensive coursework in pedagogy and practice teaching. Because of these contradictions in federal definitions, teacher preparation programs are confused as to how to produce HQT. Research needs to investigate teacher qualifications and characteristics that are the most effective. There are currently studies that have evaluated these questions, but there is little consensus on the results (Wayne & Youngs, 2003).

A study conducted by Boe, Shinn, and Cook (2007) examined the effects of teacher preparation in producing highly qualified teachers. This study was a direct result of the contradictory reports put out by the USDOE. The study investigated relationships between the amount of teacher preparation and several teacher qualifications. The authors used national data produced by the 1999-2000 Schools and Staffing Survey (SASS). The authors focused on both teachers of general education and special education (Boe et al.). The results of this study suggest that for both special education and general education teachers, those with extensive preparation in pedagogy and practice teaching earn a higher level of certification. This means that teachers with higher levels of certifications meet the basic requirement of the NCLB, to have a full state certification. Results also showed that general education teachers with more preparation were more likely to be teaching in the field of their subject matter expertise. Also, those teachers with more preparation in training felt like they were better prepared to teach assigned subject matter (Boe et al.). Thus, according to the definition of a HQT by the NCLB, those with more extensive preparation were able to meet these standards than those with little or no preparation. Boe et al. believe that requiring less instruction in pedagogy and

also little time practicing and being supervised teaching could potentially lower the national number of HQTs, which is directly the opposite effect the USDOE is trying to achieve. Having an extensive training background appears to be a characteristic of highly qualified teachers that should not be overlooked.

Similar to teacher preparation studies, Wayne and Youngs (2003) reviewed an extensive body of literature that examined the relationship between student achievement gains and teacher characteristics. Specifically, the authors looked at college ratings (simply refers to how well rated the undergraduate college attended was ranked), tests scores, degrees and coursework, and certification status. Only 21 studies met criteria for inclusion in the review. The findings suggest that there appears to a positive relationship between college ratings and student performance. However, in these studies, a relationship between these two were not always found, but when there was a relationship it tended to be positive (Wayne & Youngs).

Other investigations into what makes for highly qualified or effective teachers examined different variables. Some research has examined the impact of having a subject-specific, advanced degree and the affect it has on a student's performance. Goldhaber & Brewer (1997, 2000) investigated the effects of teachers holding masters degrees on high school students' mathematics achievement. According to their findings, the area in which the degree was awarded is very important. They found that students who were instructed by a teacher with an advanced degree in mathematics made greater performance gains than did those students with teachers with no advanced degree or a degree earned in another subject. Another study conducted by Goldhaber & Brewer (1998) confirmed their initial findings that, in particular, for mathematics the degree earned by the teacher has an effect on student performance. Subject-specific degrees

earned by teachers had a more positive effect on student achievement outcomes regardless of whether it was a bachelor's or master's degree.

Although these findings are positive in suggesting what makes an effective teacher, most of these studies have been conducted with high school teachers and students. Croninger, Rice, Rathbun, and Nishio (2007) conducted a study investigating teacher qualifications that have an impact on elementary student performance. In this study the authors look at specific teacher qualifications such as teacher certification status, degree attainment, degree program, subject-specific coursework in reading and mathematics, and years of first-grade teaching experience (Croninger et al.). The findings of this study suggest that certain teacher qualifications matter. Teacher experience, coursework taken in preparation for the profession, and the specific type of degree earned all had a positive effect on reading achievement. In contrast, qualifications such as certification status and possession of advanced degree were not found to be significantly related to reading achievement. "An important implication of our findings, though, is that teacher qualifications may influence student achievement through effects associated with individual teacher characteristics or through the effects associated with collective teacher characteristics" (Croninger et al., 2007, p. 321).

Teacher quality is important and finding what affects teacher quality should be at the top of the educational research agenda. Strauss and Sawyer (1968) used cross-sectional data and examined the impact that teacher quality had on students standardized test scores. The study specifically looked at mean achievement of students and the number of students who failed the standardized tests. Based on this study an increase in teacher quality results in a decline in the rate of student failure. Their results concluded that just an increase of 1% in teacher quality, as measured by a teacher evaluation score,

results in a 5% decline in the rate of student failure on standardized test scores (Strauss & Sawyer). It is important to keep in mind that this study is very dated. Much has changed in the educational world since the results of this study were published. The impact our teachers are having on our students is critical and thus the qualities a teacher needs should be investigated to improve our education system.

It is also important to keep in mind that research has shown that teacher effects are cumulative over time. "Groups of students with comparable abilities and initial achievement levels may have vastly different academic outcomes as a result of the sequence of teachers to which they are assigned. These analyses also suggest that the teacher effects are both additive and cumulative with little evidence of compensatory effects of more effective teachers in later grades" (Sanders, & Rivers, 1996, p. 6). Not only does this research suggest that teachers play a major role in student achievement over time, it may also lead to important implications for student assignment. For example, in a study conducted by Sanders and Rivers (1996) found that lower achieving students benefit first from an increase in teacher effectiveness. Similarly, teachers who are found to be most effective result in appropriate to excellent gains for all students. If characteristics can be identified as to what specifically makes an effect teacher, students could benefit from being placed in these teachers' classrooms if they had been previously placed in classrooms with less effective teachers.

The studies discussed above and others indicate the importance of teacher qualifications for students in high school and elementary grades, but what about special education students? To date, little research has been conducted regarding teacher effectiveness with special education students. Students with disabilities are in need of more intensive services and thus may require different kinds of teachers than non-

disabled students. General education and special education teachers may require different qualifications. In fact, some programs that are used in the general education and taught to these teachers are hardly recognized or used within special education. Just as there are some programs used with exceptional general education programs there are also some unique features of special education programs that are not introduced to general education programs (Brownell, Ross, Colon, & McCallum, 2005). Because of these facts, more research should be conducted to determine teacher qualifications that are the most effective and lead to positive performances for special education students.

Value-Added Models

Value-added modeling (VAM) is a promising technique to answer some of these teacher effectiveness questions. VAM is a "collection of complex statistical techniques that use multiple years of students' test score data to estimate the effects of individual schools or teachers" (McCaffrey, Lockwood, Koretz, & Hamilton, 2003, p.xi). VAMs are being used more and more to examine teacher and school effectiveness (Noell & Burns, 2006). "VAMs have been employed extensively in domains such as biology or economics to model growth, change, or production for complex systems in which a number of variables may moderate change and in which change may be nested within specific nonequivalent units" (Noell & Burns, 2006, p. 41). These models can be effective for evaluating education because the modeling approach closely resembles the structure of education.

Value-added models have two key features when examining student success. The first feature is that the models allow for dependent variables in the analysis to be created to assess the amount of change in student achievement that occurs over the year that the student is in the classroom under study, although this is not required for VAM. The

second feature is that variables such as students' prior achievement in other classrooms, student demographic variables, and also social composition variables of the schools the student attended can be adjusted for in the dependent variable when used in the model (Rowan, Correnti, & Miller, 2002). Value-added models estimate the proportion of variance changes in student achievement lying among classrooms or schools, when controlling for the effects of confounding variables.

Policy makers and researchers alike have taken a great interest in VAMs. Policy makers see VAMs as a window for educational reform through improved teacher evaluations or as part of a test-based accountability program (McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Of particular interest is the fact that because of the complex statistical techniques offered by VAM teacher and school effects can be evaluated with less distraction by other confounding variables (i.e. family background (McCaffrey et al., 2003).

These models are relatively new methods for evaluating teachers and schools within psychology, but seem to be on the rise. A few states, Louisiana being one of them, are currently using VAMs to investigate teacher preparation programs. One reason that the use of value-added models is becoming so popular is because of the availability of administrative databases that track individual student achievement over time (Harris & Sass, 2006). Many states are now creating longitudinal databases to measure student achievement. By keeping longitudinal databases that measures change at the individual level the influences of student and family characteristics can be controlled for when evaluating educational programs. With new federal mandates and accountability standards, within a few years, most states should have the capability to track student achievement over time (Harris & Sass).

Along with more access to longitudinal databases, McCaffrey, Lockwood, Koretz, and Hamilton (2003) have cited at least two other reasons why the use of VAMs are on the rise.

> "One reason is that VAM holds out the promise of separating the effects of the teacher and schools from powerful effects of such noneducational factors as family background, and this isolation of the effects of teachers and schools is critical for accountability systems to work as intended. The second is that early VAM studies purport to show very large differences in effectiveness among teachers. If these differences can be substantiated and causally linked to specific characteristics of teachers, the potential for improvement of education could be great" (p. xi).

Although VAM appears to be very promising it does have some shortcomings when investigating teacher effects. One drawback is known as estimating the counterfactual. This refers to estimating what would have happened to a student's achievement score under different circumstances while isolating the teacher's effects (Murnane & Steele, 2007). The specification of the meaningful counterfactual is not always clear and it is not necessarily the case that the teacher's effect would the same with different students.

A second shortcoming of VAM is that because there are correlations in students test scores from one year to the next, models need to be specified as to take these correlations into account (Harris & Sass, 2006). These models require that statistical assumptions be met about teacher effects over time. However, using different assumptions can create different estimates of teacher effectiveness.

Third, VAMs must take into account the relatively small number of students that teachers work with (Murnane & Steele). Each individual student has more of an impact on a teacher's perceived effectiveness when the teacher has a small number of students. This may cause estimates to be less reliable for teachers with a small number of students. There are statistical procedures to minimize this effect, but these also have the disadvantage of potentially under estimating the impact of the strongest and weakest teachers.

A fourth challenge lies in the fact that teachers do not choose the students that are in their classrooms and student assignment is not random. This makes it difficult to discriminate between contextual effects, such as school and classroom characteristics, from teacher effects. Teacher effectiveness may be affected by many variables and thus it may be hard to determine effects that are due strictly to teachers and those that are due to the classroom, school, or district environment characteristics (McCaffrey et al., 2003).

Having missing data, either student test scores or links between students and their teachers, creates a fifth problem. With missing data, it is impossible to tell if these students are systematically different than their peers and this may create biased estimates of teacher effects (McCaffrey et al., 2003). Finally, a sixth challenge deals with the achievement test used to measure student performance. Achievement gains are measured from one test administration to the next. Because of this it is important that the test measures content that the teachers have taught, that the scores are measured on the same scale, and that the tests measure comparable content. This can be difficult with tests administered in upper grades when different curriculums are taught in different classrooms. Despite their limitations, applications of VAM to education provide

powerful models that fit the natural structure of educational data and whose most important limitations can be overcome in some applications (Murnane & Steele, 2007).

A study conducted by Harris and Sass (2006) evaluated model assumptions and their impact on estimates of teacher quality when using VAM. The authors noted that past studies on value-added modeling "has been significantly hampered by data limitations, which, in turn, has forced researchers to estimate mis-specified models" (p. 27). According to Harris and Sass (2006) very few authors who use value-added models test for the assumptions underlying the models. The authors investigated factors that were consistently impacting the effect teachers had on student achievement. The authors also looked at the effect prior knowledge and educational inputs had on student achievement, the measurement of schooling inputs that affect student achievement, and alternative methods for controlling student and family characteristics. From the findings the authors conclude that covariates are inadequate replacements for individual student and teacher effects, random effects models yield inconsistent estimates of model parameters, and that individual school effects should not be excluded (Harris & Sass, 2006). One large limitation to note with this study is that this study used data from standardized tests that are vertically aligned. Test data for this study were taken from Florida in which the tests are aligned and thus can be easily compared from year to year. However, many states use tests that are not vertically aligned across years. When tests do not vertically align this needs to be taken into consideration otherwise the results of the analyses will be meaningless.

A study conducted by Ballou, Sanders, and Wright (2004) investigated the Tennessee Value-Added Assessment System (TVASS) and how it measures teacher effectiveness. The TVASS measures teacher effectiveness based on student progress.

One of the main arguments against this system was that certain student level covariates were not included or accounted for, specifically socio-economic status and demographic covariates. The authors of this study ran analyses including these covariates to determine whether or not they had a large impact on the outcomes. What the authors found was that adding these covariates into the models did not have a significant impact on the results. The authors concluded that including SES and student demographic covariates in the TVASS was not necessary. Again, like the previous study discussed, the test scores used to measure student progress are vertically aligned. Because these tests are aligned the student's test history performance may be a substitute for these covariates that are left out of the model. The current study, unlike the study conducted by Ballou et al. (2004) and Harris and Sass (2003) does not have access to test scores that are vertically aligned and thus these covariates are imperative to the model.

The selection of the statistical model used in completing a value-added analysis has been a source of considerable debate. Tekwe and colleagues (2004) conducted an important study examining the performance of several models. These authors investigated different types of models used in VAM and the impact each model has when used in VAM studies. Specifically the authors looked at hierarchical linear models (HLM), layered mixed effects model (LMEM), and simple fixed effects model (SFEM). The authors used the same data set with each type of model. The data set contained two years of test scores for students. What the authors found was that there was little difference between using HLM, LMEM, and SFEM using only two years of data. The SFEM is less complex than LMEM and so it would be recommended over LMEM with only two years of data. The results of comparing SFEM and HLM showed that they were

also similar. However, if demographic variables are used the results are different. HLM is recommended over SFEM when including demographic variables.

The results show that each model has its hypothetical advantages for different data and analyses. HLM is highly recommended when using demographic variables and only a few years of data. SFEM is recommended over the more complex LMEM with only a few years of data and when demographic variables are not included. LMEM is recommended when change scores can be used and possibly when more years of data are used. Each model has its hypothetical advantages and should be investigated before choosing a model.

A study conducted by Lockwood and McCaffrey (2007) examined the difference between using fixed effects and random effects models (mixed model). Random effects models have been criticized for treating individual heterogeneity as part of the models error term. Critics believe that this leads to biased parameter estimates, which might be alleviated with some forms of fixed effects models. The authors found that the random effects models have a bias compression feature when a large number of correlated observations are used. These results showed that a mixed model approach may be beneficial when using longitudinal achievement data and that fixed and random effects models may converge when a considerable amount of data is available.

Value-added models can be very useful when looking at educational issues. Questions regarding teacher effectiveness have emerged into the spotlight with new accountability programs. Previous VAM studies have suggested that teachers are important when considering student outcomes because they are a source of variance (McCaffrey et al., 2003). Value-added model research shows promise for the future of educational reform.

RATIONALE AND PURPOSE OF CURRENT STUDY

The purpose of the current study was to investigate variables that affect special education students' performance on high-stakes testing in the state of Louisiana. Now that students with disabilities are required to be included in statewide testing, they should be given the same chance of success as general education students. Many factors are suggested to affect student performance on these high-stakes tests. Many studies have been conducted investigating the performance of students in the general education classroom, but little has been done with special education students.

The current study examined specific variables to determine their affect on special education students' high-stakes testing performance. It was investigated whether or not those students with disabilities who were taught within the general education classroom performed at higher levels than their peers who were segregated into special education classrooms. Studies have shown that inclusion has social benefits for both disabled and nondisabled students, and this study hoped to add to the literature by determining if there were also academic benefits involved with inclusion.

Other variables that were examined were related to teacher qualifications and effectiveness. Certification in special education and specific content areas were investigated to see whether or not these qualifications affect student performance on high-stakes tests. It was investigated as to whether students with disabilities scored higher on tests if instructed by a teacher with a special education certification. Also, the type of certification the teacher holds was investigated. Did the teacher have an alternative certification or a regular certification? Lastly, whether or not the teacher had been trained and certified to teach specific content areas was examined.

METHODS

Participants

Analyses for the current study began with the use of a large pre-existing multivariate longitudinal database with modifications being made specific to the current study. The data used to construct this database were obtained from the Louisiana Department of Education.

Data were analyzed for students enrolled in grades 4 through 9 for the school years of 2004-2005 and 2005-2006. These grades were selected to permit the availability of one-year prior achievement data (grade 3).

Additionally, students who were retained at the end of the prior school year and students who moved during the school year were dropped from the dataset. Students who were retained were dropped because the meaning of assessment data for students who are repeating the same grade is different from students who were promoted. Students who moved were dropped due to the fact that student-teacher-course nexus data were only collected once per year, thus, once a student moves it is not possible to ascribe subsequent instruction to a particular teacher (Noell, Porter, & Patt, 2007).

Measures

The Louisiana Educational Assessment Program for the 21st Century (LEAP-21) and the Integrated Louisiana Educational Assessment Program (*i*LEAP) are given to students in the state of Louisiana to measure how well the student has mastered the state content standards. The LEAP-21 is a criterion-referenced test that was initiated in 1997 to align with new content standards (Mitzel & Borden, 2000). Students are tested in English language arts, mathematics, science, and social studies. The LEAP-21 test is given to students in the fourth and eighth grades. The LEAP-21 test is validated based on content

validity. Content validity is verified by a content review committee to determine whether the test aligns with state standards. Thus, content validity is considered to be built into the test during development. Reliability for the LEAP-21 was assessed using a traditional, Cronbach's alpha, and ranges from .87 to .94 (Louisiana Department of Education, 2006a). Reliability coefficients above .85 are considered excellent, and thus the LEAP-21 has excellent reliability (Louisiana Department of Education, 2006a). More detailed reliability, validity, and test development data for the LEAP-21 are available at the Louisiana Department of Education website located at: http://www.doe.state.la.us/lde/saa/2273.html and in Mitzel and Borden (2000).

The *i*LEAP was initiated in the spring of 2006 to fulfill the *No Child Left Behind Act* (NCLB) standards. Students are tested in English language arts, mathematics, science, and social studies. The *i*LEAP is given to students in the third, fifth, sixth, seventh, and ninth grades. Prior to the *i*LEAP the Iowa Tests of Basic Skills (*ITBS*) was administered. The *ITBS* tests are norm-referenced tests, which were not aligned to state content standards, and thus the *i*LEAP was developed to take its place. "The *i*LEAP includes a subset of *ITBS* items that make up a bulk of that scale, plus some augmentation for Louisiana. By making this change in assessment standards, this should improve the content validity of the assessment by assuring tighter alignment between what is expected to be taught and what is assessed" (Noell et al., 2007). Validity for the *i*LEAP is considered built in to the test in the same way it is for the LEAP-21. Reliability, assessed with Cronbach's alpha, for the *i*LEAP ranges from .80 to .96 (Louisiana Department of Education, 2006b). Reliability coefficients above .80 are considered good while those above .85 are considered excellent (Louisiana Department of Education, 2006b). More detailed information regarding test development, reliability, and validity data for the

*i*LEAP can be found at the Louisiana Department of Education's website at: http://www.doe.state.la.us/lde/saa/2273.html.

The *ITBS* is a nationally standardized test of achievement. Its use was discontinued because it was not specifically aligned with Louisiana's grade level expectations. The *ITBS* tested students in English language arts, mathematics, science, and social studies. The *ITBS* was administered to students in the third, fifth, sixth, seventh, and ninth grades. The *ITBS* is considered valid and reliable. The internal consistency coefficients of the *ITBS* are high as assessed on Kuder-Richardson Formula 20 (KR20) ranging from the middle .80s to low .90s (Engelhard, 2004). Detailed technical data is available from Hoover, Hieronymus, Frisbie, and Dunbar (1996).

Students receive one of five achievement ratings ranging from "Unsatisfactory" to "Advanced." This score is then used to make high stakes decisions in the fourth and eight grades. This information along with identifying information about each student is retained in a large database by the Louisiana Department of Education. Along with this database, there is a database containing Louisiana teacher information and curriculum information.

The current study used students' scores on the English Language Arts (ELA), reading, and the mathematics domains from the test that was administered to them to evaluate the effects that teacher and classroom characteristics had on their scores. These domains were chosen because they are used for retention/promotion decision-making and as a result are arguably the most important domains.

For the school year 2004-2005, when *ITBS* was still being administered, it is important to note that the *ITBS* and LEAP-21 do not report scores on comparable scales.

Because of this, standard scores for each domain within each test and grade have already been converted to *z*-scores based on the students who contributed to the analyses.

Special Education status is classified into numerous primary disability categories. For the purposes of this study, only the categories of Specific Learning Disability (SLD), Mild Mental Retardation (MMR), Speech and Language Disability (SPLD), Emotional Disturbance (ED), and Other Health Impairments (OHI) were used because each of these categories contained at least one thousand cases in the data set. Students in the special education categories were examined along with students not identified as being special education students (see below).

Database Construction

The database links data points from Louisiana's student achievement, teacher, and curriculum databases. The student database included student demographic information, and testing information for each year (2004-2005 and 2005-2006). Student demographic information in the database included the student's race, gender, poverty level (as indicated by free/reduced lunch status), grade, gifted status, special education status, and information about what school the student attended.

Multiple teacher databases were used to obtain the teacher information. The teacher information included demographic variables, teacher degree information, and teacher certification information. Degree information contained the college program(s) of attendance, year of graduation, and degree(s) obtained from each university. The type of certification that each teacher holds was also obtained and used.

The curriculum database was used to obtain information regarding classes each student took and the teacher who instructed the course.

Preliminary work was conducted to resolve the issue of duplicate records and multiple partially complete records that described the same student. Following this work, the LEAP and *i*LEAP data files were merged followed by an additional round of duplication resolution. Students' data were linked across years (2004-2006) based upon unique matches on multiple identifiers used in each stage of the matching process. A five step matching process was used. The first match consisted of trying to match students on their 12-digit identification number, their last name, and gender. Students who did not match uniquely on this step were then matched on their identification number, gender, and birthday. Again, a number of students who did not match uniquely on this sequence were then matched on their last name, first name, gender, and birthday. The next step was to match any unmatched students on their identification number, last name, and birthday. Finally any unmatched students were matched using their identification number, last name, and first name. Those student records that did not uniquely match at any stage were retained as isolated records of student performance and were not used in the current analyses (Noell & Burns, 2006).

In addition to achievement data, a number of additional variables were gathered and/or computed from the available database. These were student free/reduced lunch status, gifted status, special education status, limited English proficiency status, gender, and minority status. The percentage of students identified as being disabled per classroom was measured to use in the analysis. Classroom type (special education versus general education) was viewed as a continuous variable in the model.

From the preliminary data base construction, most students (93.2%) for whom assessment data were available for the school year 2005-2006 were matched with assessment records from the school year 2004-2005 (Noell et al., 2007).

Procedure

The data were analyzed using a hierarchical linear model (HLM).

> "HLM or mixed linear models have several important advantages over traditional analytic approaches. First, they readily capture the grouping of students within classrooms. Second, they permit appropriate modeling of variables at multiple levels such as student, teacher, and school. Third, they provide a model in which estimates of teacher effectiveness can be adjusted to account for unreliability of estimates" (Noell et al., 2007, p. 12).

The model that was used in the current analysis was a three-layered structure. Students were grouped within teachers' classes, who were in turn grouped within schools (see Figure 1). This three-layer model was chosen for several reasons. First, the school building level was used to account for the variance component at the school building level. Prior analyses have demonstrated that this effect may be small, but still important (Noell, 2006). The teacher level allowed for the analysis of various teacher characteristics that may have affected the student score. Finally, the student level containing student scores on high stakes testing was examined to see how it was affected by factors at level two.

Building the Models. The modeling approach for the current study followed similar procedures as in Noell (2006) and Noell, Porter, and Patt (2007). The same approach was used for ELA, reading, and mathematics. Error at each level (student, teacher, and school) was assumed to be normally distributed with a mean of 0 and common variance at that level. First, an initial 3 level model was specified in which

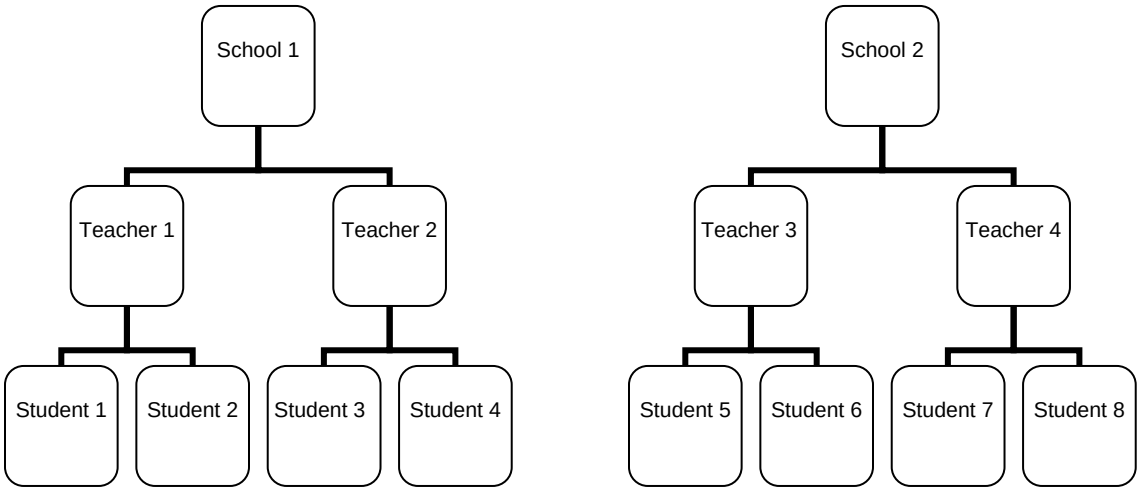

Figure 1: Nesting Structure of Students within Teachers and Teachers within Schools (Figure reprinted with permission from Noell, Porter, & Patt, 2007)

achievement was modeled with no prior predictors to use as a basis for comparison with more complex school) was assumed to be normally distributed with a mean of 0 and common variance at that level. First, an initial 3 level model was specified in which achievement was modeled with no prior predictors to use as a basis for comparison with more complex models. Next, prior achievement was added in blocks as fixed effects. Then, demographic variables were added as a block. Variables were removed one at a time in order of the lowest *t* value until only variables with significant effects, p = .01, were remaining. This same procedure was conducted for each level. The variables that were examined at each level are presented in the following tables.

A simplified presentation of the models that were used is provided below. Equations for intercepts and for the teacher level (Level 2) effects for variables modeling the impact of disability status are presented. In the equations presented below Σ is used to indicate summing across the p, q, and s coefficients at the student, teacher, and school

Table 1: Student Level Demographic Variables Examined

Variables
Gender (Male)
African American
Asian American
Hispanic
Native American
Receiving Free Lunch
Reduced Lunch
Gifted
Special Education
Section 504 Identification
Limited English Proficiency
Student attendance

Table 2: Classroom Level Variables Examined

Variables
Percentage of students who are male
Percentage of students who are minorities
Percentage of students who received free lunch
Percentage of students who received reduced priced lunch
Percentage of students who were identified as gifted
Percentage of students who exhibited limited English proficiency
Class mean prior achievement in ELA
Class mean prior achievement in mathematics
Teacher attendance

Table 3: School Level Variables Examined

Variables
Percentage of students who are male
Percentage of students who are minorities
Percentage of students who received free lunch
Percentage of students who received reduced priced lunch
Percentage of students who were identified as gifted
Percentage of students who exhibited limited English proficiency
Class mean prior achievement in ELA
Class mean prior achievement in mathematics

levels of the model respectively (Noell et al., 2007). The equation for the student level is broken in to two parts for presentation purposes only. For the actual equation, all of the special education categories coefficients were included with the student level predictor coefficients. Each model is only presented with Special Education category of Specific Learning Disability (SLD) as an example and for purposes of space. However, the model that was implemented contained each of the five Special Education categories under investigation.

Level 1: Students

$$Y_{ijk} = \pi_{0jk} + \sum(\pi_{pjk})a_{pijk} + \sum(\pi_{SLD \cdot jk})\, a_{SLD \cdot ijk} + e_{ijk}$$

where

Y_{ijk} is the achievement of student i in class j at school k in the target subject

π_{0jk} is the mean achievement for classroom j at school k

π_{pjk} are the *p* coefficients that weight the contribution of the student level data in the prediction of Y for $p = 1$ to the total number of coefficients

a_{pijk} are the student level data (prior achievement, demographic variables, and attendance) that predict achievement for $p = 1$ to the total number of data points for all variables other than special education disabilities

$\pi_{SLD \cdot jk}$ the coefficient for Specific Learning Disability summed across the j classrooms and k schools

$a_{SLD \cdot ijk}$ student level data indicating the presence of SLD

e_{ijk} the student level random effect, the deviation of the predicted score of student i in classroom j in school k from the obtained score

Level 2: Classrooms

$$\pi_{0jk} = \beta_{00k} + \sum(\beta_{q0k})X_{q0jk} + r_{0jk}$$

where

π_{0jk} is the mean achievement for classroom j at school k

β_{00k} is the mean achievement for school k

β_{q0k} are the q coefficients that weight the relationship between the classroom characteristics and π_{0jk}, $q = 1$ to the total number of coefficients

X_{q0jk} are the classroom level data that are used to predict achievement

r_{0jk} the classroom level random effect, the deviation of classroom jk's measured classroom mean from its predicted mean

$$\pi_{SLD \cdot jk} = \beta_{SLD \cdot 0k} + \sum(\beta_{q \cdot SLD \cdot k}) X_{q \cdot SLD \cdot jk} + r_{SLD \cdot jk}$$

where

$\pi_{SLD \cdot jk}$ is the mean achievement adjustment for SLD for classroom j at school k

$\beta_{SLD \cdot 0k}$ is the mean achievement for SLD students school k

$\beta_{q \cdot SLD \cdot k}$ is the type of certification (Special Education), classroom type (Special Education vs. General Education), teacher content domain certification, alternative certification or regular certification, and number of teacher years experience.

$X_{q \cdot SLD \cdot jk}$ are the classroom level data that are used to predict achievement

$r_{SLD \cdot jk}$ the classroom level random effect, the deviation of classroom jk's measured classroom mean from its predicted mean

Level 3: Schools

$$\beta_{00k} = \gamma_{000} + \sum(\gamma_{s00}) W_{s00k} + u_{00k}$$

where

β_{00k} is the mean achievement for school k

γ_{000} is the grand mean achievement in the target subject

γ_{s00} are the s coefficients that weight the relationship between the school characteristics and β_{00k} for s = 1 to the total number of coefficients

W_{s00k} are the school level data that are used to predict achievement

u_{00k} the school level random effect, the deviation of school k's measured classroom mean from its predicted mean

Analysis Plan

Once the final models for student achievement independent of the variables of interest in this study for ELA, reading, and mathematics were extracted, models were developed that examined the specific research questions targeted by this study. The coefficients were evaluated to determine the effects on student scores. A coefficient that was negative indicated a variable that was correlated with poorer test performance. Conversely, a positive coefficient indicated that the variable in question was correlated with improved test performance in an analytic context in which all of the other variables

are present. For example, if coefficient for Specific Learning Disability (SLD), which is generally negative, in the final model was -10.00 that would indicate that being identified as SLD was associated with a score that was 10 points lower than would be suggested by prior achievement and demographic variables.

Additional analyses were conducted examining the variables of interest for the current study, which were classroom type, teacher certification in special education, teacher certification in content area, teacher regular certification versus alternative certification, and teacher year's experience. These variables were included at Level 2 of the model. The variables were fixed effects added onto the five special education categories of interest in Level 1. The coefficients from these variables of interest were only relevant to the scores of students identified as exhibiting one of the five disability categories. Students were identified in the data under these categories using categorical codes. For example, there was a variable for Emotionally Disturbed in the database. If a student had been identified as being emotionally disturbed he or she had a 1 in this column in the data. If the coefficient for teacher certification in special education was found to be +5.00 at level 2 for emotional disturbance, that indicated that on average emotionally disturbed students who took the test and were taught by a teacher certified in special education, scored five points higher than students with similar prior achievement and demographic factors who were taught by a teacher who was not special education certified. A positive coefficient suggests a positive impact on the student's score and a negative coefficient represents a negative impact on the student's score.

Two sets of analyses were run. The analyses described above were first run with all students. A portion of the students had multiple teachers and their link to each teacher was weighted in proportion to their total number of teachers. For example, if a student

had two teachers each that student was weighted as .50 to account for the contribution of each teacher. The model was estimated using this weight for each case.

The second series of analyses were run using the model described above, but only retaining students with one teacher in a content domain to test for the importance of having a special education teacher without the confound of having multiple teachers who vary in certification. These students were divided into three categories based on prior achievement. The first category contained those students who scored one standard deviation or more above the mean in the content domain. The second category contained those students who scored within one standard deviation of the mean for the content domain. Finally, the last group consisted of students who scored at least one standard deviation below the mean for the content domain. The primary group of concern was the last group because these students were performing poorly in the specified content area. Analyses examined whether these students did better with a special education teacher instructing them relative to students who were taught by general education teachers. The three groups were divided up this way to specifically investigate the effects of special education teachers on students who were performing poorly in the target academic content. These groups allowed for the identification of students who were performing poorly in a specific content. Analyses examined whether these students performed better when instructed by a teacher who had a special education certificate, had a certificate to teach in that content domain, had an alternative certificate versus a regular certificate, has been teaching for a number of years, and what percentage of the classroom was special education students.

RESULTS

A series of ordinary least squares (OLS) regression analyses were conducted prior to the analyses of question for this study. These OLS analyses were conducted to examine general patterns in the data. The criterion variable was the current year achievement (2005 and 2006) with adding progressively larger blocks of predictor variables to examine the relationship. Test scores were standardized to a mean of zero and a unit standard deviation within grade and year. Demographic variables were entered as dummy codes. First and second order polynomial terms for prior achievement were examined and not found to be statistically significant. Similarly, a large family of demographic interaction terms was examined, with prior achievement and demographic factors included in the equations, and were not found to be statistically significant. As a result, polynomial predictors for prior achievement in the content area assessed and demographic interaction terms are not presented or discussed below.

To examine the predictive power of conceptually meaningful blocks of variables all variables were entered sequentially in blocks. The variables of interest were: prior achievement, demographic variables, and attendance data. Results for all content areas in each year are presented along with a brief description.

Across all content domains and both years prior year achievement in the content area is strongly related with current year achievement. The relationship was strongest for mathematics and weakest for writing. Including all four prior year achievement scores increased multiple *r* for an even stronger correlation. Interestingly, in all three contents across both years, adding a block of demographics increased multiple *r* by only a small increment ranging from +.006 to +.020. When including all prior year achievement and

Table 4*:* Reading Regression Analyses for 2005 & 2006

Predictors	Multiple correlation (Number of Students) 2005	Multiple correlation (Number of Students) 2006
Z-score Prior Year Reading	.753 (186,381)	.729 (139,950)
Z-scores Prior Year Achievement	.781 (186,381)	.765 (139,950)
Z-scores Prior Year Achievement & Demographic	.790 (186,381)	.775 (139,950)
Z-scores Prior Year Achievement Demographic & attendance	.791 (186,381)	.776 (139,950)
Z-score: Two Prior Year Reading	.801 (137,664)	.773 (101,490)
Z-scores: Two Prior Year Achievement	.815 (137,663)	.793 (101,447)
Z-scores Two Prior Year Achievement Student demographic factors	.820 (137,663)	.797 (101,447)
Z-scores Two Prior Year Achievement Demographic & attendance	.820 (137,663)	.797 (101,447)

Table Note. Prior Year achievement includes the *Z*-scores for reading, writing, and mathematics. Student demographic factors included were free lunch status, reduced price lunch, gifted status, primary special education diagnosis (codes for emotionally disturbed, specific learning disability, mild mental retardation, other health impaired, and speech/language concerns), limited English proficiency status, gender, Section 504 eligibility, and minority status (codes for Asian American, African American, Hispanic, and Native American). Only a combined free/reduced lunch status variable was available for 2004-2005. Attendance was the number of days the student was absent.

student demographics roughly 60% of the variance is accounted for in all content domains across both years. Adding another year of achievement data resulted in a modest increase for both years in mathematics, but only for the 2004-2005 year in writing and reading. There was a slight decrease in multiple r for the 2005-2006 year data in writing and reading. Again, adding two years prior year achievement in all contents resulted in an increase in the relationship. In all contents in both years the multiple r increased in small increments by adding demographic variables and attendance. The

final relationship between all variables and current year achievement is strong with a multiple *r* ranging from .797 to .824.

Table 5*:* Writing Regression Analyses for 2005 & 2006

Predictors	Multiple correlation (Number of Students) 2005	Multiple correlation (Number of Students) 2006
Z-score Prior Year Writing	.718 (252,330)	.710 (203,488)
Z-scores Prior Year Achievement	.749 (252,330)	.748 (203,488)
Z-scores Prior Year Achievement & Demographic	.762 (252,330)	.768 (203,488)
Z-scores Prior Year Achievement Demographic & attendance	.764 (252,330)	.771 (203,488)
Z-score: Two Prior Year Writing	.776 (197,216)	.769 (158,741)
Z-scores: Two Prior Year Achievement	.790 (197,215)	.787 (158,683)
Z-scores Two Prior Year Achievement Student demographic factors	.796 (197,215)	.798 (158,683)
Z-scores Two Prior Year Achievement Demographic & attendance	.797 (197,215)	.801 (158,683)

Table Note. All variables were entered as in Table 4, see the note above.

The final models for each content and year were specified by procedures described earlier. The values presented in the tables below were obtained before entering the variables of question for the study (these variables are presented later).

The coefficients are scaled to the approximate standard deviation of the educational assessments (*i*LEAP and *LEAP*) used in Louisiana: 50. It is also important to note that differences in how variables were scaled create the need for caution in comparing the coefficients across different types of predictors. Demographic variables at the student level were coded 1 if present and 0 if absent. Prior achievement is measured

in standard deviation units from the grand mean prior achievement. Classroom percentages are measured in 10% units, so that the value presented would be the expected change in students' scores if the percentage of the indicated group increased by ***10%***.

Table 6*:* Mathematics Regression Analyses for 2005 & 2006

Predictors	Multiple correlation (Number of Students) 2005	Multiple correlation (Number of Students) 2006
Z-score Prior Year Mathematics	.778 (219,816)	.768 (207,067)
Z-scores Prior Year Achievement	.799 (219,816)	.789 (207,067)
Z-scores Prior Year Achievement & Demographic	.805 (219,816)	.798 (207,067)
Z-scores Prior Year Achievement Demographic & attendance	.807 (219,816)	.800 (207,067)
Z-score: Two Prior Year ELA	.824 (167,795)	.809 (123,105)
Z-scores: Two Prior Year Achievement	.834 (167,795)	.817 (123,105)
Z-scores Two Prior Year Achievement Student demographic factors	.834 (167,795)	.822 (123,105)
Z-scores Two Prior Year Achievement Demographic & attendance	.835 (167,795)	.824 (123,105)

Table Note. All variables were entered as in Table 4, see the note above.

Due to differences in scales of measurement and the meaning of the measurements it is difficult o make direct comparisons across different types of measures (Noell, 2006; Noell, Porter, and Patt, 2007).

The variables are in order of the largest to smallest coefficient. As can be seen in the table above, prior year reading achievement is the largest contributor to a student's current achievement among the achievement variables. The coefficient for prior year reading achievement was more than three times the value of any other prior year content area coefficient.

Table 7: Hierarchical Linear Model for Reading Achievement 2005

Model Level	Variables Entered	Coefficient	(CI)
Student level variables	Prior Year Reading Test	19.4	(18.9, 19.9)
	Prior Year Science Test	6.9	(6.6, 7.2)
	Gifted	6.5	(5.6, 7.3)
	Prior Year Social Studies Test	6.0	(5.6, 6.4)
	Prior Year Writing Test	4.3	(4.0, 4.5)
	Prior Year Mathematics Test	2.0	(1.8, 2.3)
	Student Absences	-0.1	(-0.1, -0.1)
	Gender (Male)	-0.8	(-1.1, -0.4)
	Free/Reduced Lunch	-2.3	(-2.7, -2.0)
	Limited English Proficiency	-3.5	(-6.0, -1.1)
	Speech Language Impairment	-3.6	(-4.6, -2.6)
	African American	-4.1	(-4.5, -3.6)
	Emotionally Disturbed	-9.3	(-11.8, -6.8)
	Other Health Impaired	-11.5	(-13.0, -10.0)
	Specific Learning Disability	-16.8	(-17.9, -15.7)
	Mild Mental Retardation	-23.3	(-25.7, -21.0)
Classroom variables	% Gifted	12.7	(9.5, 15.9)
	Teacher Absences	0.0	(-0.1, 0.0)
	Class Mean Prior Year Reading Test	-6.9	(-8.2, -5.6)
	% Gender (Male)	-7.9	(-11.9, -3.9)
	% Free/Reduced Lunch	-11.3	(-15.2, -7.4)
Building Variables	School Mean Prior Year Reading Test	10.4	(8.2, 12.5)
	% Free/Reduced Lunch	1.2	(0.7, 1.7)
	% Gifted	-1.4	(-2.3, -0.5)

Among the demographic variables, being gifted was the only positive coefficient. The special education disabilities Speech Language Impairment, Emotionally Disturbed, Other Health Impairment, Specific Learning Disability, and Mild Mental Retardation all had negative coefficients. Mild Mental Retardation had the largest negative coefficient at -23.3. It is interesting to note that being African American has a larger negative coefficient than having Speech Language Impairment. When examining the student absences coefficient it must be noted that the effect is for every day missed. This means

that a student who had missed 10 days of school would score at least 1 point lower on the test than a student with perfect attendance.

Among the classroom variables it would be expected that having a larger percentage of gifted students in the classroom would result in higher scores just as the values demonstrate. It may come as a surprise that students in a classroom with higher achieving students may perform at lower levels, however, this most likely is attributed to a correction loading to attenuate strong positive loadings on both the student and school levels. Students in classes with a higher percentage of students receiving free/reduced lunch also performed lower on the reading test. It is interesting to note that students in classrooms with a larger portion of males score lower on the reading test than those students in classrooms with fewer male students.

The school building coefficients demonstrate that students attending schools with students who have higher prior achievement are predicted to perform better on current year assessments. Having a higher percentage of free lunch students had a slightly positive effect while a larger portion of gifted students had a slightly negative effect. As noted above, it is important to keep in mind the difference in scaling among the coefficients between differing types of predictors.

Again, in the content of writing, a student's prior year achievement in that content was the single largest contributor to current year achievement. This effect was similar to the reading content because the prior year achievement in writing was more than three times the value of any other prior year achievement in the other content areas. Being gifted had a large positive effect on current achievement as well as being Asian. Being African American had a slightly positive effect. Having absences, being male,

Table 8: Hierarchical Linear Model for Writing Achievement 2005

Model Level	Variables Entered	Coefficient	(CI)
Student level variables	Prior Year Writing Test	19.9	(19.3, 20.5)
	Gifted	10.0	(9.1, 11.0)
	Prior Year Reading Test	6.5	(6.2, 6.8)
	Prior Year Mathematics Test	6.0	(5.8, 6.3)
	Asian	5.7	(4.6, 6.9)
	Prior Year Social Studies Test	3.1	(2.7, 3.4)
	Prior Year Science Test	1.9	(1.6, 2.1)
	African American	1.0	(0.7, 1.4)
	Student Absences	-0.3	(-0.3, -0.3)
	Free/Reduced Lunch	-2.2	(-2.6, -1.9)
	Speech Language Impairment	-2.7	(-3.6, -1.8)
	Emotionally Disturbed	-4.8	(-7.2, -2.4)
	Other Health Impaired	-7.7	(-9.0, -6.4)
	Gender (Male)	-9.3	(-9.6, -9.0)
	Specific Learning Disability	-10.8	(-11.7, -9.9)
	Mild Mental Retardation	-20.3	(-22.9, -17.7)
Classroom variables	% Gifted	12.1	(8.7, 15.5)
	Teacher Absences	-0.1	(-0.1, 0.0)
	Class Prior Mean Writing Test	-5.5	(-6.8, -4.1)
	% Free/Reduced Lunch	-11.6	(-15.5, -7.6)
	% Male	-12.0	(-16.0, -8.1)
Building Variables	School Mean Prior Writing Test	12.6	(9.1, 16.1)
	% Free/Reduced Lunch	1.0	(0.5, 1.5)
	% Gifted	-1.1	(-1.9, -0.3)
	School Prior Mathematics Test	-4.7	(-7.9, -1.6)

and being classified as one of the five special education disability categories all had negative effects. Mental retardation again, had the largest negative effect.

Within the classroom variables, being in a classroom with a higher percentage of gifted students had a large positive effect. Small negative effects were found for prior year achievement in writing and teacher absences. Large negative effects were found for being in a classroom with more free lunch and male students.

School variables revealed a large positive effect for a school's prior year aggregate achievement. The large effect for school building prior year achievement appears to be partially attributable to scaling. Moving an entire school up a standard deviation in achievement would be a large effect. A small positive coefficient for free/reduced lunch status and a small negative coefficient for gifted was found. These coefficients are unexpected, but may be explained by a corrective loading to attenuate for the strong negative loadings for free/reduced lunch status and strong positive loadings for gifted at the student and classroom levels. Also, a moderate sized negative coefficient was found for prior year aggregate achievement in mathematics.
Again, as above, caution is warranted when comparing across differing predictors. Keep in mind the differences in scaling.

Like the other 2004-2005 year content data, prior year achievement in mathematics was the single largest variable for a student's current achievement among the achievement variables. This effect was slightly larger than it was for reading and writing contents with the mathematics coefficient being more than four times the value of the other prior year achievements in other content domains. For mathematics, being gifted, Asian, and male lead to positive performances on current assessments. Student absences, being African American, and being classified as one of the five special education disabilities all had negative coefficients. As could be expected, Mild Mental Retardation had the largest negative coefficient.

Classroom variables revealed a large positive coefficient for a high percentage of gifted students in the classroom. Large negative coefficients were found for a larger percentage of male students and students who receive free/reduced lunches. There was little to no effect for teacher absences.

Table 9: Hierarchical Linear Model for Mathematics Achievement 2005

Model Level	Variables Entered	Coefficient	(CI)
Student level variables	Prior Year Mathematics Test	25.5	(25.1, 25.9)
	Gifted	6.4	(5.6, 7.2)
	Asian	5.5	(4.3, 6.7)
	Prior Year Writing Test	5.5	(5.3, 5.8)
	Limited English Proficiency	4.3	(2.4, 6.1)
	Prior Year Science Test	3.9	(3.7, 4.2)
	Prior Year Reading Test	2.7	(2.4, 3.0)
	Prior Year Social Studies Test	2.2	(1.9, 2.5)
	Gender (Male)	2.1	(1.9, 2.4)
	Student Absences	-0.3	(-0.3, -0.2)
	Free Lunch/Reduced Lunch	-1.2	(-1.5, -0.9)
	Speech Language Impairment	-2.3	(-3.2, -1.4)
	African American	-5.6	(-6.0, -5.2)
	Emotionally Disturbed	-7.2	(-10.5, -4.0)
	Specific Learning Disability	-8.8	(-9.8, -7.8)
	Other Health Impaired	-10.1	(-11.5, -8.7)
	Mild Mental Retardation	-16.6	(-20.5, -12.6)
Classroom variables	% Gifted	6.4	(4.2, 8.6)
	Teacher Absences	0.0	(-0.1, 0.0)
	% Male	-8.1	(-11.5, -4.6)
	% Free/Reduced Lunch	-12.6	(-15.6, -9.5)
Building Variables	% Free/Reduced Lunch	12.6	(8.4, 16.9)
	Mean Prior Year Mathematics Test	4.3	(2.5, 6.2)

School based variables revealed a large positive coefficient for a higher portion of free/reduced lunch status students. Although the meaning of this finding is uncertain, it would appear that this was a corrective loading to attenuate the strong negative coefficient at both the student and classroom levels. A positive coefficient was found for students attending a school with a higher aggregate prior year achievement.

As with the 2004-2005 data, the 2005-2006 data also had the same scaling differences among different predictor variables. These differences should be considered when comparing different predictors.

Table 10: Hierarchical Linear Model for Reading Achievement 2006

Model Level	Variables Entered	Coefficient	(CI)
Student level variables	Prior Year Reading Test	18.2	(17.5, 18.9)
	Prior Year Science Test	7.0	(6.6, 7.4)
	Prior Year Social Studies Test	6.6	(6.1, 7.0)
	Prior Year Writing Test	4.1	(3.8, 4.4)
	Gifted	3.9	(3.0, 4.8)
	Prior Year Mathematics Test	3.5	(3.2, 3.8)
	Student Absences	-0.1	(-0.2, -0.1)
	Gender (Male)	-1.1	(-1.5, -0.7)
	Reduced Lunch	-1.2	(-1.8, -0.6)
	Free Lunch	-3.2	(-3.6, -2.8)
	African American	-4.5	(-5.0, -3.9)
	Speech Language Impairment	-4.7	(-5.8, -3.5)
	Section 504	-7.7	(-8.6, -6.8)
	Emotionally Disturbed	-8.1	(-11.7, -4.5)
	Other Health Impaired	-11.6	(-13.3, -10.0)
	Mild Mental Retardation	-17.1	(-20.2, -13.9)
	Specific Learning Disability	-17.3	(-18.6, -16.1)
Classroom variables	% Male	-0.6	(-1.0, -0.2)
	% Free Lunch	-2.2	(-2.6, -1.7)
	Class Mean Prior Reading Test	-5.0	(-6.3, -3.6)
Building Variables	School Mean Prior Year Reading Test	7.6	(5.0, 10.2)
	% Free Lunch	2.0	(1.4, 2.6)

The single, largest contributor among the prior achievement variables to student current achievement was prior year achievement in reading. This coefficient was over two and a half times the value of the next positive coefficient, which was prior year achievement in science. Gifted was the only student demographic with a positive coefficient value from the demographic predictors in this specific model. The special education disability categories of Emotionally Disturbed, Other Health Impaired, Mild Mental Retardation, and Specific Learning Disability were the largest negative effects in this model.

There were no positive coefficients among the classroom level variables. Being in a classroom with a high percentage of males, free lunch status, and higher achieving peers one would be expected to perform more poorly on the reading test. Again, the paradoxical effect of performing more poorly when in a classroom of higher achieving peers was seen as in the 2004-2005 reading data. It must be noted that this effect was a phenomenon that only occurs in the context of a system in which a tremendous amount of information is already available regarding student achievement.

In the school level variables, prior year achievement in reading resulted in a positive coefficient and percentage of students receiving free lunch appeared to load positively.

As noted above, be cautious when comparing across differing predictors. Keep in mind the differences in scaling.

Similar to the other contents and years, prior year writing achievement was the single largest contributor to current year achievement in writing among the prior achievement variables. Prior year achievement coefficients in all content areas were positive effects with prior year achievement in science being the lowest coefficient value. Being Asian and gifted were the only two demographic variables resulting in positive coefficients. However, the gifted coefficient was lower in for this content than in previous contents and years. For every student absence a loss of -0.4 points from the current year writing test can be expected. For example, if a student misses 10 days, that student would be expected to perform at least 4 points lower on the test as compared to a student with perfect attendance. Mild Mental Retardation was the largest negative coefficient with a value of -25.6.

Table 11: Hierarchical Linear Model for Writing Achievement 2006

Model Level	Variables Entered	Coefficient	(CI)
Student level variables	Prior Year Writing Test	20.7	(20.3, 21.2)
	Prior Year Reading Test	7.0	(6.7, 7.4)
	Prior Year Mathematics Test	5.8	(5.5, 6.1)
	Asian	5.0	(3.9, 6.2)
	Gifted	3.5	(2.7, 4.3)
	Prior Year Social Studies Test	2.3	(2.0, 2.6)
	Prior Year Science Test	1.0	(0.7, 1.3)
	Student Absences	-0.4	(-0.4, -0.4)
	Reduced Lunch	-0.9	(-1.4, -0.4)
	Free Lunch	-1.9	(-2.3, -1.6)
	Speech Language Impairment	-5.7	(-6.8, -4.6)
	Section 504	-9.7	(-10.6, -8.8)
	Other Health Impaired	-13.7	(-15.1, -12.2)
	Emotionally Disturbed	-16.3	(-19.5, -13.2)
	Specific Learning Disability	-18.1	(-19.1, -17.1)
	Mild Mental Retardation	-25.6	(-28.5, -22.6)
Classroom variables	% Male	0.8	(0.3, 1.4)
	Teacher Absences	0.1	(0.0, 0.1)
	% Reduced Lunch	0.0	(-0.1, 0.0)
	% Section 504	-1.1	(-1.3, -0.8)
	% Limited English Proficiency	-1.8	(-2.2, -1.5)
Building Variables	School Mean Prior Year Writing Test	4.0	(2.2, 5.8)
	% Minority	0.4	(0.2, 0.6)

Among the classroom variables higher percentage of males and teacher absences had a minute positive effect on current year achievement. Percentage of students in the classroom receiving free lunch had no effect while more students in the classroom with a Section 504 status and Limited English Proficiency had small negative effects.

As would be expected, attending a school with a higher aggregate prior year achievement in writing would result in a higher score. The school building coefficient for every 10% of the student population identified as being a minority was small and positive.

Table 12: Hierarchical Linear Model for Mathematics Achievement 2006

Model Level	Variables Entered	Coefficient	(CI)
Student level variables	Prior Year Mathematics Test	25.0	(24.6, 25.3)
	Gifted	8.8	(7.8, 9.8)
	Asian	7.8	(6.4, 9.2)
	Prior Year Writing Test	5.6	(5.3, 5.8)
	Prior Year Science Test	4.1	(3.9, 4.4)
	Gender (Male)	2.8	(2.5, 3.1)
	Prior Year Reading Test	2.1	(1.8, 2.4)
	Prior Year Social Studies Test	2.0	(1.7, 2.3)
	Student Absences	-0.3	(-0.4, -0.3)
	Free Lunch	-1.4	(-1.8, -1.1)
	Native American	-1.8	(-3.2, -0.4)
	Speech Language Impairment	-3.2	(-4.3, -2.1)
	Section 504	-5.2	(-6.1, -4.3)
	African American	-5.9	(-6.3, -5.5)
	Emotionally Disturbed	-11.1	(-14.9, -7.2)
	Other Health Impaired	-12.0	(-13.4, -10.6)
	Specific Learning Disability	-12.5	(-13.7, -11.4)
	Mild Mental Retardation	-25.7	(-30.5, -20.9)
Classroom variables	% Gifted	10.1	(7.2, 13.0)
	Teacher Absences	-0.1	(-0.1, 0.0)
	% Free Lunch	-7.8	(-10.4, -5.2)
Building Variables	School Mean Prior Year Mathematics Test	5.6	(3.4, 7.7)
	% Minority	5.5	(3.1, 8.0)
	% Gifted	-14.5	(-23.2, -5.8)

As noted above, be cautious when comparing across differing predictors. Keep in mind the differences in scaling.

Like all other contents in both years, the single largest predictor for current year mathematics achievement among prior year achievement scores was prior year mathematics achievement by a substantial amount. Prior year reading, writing, science, and social studies achievement were also positive effects. Similar to 2005-2006 writing data, giftedness and Asian were both positive coefficients. Along with these

demographic variables, being male was also a positive coefficient. All other demographic variables were negative with Mild Mental Retardation being the largest negative coefficient.

The contribution of classroom demographic variables to the predictions was similar to the 2004-2005 mathematics data and generally in the expected direction. Students who attend class with more gifted students would perform better. Students in classrooms in which the teacher was absent would perform more poorly as well as if the classroom had a higher population of students receiving free lunch.

Attending a school with more students who perform better on the prior year test would result in a higher score. Also, the school building coefficient for every 10% increase in the school's population of minority students was moderately positive. The school building level gifted status loaded negatively.

Summary*:* Generally, the student level variables were similar. Among all content areas and both years the prior year achievement for the target content was the largest contributor to current year achievement among the achievement variables. All of the special education disabilities resulted in negative effects and many were large. Giftedness and being an Asian American resulted in positive effects to varying degrees among contents and years. Student absences, free/reduced lunch status, and section 504 generally exhibited negative coefficients. Variables at the classroom level were not as consistent. However, teacher absences generally resulted in small negative effects as well as percentage of students receiving free/reduced lunch. Percent of students who were gifted at the classroom level generally resulted in positive effects as well as prior year achievement for the target content. There was even less consistency among school level variables. Generally, prior year aggregate achievement was positive.

Once the final models for student achievement were developed with students nested within classrooms and schools they were used to examine the adjustment in students' predicted achievement that would be predicted when different teacher and classroom characteristics were in place. The models examined adjustments to students' predicted achievement when students were taught by a teacher who has a special education certificate, has a regular versus alternative certificate, has taught for a number of years, and is certified to teach in the target content. The models also examined the adjusted predicted achievement when in a classroom with a higher percentage of special education students versus a class full of regular education students.

It is important to remember that for the current study the teacher and classroom variables of interest were only examined for students who were identified as being classified as one of the five special education categories: Speech Language Impairment, Specific Learning Disability, Emotionally Disturbed, Other Health Impaired, and Mild Mental Retardation. Dummy codes were used to identify these students. For example, if a student had been identified as being Emotionally Disturbed, that student would have a 1 in that data column.

For the analyses for each predictor variable and special education category there had to be at least 50 students in the diagnostic category (i.e. Emotionally Disturbed) and at least 10 teachers each who exhibited values of the dichotomous specific predictor variable (i.e. 10 teachers who were certified and10 teachers who were not certified in special education teaching the specified disability group) to analyze. If there were not enough of either students or teachers in the category it was left out of the analysis. This was done because if there were not enough cases for the analysis the data were unstable and the results were unclear.

The following tables represent the values obtained for each special education category. The coefficient and confidence interval (95% CI), p-value, and number of teachers for the target predictor are presented. A table is presented for each content domain and each year. The first three tables represent final values after analyzing all students in that specified content domain for the school year 2004-2005. The next three tables represent the final values obtained after analyzing students identified as being in category three, those students who scored one or more standard deviations below the mean on the target content domain for the school year 2004-2005. The next three tables represent final values after analyzing all students in that specified content domain for the school year 2005-2006. Finally, the last three tables represent the final values obtained after analyzing students identified as being in category three, those students who scored one or more standard deviations below the mean on the target content domain for the school year 2005-2006. A brief discussion follows each table with a summary of each set of three after the set of tables are displayed.

The tables presented below should be interpreted as follows: the first column represents the special education category that was analyzed. The second column presents the predictor variables that contained enough cases to be analyzed. The third column is the coefficient that was obtained in the analysis. The fourth column depicts the p-value for that coefficient. Finally, the last column presents the number of teachers used for each predictor variable analysis. For example, in the first table, there were a total of 762 teachers who were analyzed when investigating the predictor variable teachers' years experience for the group Emotionally Disturbed. However, for the predictor variable special education certification, there were 272 teachers who had a special education certification and 490 teachers who did not have a special education certification

instructing students with Emotional Disturbance. The number presented in the last column presents the number of teachers instructing the specific disability category who displayed that specific predictor variable. Also, the title of the table presents the school year, the content area, and also the group of students in the analysis. When the table is titled as "All Students" this means that all students identified as one of the special education categories were used in the analysis. When the table is titled "Category 3 Students" this means that only students who scored one or more standard deviations below the mean, who were identified as one of the special education categories, were used in the analysis.

The first table presented below represents the final values obtained for the school year 2004-2005, for all students in the content domain reading.

For students identified as being Emotionally Disturbed the only significant finding was for the predictor variable reading certification and it was -9.4. For students being identified as having a Specific Learning Disability the results are similar to those for Emotionally Disturbed students. There were two significant findings for this group. These significant predictors were being in a classroom with a higher percentage of special education students with a coefficient of -7.4 and having a teacher with a reading certificate with a coefficient of -5.4. The result of a negative coefficient for teachers with a reading certification was an unusual finding and will be discussed further in the discussion section.

For the Mild Mental Retardation group none of the predictor variables were significant, although the predictor reading certification was close to reaching significance and reflects the same negative effect as the first two disability categories.

Table 13: Reading 2004-2005 Predictor Variable Effects, All Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	-0.1 (-0.3, 0.1)	0.414	762
	Percent Special Education	-5.8 (-15.2, 3.5)	0.221	762
	Special Education Certified	-2.0 (-8.9, 4.9)	0.577	272
	Reading Certification	-9.4 (-17.1, -1.7)	0.017*	692
	Alternative Certification	-6.7 (-23.5, 10.1)	0.433	18
Specific Learning Disability	Years Experience	0.0 (0, 0.1)	0.346	10,257
	Percent Special Education	-7.4 (-11.3, -3.5)	<0.001*	10,257
	Special Education Certified	2.6 (-0.2, 5.5)	0.073	2,234
	Reading Certification	-5.4 (-7.7, -3)	<0.001*	8,980
	Alternative Certification	-1.0 (-8.1, 6)	0.775	209
Mild Mental Retardation	Years Experience	0.0 (-0.2, 0.2)	0.937	1,353
	Percent Special Education	-4.2 (-11.4, 3)	0.255	1,353
	Special Education Certified	-0.7 (-6.2, 4.8)	0.812	531
	Reading Certification	-7.7 (-15.5, 0.1)	0.053	1,213
	Alternative Certification	-4.4 (-14.4, 5.5)	0.383	61
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.454	4,773
	Percent Special Education	-3.9 (-12.1, 4.4)	0.358	4,773
	Special Education Certified	-0.5 (-4.1, 3.1)	0.794	353
	Reading Certification	4.2 (1.1, 7.3)	0.008*	4,498
	Alternative Certification	-6.7 (-14.8, 1.5)	0.108	55
Other Health Impairments	Years Experience	-0.1 (-0.2, 0.1)	0.302	3,079
	Percent Special Education	-7.4 (-13.4, -1.3)	0.017*	3,079
	Special Education Certified	-0.3 (-4.5, 3.8)	0.878	572
	Reading Certification	-0.5 (-4.3, 3.3)	0.781	2,758
	Alternative Certification	-5.4 (-17, 6.3)	0.367	42

Students with a Speech Language Disability also had similar results to the previous groups. Interestingly, unlike the other groups, a Speech Language Disabled student would be predicted to score higher (+4.2) when instructed by a teacher with a reading certification, which was the only significant finding.

For the final group, Other Health Impaired, results were different from the other categories. The only significant finding was for the predictor variable of being placed in a classroom with a higher percentage of special education students at -7.4.

The following table presents the values obtained for the school year 2004-2005 for all students in the target content writing.

For students identified as Emotionally Disturbed there were not enough teachers without a writing certification for this predictor variable to be analyzed. Teachers' years' experience was the only significant effect at -0.3.

For the group of Specific Learning Disabled students' being in a classroom with a higher percentage of special education students led to a poorer performance on the current test (-4.1) and this variable was the only significant finding.

Again, for students identified as Mild Mental Retardation the predictor variable classroom percent special education was the only significant finding at -11.2.

For students identified as having a Speech Language Disability there was a large positive effect for having a teacher with a writing certification (10.7), which was the only significant result.

None of the predictor variables for the group Other Health Impaired resulted in a significant finding.

Table 15 presents the final values obtained for the school year 2004-2005, for all students in the target content mathematics.

Table 14: Writing 2004-2005 Predictor Variable Effects, All Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	-0.3 (-0.5, 0.0)	0.024*	1,009
	Percent Special Education	-5.5 (-15.4, 4.4)	0.276	1,009
	Special Education Certified	4.8 (-3.0, 12.6)	0.229	335
	Alternative Certification	1.7 (-7.2, 10.6)	0.702	984
Specific Learning Disability	Years Experience	0.0 (-0.1, 0.1)	0.611	4,077
	Percent Special Education	-4.1 (-7.5, -0.7)	0.018*	4,077
	Special Education Certified	-1.8 (-4.0, 0.5)	0.124	1,096
	Writing Certification	1.8 (-3.3, 6.8)	0.494	4,004
	Alternative Certification	3.5 (-1.3, 8.3)	0.152	110
Mild Mental Retardation	Years Experience	0.0 (-0.2, 0.2)	0.998	1,565
	Percent Special Education	-11.2 (-20.4, -1.9)	0.018*	1,565
	Special Education Certified	3.4 (-4.0, 10.8)	0.371	581
	Writing Certification	9.4 (-8.9, 27.7)	0.316	1,511
	Alternative Certification	-9.6 (-19.8, 0.6)	0.066	67
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.473	2,942
	Percent Special Education	-4.2 (-13.0, 4.6)	0.354	2,942
	Special Education Certified	0.1 (-3.3, 3.6)	0.935	199
	Writing Certification	10.7 (2.0, 19.5)	0.016*	2,914
	Alternative Certification	-0.9 (-10.2, 8.4)	0.847	33
Other Health Impairments	Years Experience	-0.1 (-0.2, 0.1)	0.359	4,003
	Percent Special Education	-4.1 (-9.9, 1.7)	0.164	4,003
	Special Education Certified	0.9 (-3.3, 5.1)	0.673	697
	Writing Certification	-3.2 (-12.3, 5.9)	0.490	3,923
	Alternative Certification	-0.1 (-8.4, 8.3)	0.990	64

Table 15: Mathematics 2004-2005 Predictor Variable Effects, All Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.0 (-0.2, 0.2)	0.968	766
	Percent Special Education	1.4 (-8.9, 11.7)	0.787	766
	Special Education Certified	-5.6 (-15.3, 4.0)	0.253	226
	Alternative Certification	12.0 (-4.5, 28.5)	0.154	35
Specific Learning Disability	Years Experience	0.0 (0, 0.1)	0.409	10,302
	Percent Special Education	-1.6 (-5.9, 2.7)	0.466	10,302
	Special Education Certified	2.5 (-0.8, 5.7)	0.144	1,779
	Mathematics Certification	-2.0 (-7.2, 3.2)	0.459	10,013
	Alternative Certification	3.5 (-2.6, 9.6)	0.258	316
Mild Mental Retardation	Years Experience	-0.3 (-0.5, 0.0)	0.041*	1,203
	Percent Special Education	6.1 (-7.8, 20.0)	0.389	1,203
	Special Education Certified	-8.3 (-19.9, 3.4)	0.164	429
	Mathematics Certification	-0.7 (-24.0, 22.6)	0.955	1,156
	Alternative Certification	9.1 (-13.0, 31.2)	0.420	58
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.583	4,744
	Percent Special Education	-3.0 (-11.5, 5.4)	0.482	4,744
	Special Education Certified	1.1 (-2.5, 4.6)	0.558	282
	Mathematics Certification	-8.1 (-17.1, 0.9)	0.077	4,692
	Alternative Certification	-1.2 (-10.7, 8.4)	0.812	58
Other Health Impairments	Years Experience	0.1 (0.0, 0.2)	0.240	3,101
	Percent Special Education	-10.6 (-16.2, -5.1)	<0.001*	3,101
	Special Education Certified	7.9 (2.9, 12.9)	0.002*	480
	Mathematics Certification	-2.7 (-17.0, 11.7)	0.716	3,040
	Alternative Certification	-3.3 (-11.3, 4.7)	0.421	73

For students identified as being Emotionally Disturbed there were not enough teachers without a mathematics certification to analyze. None of the results were found to be significant.

Again, like the Emotionally Disturbed group, none of the predictor variables for the group of students identified as Specific Learning Disability resulted in significant findings.

There was a slightly larger, negative effect for teachers' years experience than the past contents (-0.3) with a significant finding for the group identified as Mild Mental Retardation. However, this predictor variable was the only significant finding.

For the group Speech Language Impairment none of the predictor variables were found to be significant.

Two of the predictor variables for the group Other Health Impaired resulted in significant findings. These two predictors were percent special education (-10.6) and a having a teacher with a special education certification (7.9).

Summary. The following table presents a summary of all the significant findings for all students analyzed in all content areas for the school year 2004-2005.

For all groups, the predictor variable Percent Special Education resulted in negative coefficients. For the groups Emotionally Disturbed and Specific Learning Disability, the content certification resulted in negative findings, however, for the group of Speech Language Disability, the findings resulted positive coefficients. Negative results were also found for the predictor variables years experience and special education certification.

Table 17 presents the final values obtained for the school year 2004-2005, for students in category three in the target content reading.

Table 16: Significant Findings for All Students in 2004-2005 for All Content Areas

Special Education Group	Reading (coefficient)	Writing (coefficient)	Mathematics (coefficient)
Emotionally Disturbed	1.Reading Certification (-9.4)	1. Years Experience (-0.3)	None
Specific Learning Disability	1.Percent Special Education (-7.4) 2. Reading Certification (-5.4)	1. Percent Special Education (-4.1)	None
Mild Mental Retardation	None	1. Percent Special Education (-11.2)	1. Years Experience (-0.3)
Speech Language Disability	1. Reading Certification (4.2)	1. Writing Certification (10.7)	None
Other Health Impaired	1. Percent Special Education (-7.4)	None	1. Percent Special Education (-10.6) 2. Special Education Certification (7.9)

For the Emotionally Disturbed group only three of the predictor variables had enough cases to analyze. There were not enough teachers who did not have a reading certificate or an alternative certificate to analyze and thus these variables were left out. Of the three predictor variables analyzed, none were found to be significant.

All predictor variables for the Specific Learning Disability group contained enough cases and were therefore analyzed. Three of the predictor variables resulted in

Table 17: Reading 2004-2005 Predictor Variable Effects, Category 3 Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.0 (-0.2, 0.2)	0.746	261
	Percent Special Education	-5.1 (-12.4, 2.2)	0.174	261
	Special Education Certified	-4.0 (-9.5, 1.4)	0.149	121
Specific Learning Disability	Years Experience	0.0 (0.0, 0.1)	0.544	3,703
	Percent Special Education	-3.5 (-6.5, -0.6)	0.020*	3,703
	Special Education Certified	2.2 (0.1, 4.2)	0.040*	978
	Reading Certification	-3.2 (-5.3, -1.1)	0.003*	3,295
	Alternative Certification	-2.5 (-6.5, 1.5)	0.222	94
Mild Mental Retardation	Years Experience	0.0 (-0.2, 0.1)	0.596	871
	Percent Special Education	-2.9 (-8.8, 3.0)	0.335	871
	Special Education Certified	3.3 (-1.0, 7.7)	0.135	382
	Reading Certification	-3.0 (-7.2, 1.2)	0.166	797
Speech Language Disability	Years Experience	-0.1 (-0.2, 0.0)	0.192	765
	Percent Special Education	1.0 (-6.4, 8.4)	0.795	765
	Special Education Certified	-3.2 (-7.4, 1.0)	0.139	65
	Reading Certification	-3.6 (-7.7, 0.4)	0.078	701
Other Health Impairments	Years Experience	-0.1 (-0.2, 0.1)	0.399	732
	Percent Special Education	-5.6 (-12, 0.8)	0.085	732
	Special Education Certified	4.2 (-0.4, 8.7)	0.071	209
	Reading Certification	-0.5 (-5.5, 4.5)	0.845	670

significant results. Being placed in a classroom with a higher percentage of special education students resulted in a -3.5 effect, having a teacher with a special education certification resulted in a 2.2 effect, and lastly, having a teacher with a reading certification resulted in a -3.2 effect. Having a teacher with a reading certification result

in a negative coefficient is an unusual finding and will be discussed further in the discussion section.

There were not enough teachers with an alternative certification teaching students with Mild Mental Retardation to run in the analysis. Again, none of the results for this group were found to be significant.

For students identified as having Speech Language Impairment there were not enough teachers with an alternative certification and therefore this predictor variable was not analyzed. Of the remaining predictor variables, none of them were found to be significant. However, although not significant, having a teacher with a reading certification was close to reaching significance and similar to the group of Specific Learning Disabled, which resulted in a -3.6 effect.

Finally, for the students identified as Other Health Impaired the results were similar to prior groups and there were not enough teachers to analyze the predictor variable of having an alternative certification. Although none of the results were found to be significant, having a teacher with a special education certification was close to reaching significance with a coefficient of 4.2, similar to the Specific Learning Disabled group.

Table 18 presents the final values obtained from the school year 2004-2005 for students scoring one or more standard deviations below the mean on the writing test.

For this category of students, those identified as being Emotionally Disturbed, there were not enough teachers with an alternative certification or without a writing certification to analyze. Although none of the predictor variables reached significance, having a teacher with a special education certification was close to reaching significance with a high positive coefficient (7.1).

Table 18: Writing 2004-2005 Predictor Variable Effects, Students in Category 3

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	-0.2 (-0.5, 0.1)	0.116	350
	Percent Special Education	-5.5 (-16.4, 5.4)	0.322	350
	Special Education Certified	7.1 (-0.3, 14.6)	0.058	155
Specific Learning Disability	Years Experience	0.1 (0.0, 0.1)	0.026*	4,077
	Percent Special Education	-2.7 (-5.8, 0.3)	0.074	4,077
	Special Education Certified	1.1 (-1.1, 3.2)	0.342	1,096
	Writing Certification	2.5 (-2.9, 7.9)	0.365	4,004
	Alternative Certification	1.2 (-2.4, 4.9)	0.515	110
Mild Mental Retardation	Years Experience	0.0 (-0.2, 0.2)	0.847	885
	Percent Special Education	-4.5 (-12.4, 3.3)	0.259	885
	Special Education Certified	4.5 (-1.6, 10.6)	0.147	380
Speech Language Disability	Years Experience	0.0 (-0.2, 0.1)	0.501	588
	Percent Special Education	-4.6 (-13.1, 3.8)	0.284	588
	Special Education Certified	-4.1 (-10.4, 2.1)	0.196	41
Other Health Impairments	Years Experience	-0.1 (-0.2, 0.0)	0.110	884
	Percent Special Education	-4.4 (-10.4, 1.5)	0.145	884
	Special Education Certified	0.2 (-3.8, 4.2)	0.918	241

For the group identified as Specific Learning Disability, teacher years experience had a small positive effect of 0.1 at a significant level. Being placed in a classroom with a higher percentage of special education students neared significance with a negative coefficient of -2.7, which was similar to findings in other content areas. None of the other predictor variables reached significance.

For the group Mild Mental Retardation only three predictor variables had enough cases to analyze. None of the variables analyzed reached significance.

The results for the Speech Language Impairment group were like the prior group. Only three variables had enough teachers to analyze and none of the predictor variables analyzed reached a significant level.

Finally, for the group, Other Health Impaired, similar results were found. Again, only three variables had enough cases to be analyzed, however, none of the results reached significance.

Table 19 presents the final values obtained from the school year 2004-2005 for students scoring one or more standard deviations below the mean on the mathematics test.

For students identified as being Emotionally Disturbed and scoring more than one standard deviation below the mean on the mathematics test there were not enough teachers without a mathematics certification or an alternative certification to analyze. For the remaining predictor variables, none of the results reached significance.

For the group Specific Learning Disability two of the predictor variables resulted in significant results. The predictor variable percent special education was significant with a coefficient of -6.3. The variable alternative certification was also significant with a coefficient of 6.5.

For all of the final three groups of students, Mild Mental Retardation, Speech Language Impaired, and Other Health Impaired, only three of the predictor variables were analyzed due to a lack of cases. Of the predictor variables analyzed, none of the results were found to be significant.

Summary. Table 20 presents a summary of all the significant findings for all students analyzed in all content areas for the school year 2004-2005.

Table 19: Mathematics 2004-2005 Predictor Variable Effects, Category 3 Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.0 (-0.2, 0.2)	0.818	236
	Percent Special Education	4.7 (-4.1, 13.6)	0.292	236
	Special Education Certified	-7.5 (-17.5, 2.5)	0.143	100
Specific Learning Disability	Years Experience	0.0 (-0.1, 0.1)	0.684	3,157
	Percent Special Education	-6.3 (-11.4, -1.2)	0.016*	3,157
	Special Education Certified	3.6 (-1.0, 8.1)	0.127	669
	Mathematics Certification	-4.0 (-11.6, 3.6)	0.308	3,105
	Alternative Certification	6.5 (2.4, 10.5)	0.002*	99
Mild Mental Retardation	Years Experience	-0.2 (-0.4, 0.1)	0.143	703
	Percent Special Education	-2.7 (-15.3, 9.9)	0.676	703
	Special Education Certified	-4.9 (-15.6, 5.8)	0.368	240
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.407	662
	Percent Special Education	-10.8 (-25.2, 3.7)	0.144	662
	Special Education Certified	0.9 (-5.4, 7.2)	0.782	32
Other Health Impairments	Years Experience	0.1 (-0.1, 0.2)	0.598	726
	Percent Special Education	-5.6 (-12.7, 1.5)	0.123	726
	Special Education Certified	0.6 (-5.7, 6.8)	0.858	157

Few variables resulted in significant findings for this group; only the group Specific Learning Disability had significant findings. For this group, the predictor variable percent special education consistently resulted in negative coefficients. A teacher with a reading certification was a negative result while a teacher with an alternative certification

was a positive result. Finally, teacher's years experience also resulted in a positive finding.

Table 20: Significant Findings for Category 3 Students in 2004-2005 for All Content Areas

Special Education Group	Reading (coefficient)	Writing (coefficient)	Mathematics (coefficient)
Emotionally Disturbed	None	None	None
Specific Learning Disability	1.Percent Special Education (-3.5) 2. Reading Certification (-3.2)	1. Years Experience (0.1)	1. Percent Special Education (-6.3) 2. Alternative Certification (6.5)
Mild Mental Retardation	None	None	None
Speech Language Disability	None	None	None
Other Health Impaired	None	None	None

The following table presents the values obtained for the school year 2005-2006 for all students in the target content reading.

Table 21: Reading 2005-2006 Predictor Variable Effects, All Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.3 (-0.1, 0.6)	0.104	403
	Percent Special Education	-1.2 (-2.6, 0.3)	0.125	403
	Special Education Certified	9.5 (-2.4, 21.4)	0.118	153
Specific Learning Disability	Years Experience	0.0 (-0.1, 0.1)	0.824	7,226
	Percent Special Education	-1.0 (-1.5, -0.6)	<0.001*	7,226
	Special Education Certified	2.0 (-1.5, 5.5)	0.270	2,014
	Reading Certification	-5.3 (-13.2, 2.5)	0.185	7,131
	Alternative Certification	1.0 (-4.0, 6.0)	0.689	250
Mild Mental Retardation	Years Experience	0.0 (-0.3, 0.2)	0.868	830
	Percent Special Education	-1.1 (-2.4, 0.2)	0.100	830
	Special Education Certified	5.9 (-4.2, 16.0)	0.253	375
	Reading Certification	-10.7 (-23.4, 1.9)	0.096	808
	Alternative Certification	-5.1 (-18.8, 8.5)	0.460	37
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.768	3,035
	Percent Special Education	-1.0 (-2.0, 0.1)	0.074	3,035
	Special Education Certified	-0.4 (-4.4, 3.6)	0.850	271
	Reading Certification	6.9 (3.1, 10.7)	0.001*	3,011
	Alternative Certification	-14.4 (-21.7, -7.1)	<0.001*	56
Other Health Impairments	Years Experience	0.1 (-0.1, 0.3)	0.315	2,213
	Percent Special Education	-0.6 (-1.3, 0.1)	0.102	2,213
	Special Education Certified	1.8 (-4.2, 7.9)	0.556	587
	Reading Certification	-7.0 (-16.1, 2.0)	0.127	2,176
	Alternative Certification	-6.7 (-13.7, 0.3)	0.061	65

For the year 2005-2006 and students identified as being Emotionally Disturbed there were not enough teachers without a reading certification or with an alternative certification to analyze. The predictor variables teacher years experience, percent special education, and special education certification were analyzed, however, none of the results were significant.

For students identified as Specific Learning Disability a negative, significant result was found when these students were instructed in a classroom with a higher percentage of special education students (-1.0). The other predictor variables did not reach significance.

For students identified as Mild Mental Retardation none of the results reached significance.

Reading certification and alternative certification predictor variables both reached significance for the group Speech Language Impaired. Having a teacher with a reading certification resulted in a large, positive coefficient of 6.9 while having a teacher with an alternative certification resulted in a large, negative coefficient of -14.4.

Finally, for the group Other Health Impaired, none of the predictor variables reached significance. However, the predictor variable alternative certification was close to reaching significance and similar to the group Speech Language Impaired with a negative effect of -6.7.

The following table presents the final values obtained for the school year 2005-2006, for all students in the target content writing.

For the content of writing, none of the special education categories had any of the predictor variables reach significance. However, for the category of Speech Language Impaired, the predictor variable percent special education neared significance at 0.059.

Table 22: Writing 2005-2006 Predictor Variable Effects, All Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.1 (-0.2, 0.4)	0.462	585
	Percent Special Education	-0.6 (-2.1, 0.9)	0.423	585
	Special Education Certified	3.2 (-9.5, 15.8)	0.624	198
	Alternative Certification	-2.4 (-17.5, 12.8)	0.760	32
Specific Learning Disability	Years Experience	0.0 (0.0, 0.1)	0.246	9,813
	Percent Special Education	-0.4 (-0.8, 0.1)	0.083	9,813
	Special Education Certified	1.5 (-1.6, 4.7)	0.342	2,386
	Writing Certification	0.5 (-7.5, 8.5)	0.899	9,729
	Alternative Certification	-2.1 (-6.1, 1.9)	0.309	405
Mild Mental Retardation	Years Experience	-0.2 (-0.4, 0.1)	0.178	1,009
	Percent Special Education	-0.4 (-1.5, 0.8)	0.521	1.009
	Special Education Certified	5.3 (-4.0, 14.6)	0.266	395
	Alternative Certification	2.8 (-7.4, 12.9)	0.589	55
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.623	3,783
	Percent Special Education	-0.9 (-1.9, 0)	0.059	3,783
	Special Education Certified	-0.6 (-4.6, 3.4)	0.776	303
	Writing Certification	4.3 (-5.6, 14.1)	0.397	3,760
	Alternative Certification	5.0 (-1.7, 11.6)	0.141	91
Other Health Impairments	Years Experience	0.0 (-0.1, 0.2)	0.603	2,985
	Percent Special Education	-0.4 (-1.0, 0.3)	0.245	2,985
	Special Education Certified	-0.3 (-5.1, 4.5)	0.906	713
	Writing Certification	1.5 (-9.3, 12.2)	0.791	2,950
	Alternative Certification	-2.8 (-9.2, 3.7)	0.399	93

The effect, although not significant, was small and negative (-0.9).

Table 23 presents the final values obtained for the school year 2005-2006 for all students in the target content mathematics.

For the group Emotionally Disturbed there were not enough teachers without a mathematics certification to be analyzed for the study. There was a large, positive effect for students who were instructed by a teacher with an alternative certification (significant result). The other predictor variables did not reach significance.

For both groups, Specific Learning Disability and Mild Mental Retardation, none of the results reached significance.

A large, positive, significant effect was found for the predictor variable mathematics certification for the group Speech Language Impaired with a 14.9. This was the only predictor variable to reach significance.

For the group Other Health Impaired, being there was a modest, negative effect (-8.1) for students placed in a classroom with a higher percentage of special education students, which was the only significant result.

Summary. Table 24 presents the significant findings for all students in the school year 2005-2006 in all content areas.

The predictor variable percent special education resulted in negative findings for two of the groups. For Speech Language Disability, the content area certifications resulted in positive results. For the predictor variable alternative certification, mixed results were found. For the group Emotionally Disturbed, in the content of mathematics, a large, positive result was found, while for the group Speech Language Disability, for the reading content, a large negative result was found.

Table 23: Mathematics 2005-2006 Predictor Variable Effects, All Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.0 (-0.4, 0.3)	0.821	587
	Percent Special Education	14.7 (-1.1, 30.5)	0.068	587
	Special Education Certified	-10.9 (-24, 2.1)	0.100	193
	Alternative Certification	14.0 (1.0, 27.0)	0.035*	29
Specific Learning Disability	Years Experience	0.1 (0.0, 0.2)	0.104	9,799
	Percent Special Education	-2.3 (-7.6, 3.1)	0.406	9,799
	Special Education Certified	0.0 (-4.0, 4.1)	0.986	1,941
	Mathematics Certification	5.7 (-6.3, 17.6)	0.353	9,711
	Alternative Certification	-4.0 (-9.0, 1.1)	0.125	503
Mild Mental Retardation	Years Experience	0.1 (-0.3, 0.5)	0.631	968
	Percent Special Education	1.6 (-15.8, 19.0)	0.856	968
	Special Education Certified	1.6 (-11.6, 14.7)	0.816	373
	Alternative Certification	6.4 (-17.1, 29.8)	0.596	60
Speech Language Disability	Years Experience	0.0 (-0.1, 0.1)	0.484	3,660
	Percent Special Education	-6.7 (-15.7, 2.2)	0.138	3,660
	Special Education Certified	0.5 (-3.3, 4.2)	0.798	260
	Mathematics Certification	14.9 (0.7, 29.1)	0.040*	3,639
	Alternative Certification	-0.6 (-9.1, 7.9)	0.887	105
Other Health Impairments	Years Experience	0.1 (0.0, 0.2)	0.213	3,161
	Percent Special Education	-8.1 (-14.9, -1.4)	0.018*	3,161
	Special Education Certified	1.3 (-4.0, 6.5)	0.630	706
	Mathematics Certification	-1.4 (-14.7, 11.8)	0.832	3,126
	Alternative Certification	-4.3 (-11.1, 2.6)	0.223	146

Table 24: Significant Findings for All Students in 2005-2006 for All Content Areas

Special Education Group	Reading (coefficient)	Writing (coefficient)	Mathematics (coefficient)
Emotionally Disturbed	None	None	1. Alternative Certification (14.0)
Specific Learning Disability	1.Percent Special Education (-1.0)	None	None
Mild Mental Retardation	None	None	None
Speech Language Disability	1. Reading Certification (6.9) 2. Alternative Certification (-14.4)	None	1. Mathematics Certification (14.9)
Other Health Impaired	None	None	1. Percent Special Education (-8.1)

The following table presents the final values obtained for the school year 2005-2006, for the content reading and students in category three.

Only three predictor variables had enough cases to be analyzed for the group Emotionally Disturbed. None of the analyzed predictor variables reached significance although teachers' years experience neared significance with a small, positive effect.

Table 25: Reading 2005-2006 Predictor Variable Effects, Category 3 Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.2 (0.0, 0.5)	0.069	145
	Percent Special Education	-0.2 (-1.1, 0.8)	0.742	145
	Special Education Certified	2.9 (-4.2, 9.9)	0.425	61
Specific Learning Disability	Years Experience	0.0 (-0.1, 0.1)	0.757	3,063
	Percent Special Education	-1.0 (-1.5, -0.6)	<0.001*	3,063
	Special Education Certified	2.0 (-0.5, 4.4)	0.113	923
	Reading Certification	-5.4 (-12.2, 1.4)	0.118	3,029
	Alternative Certification	-0.9 (-4.2, 2.3)	0.564	128
Mild Mental Retardation	Years Experience	0.0 (-0.2, 0.1)	0.583	572
	Percent Special Education	-0.3 (-1.2, 0.7)	0.573	572
	Special Education Certified	1.3 (-6.4, 9.0)	0.746	271
	Alternative Certification	-0.9 (-8.7, 7.0)	0.829	30
Speech Language Disability	Years Experience	0.1 (-0.1, 0.2)	0.350	627
	Percent Special Education	0.0 (-0.9, 1.0)	0.935	627
	Special Education Certified	-3.2 (-7.9, 1.5)	0.184	57
	Alternative Certification	-2.0 (-7.2, 3.2)	0.458	29
Other Health Impairments	Years Experience	-0.1 (-0.3, 0.1)	0.200	640
	Percent Special Education	-0.1 (-0.8, 0.6)	0.772	640
	Special Education Certified	-1.7 (-6.7, 3.3)	0.506	200
	Alternative Certification	-2.1 (-9.3, 5.1)	0.568	31

Specific Learning Disabled students placed in a classroom with a higher percentage of special education students would be predicted to score a point lower on the test. This was the only coefficient that reached significance for this group.

Table 26: Writing 2005-2006 Predictor Variable Effects, Students in Category 3

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	0.2 (-0.2, 0.5)	0.334	206
	Percent Special Education	0.4 (-1.6, 2.5)	0.676	206
	Special Education Certified	-3.1 (-20.3, 14.1)	0.725	97
Specific Learning Disability	Years Experience	0.1 (0.0, 0.1)	0.195	3,472
	Percent Special Education	-0.3 (-0.7, 0.1)	0.147	3,472
	Special Education Certified	2.0 (-1.0, 5.1)	0.195	1,041
	Writing Certification	-4.2 (-13.3, 5.0)	0.371	3,440
	Alternative Certification	-1.8 (-5.4, 1.9)	0.344	171
Mild Mental Retardation	Years Experience	-0.1 (-0.3, 0.1)	0.465	627
	Percent Special Education	-0.8 (-1.7, 0.1)	0.075	627
	Special Education Certified	7.0 (0.2, 13.9)	0.043*	283
	Alternative Certification	4.0 (-7.2, 15.2)	0.480	40
Speech Language Disability	Years Experience	-0.1 (-0.3, 0.0)	0.131	541
	Percent Special Education	-0.4 (-1.5, 0.7)	0.434	541
	Special Education Certified	-3.1 (-10.1, 3.9)	0.383	51
Other Health Impairments	Years Experience	-0.1 (-0.3, 0.1)	0.184	748
	Percent Special Education	-0.4 (-1.2, 0.3)	0.262	748
	Special Education Certified	0.9 (-5.4, 7.1)	0.783	235
	Alternative Certification	4.0 (-2.7, 10.8)	0.240	37

For the remaining three groups, Mild Mental Retardation, Speech Language Impaired, and Other Health Impaired, none of the results reached significance.

Table 26 presents the final values obtained for the school year 2005-2006, for students in category three in the target content of writing.

For the special education category groups, Emotionally Disturbed, Specific Learning Disability, Speech Language Disability, and Other Health Impaired there were no results that were significant.

However, for students identified as Mild Mental Retardation, significance was reached for the coefficient special education certification with a positive coefficient of 7.0.

Table 27 presents the final values obtained for the school year 2005-2006, for students who scored one standard deviation or more below the mean in the target content of mathematics.

Only three predictor variables had enough cases to be analyzed for the group Emotionally Disturbed. Only one coefficient reached significance, special education certification. This coefficient resulted in a large, negative effect (-15.7) for students identified as being Emotionally Disturbed.

For the remaining four groups there were no coefficients that reached a significant level.

Summary. Table 28 presents the significant findings for students who scored one or more standard deviations below the mean on the test for the school year 2005-2006 in each content area.

Again, few predictor variables resulted in significant findings for Category 3 students for the school year 2005-2006. The group Specific Learning Disability revealed a negative coefficient for the variable percent special education. The group Emotionally Disturbed found a large, negative effect for the predictor variable special education certification, while the group Mild Mental Retardation found a large, positive effect for

Table 27: Mathematics 2005-2006 Predictor Variable Effects, Category 3 Students

Primary Disability	Variable	Coefficient (CI)	P-Value (* indicates significant effect)	Number of Teachers
Emotional Disturbance	Years Experience	-0.1 (-0.6, 0.3)	0.548	204
	Percent Special Education	17.0 (-0.6, 34.5)	0.058	204
	Special Education Certified	-15.7 (-29.6, -1.8)	0.027*	88
Specific Learning Disability	Years Experience	0.1 (-0.1, 0.2)	0.318	3,660
	Percent Special Education	-0.1 (-7.9, 7.8)	0.985	3,660
	Special Education Certified	-3.2 (-9.2, 2.9)	0.304	903
	Mathematics Certification	7.9 (-6.3, 22.1)	0.278	3,626
	Alternative Certification	-3.4 (-9.5, 2.7)	0.271	229
Mild Mental Retardation	Years Experience	0.1 (-0.2, 0.5)	0.411	708
	Percent Special Education	-1.6 (-17.4, 14.2)	0.845	708
	Special Education Certified	-0.4 (-12.1, 11.4)	0.953	293
	Alternative Certification	-7.2 (-29.1, 14.6)	0.516	43
Speech Language Disability	Years Experience	-0.1 (-0.3, 0.1)	0.338	594
	Percent Special Education	2.8 (-10.7, 16.3)	0.682	594
	Special Education Certified	0.7 (-6.5, 8.0)	0.843	42
	Alternative Certification	-1.0 (-14, 12)	0.884	30
Other Health Impairments	Years Experience	0.0 (-0.2, 0.3)	0.830	865
	Percent Special Education	-4.8 (-13.7, 4.1)	0.287	865
	Special Education Certified	-2.9 (-9.9, 4.2)	0.424	238
	Alternative Certification	0.4 (-11.2, 12.0)	0.941	48

special education certification.

Table 29 is a summary table and it presents all of the significant findings from all of the analyses. Each specific disability category is presented, by school year, and by

group (all students and category 3 students labeled as struggling students). The coefficient is also presented below the variable name.

Table 28: Significant Findings for Category 3 Students in 2005-2006 for All Content Areas

Special Education Group	Reading (coefficient)	Writing (coefficient)	Mathematics (coefficient)
Emotionally Disturbed	None	None	1. Special Education Certification (-15.7)
Specific Learning Disability	1.Percent Special Education (-1.0)	None	None
Mild Mental Retardation	None	1. Special Education Certification (7.0)	None
Speech Language Disability	None	None	None
Other Health Impaired	None	None	None

Table 29: Significant Findings from All Analyses

		Reading		Writing		Mathematics	
		All Students	**Struggling Students**	**All Students**	**Struggling Students**	**All Students**	**Struggling Students**
Emotionally Disturbed	2004-2005	Reading Certification (-9.4)	None	Years Experience (-0.3)	None	None	None
	2005-2006	None	None	None	None	Alternative Certification (14.0)	Special Education Certification (-15.7)
Specific Learning Disability	2004-2005	1.Percent Special Education (-7.4) 2.Reading Certification (-5.4)	1.Percent Special Education (-3.5) 2. Reading Certification (-3.2)	Percent Special Education (-4.1)	Years Experience (0.1)	None	1. Percent Special Education (-6.3) 2. Alternative Certification (6.5)
	2005-2006	Percent Special Education (-1.0)	Percent Special Education (-1.0)	None	None	None	None
Mild Mental Retardation	2004-2005	None	None	Percent Special Education (-11.2)	None	Years Experience (-0.3)	None
	2005-2006	None	None	None	Special Education Certification (7.0)	None	None
Speech Language Disability	2004-2005	Reading Certification (4.2)	None	Writing Certification (10.7)	None	None	None
	2005-2006	1. Reading Certification (6.9) 2. Alternative Certification (-14.4)	None	None	None	Mathematics Certification (14.9)	None
Other Health Impaired	2004-2005	Percent Special Education (-7.4)	None	None	None	1. Percent Special Education (-10.6) 2. Special Education Certification (7.9)	None
	2005-2006	None	None	None	None	Percent Special Education (-8.1)	None

DISCUSSION

The No Child Left Behind Act has put new demands on special education students and teachers by holding these students to the same standards as regular education students. Special Education students are now required to take high-stakes tests just like their general education peers. Also, inclusion and mainstreaming of special education students has become more popular in the schools. There are many variables that may affect a student's high-stakes testing score. One of the main focuses of this study was to determine the effect of special education students being placed in classes with a higher percentage of other special education students (special education classroom) versus being placed in a classroom with more general education students. There were only significant results for the special education groups Specific Learning Disability, Mild Mental Retardation, and Other Health Impaired. These results suggest that for these special education students, being placed in a special education classroom may not be as beneficial as being placed in a general education classroom, which also supports the intended benefits of inclusion to improve academic success as reported by Begeny & Martens (2007). The results were most consistent across years and content areas for the Specific Learning Disability group, which may suggest that this group would benefit most from inclusion. The results replicated four times for the school year 2004-2005 and twice for the school year 2005-2006. These results support what Rogers (1993) reported, with inclusion and supports being brought to the student in the general education classroom, students can benefit from remaining in the general education classroom.

Not only do classroom variables play a large part in student outcomes, but also so do teacher characteristics. A surprising result found from this study was that having a

teacher with a special education certification was not always beneficial. The effect for having a teacher with a special education certification was variable across special education categories. Specifically, the only groups that resulted in significant effects for special education certification were Emotionally Disturbed, Mild Mental Retardation, and Other Health Impaired. The results were positive for the groups Mild Mental Retardation (writing, category 3 for the year 2005-2006) and Other Health Impaired (mathematics, all students, for the year 2004-2006). It is interesting to note that the result for the Emotionally Disturbed group was a large, negative effect for those students who scored one ore more standard deviations below the mean in mathematics. It is unclear why these results were found for these specific students. Research examining special education certification programs should be investigated in the future. Many of our special education students are being instructed by special education certified teachers and thus it should be expected that these teachers are creating positive effects for the students.

Another interesting finding was that teachers' years experience resulted in some negative effects. These findings go against what most other research has found; that teachers with more years experience are generally more effective teachers (Ding & Sherman, 2006; Croninger et al., 2007). However, it is also important to note that the significant findings for teachers' years experience were few and those that were significant resulted in small effects. Again, the results were inconsistent across years, content domains, and special education category.

Another finding was that results were inconsistent for the predictor variable of content certification. It would be expected that students would perform better in a specific content domain when instructed by a teacher who has that specific certification. Specific Learning Disabled students and Emotionally Disturbed students, for the school

year 2004-2005, would be predicted to score lower on the reading test when instructed by a teacher with a reading certification. Similar results were found the Specific Learning Disability group in reading for students who scored one or more standard deviations below the mean. These findings go against the results found by Goldhaber and Brewer (1997), which found that subject-specific degrees result in better performance by students. However, their results were also found for general education, high school students. These results may suggest that subject-specific degrees are not as important for special education students. However, after further examination of the data, these teachers with content specific certifications that resulted in negative findings had higher percentages of low SES students than the other groups. Teachers who instructed the Specific Learning Disabled students and Emotionally Disturbed students had over 70% of their population who were low SES, whereas the other groups had only a 60% or lower population of low SES students. This difference in population of students may account for the negative results found.

These results suggest that teacher certification may not be a good predictor of teacher effectiveness. There are numerous possibilities why this may be the case. For example, the special education category may be very chaotic, especially for the Specific Learning Disabled group, which is heterogeneous in its self, and this may not lead for a good connection between teacher credentials and teacher effectiveness.

Even though negative results were found, there were also some positive findings. For the group of Speech Language Disability positive effects were found in the content areas of reading and writing for subject-specific teacher degrees. It is unclear why subject-specific degrees may be important for this group, but it would be interesting for future research.

The final variable examined was having a teacher with an alternative certification versus a regular certification. The USDOE policy argues that creating "fast track" programs with alternative certifications is a good method for producing highly qualified teachers. However, this policy is contradictory in meeting the No Child Left Behind definition of a highly qualified teacher because of the discrepancy in focus on coursework. The results of this study revealed that teachers with alternative certifications instructing special education students did result in positive effects, specifically in the content domain of mathematics. However, results were not consistent as significant, positive results were only found Emotionally Disturbed students for the school year 2005-2006 in mathematics and again for Specific Learning Disabled students who scored one or more standard deviations below the mean in mathematics in the year 2004-2005. There was one instance in which a significant, negative result was found. This was for Speech Language Disabled students in the content of reading for the school year 2005-2006. Again, the results are hard to interpret because of the lack of consistency.

Overall, one of the most consistent, critical findings from the study was that special education students appear to score lower than expected on the tests when placed in a classroom with a higher percentage of special education students, particularly for the Specific Learning Disabled group. The results from this study support what Begeny and Martens (2007) reported, which was that an intended benefit of inclusion is to improve academic success. Another interesting finding was that subject-specific certification resulted in some negative results. Again, this may be due to a number of factors. This certification may not be as strong a predictor as other variables, such as type of certification (general versus special education) and/or the percentage of special education students in the class. Other characteristics previously identified by Ding and Sherman

(2006) (years of teaching, major of undergraduate study, ACT or SAT scores, course work or degree obtained, quality of high school, earning of a license, and verbal ability) may be better at predicting student outcomes than subject-specific certification.

Also, this study demonstrated that using a value-added model was effective when examining effective teacher qualifications. As McCaffrey et al. (2003) explained, using Value Added Models allowed for the control of confounding variables, such as family background. Although most results were inconsistent, the results lead to some interesting questions that may be answered by future research, which are discussed below.

Limitations and Future Research

One major limitation of this study was the limited number of cases to be analyzed for each value of specific variables of interest. The small number of cases at specific values of specific variables may have contributed to inconsistent results, specifically for the variables of target content certification and alternative certification.

Another limitation of this study was the limited external validity. All data analyzed were specific to Louisiana. Other states data may display different findings.

Future research should target differences in training programs for general education versus special education certification programs. Special education certification programs should be creating teachers who are more effective with special education students. However, based on this study, this may not be the case. What qualities should be in special education certification programs should be studied.

The effects of inclusion should continue to be investigated. Based on this study, with the specific models used, inclusion for some special education categories is supported. In some instances, students who were placed in classrooms with a higher percentage of special education students resulted in large, negative effects.

More research should also be conducted regarding special education students and how teachers' years experience affects their performance. Most research suggests that teachers with more experience are more effective, but the results from this study were inconsistent with this. Is this just a finding specific to this study or may it be a finding common among special education students? More research needs to be conducted simply investigating variables that affect special education students' performance.

Future research should also focus on increasing the external validity of this study. Would other states high-stakes testing data reveal similar results? In finding out what other states data may reveal, this could lead to more positive findings on what specific factors affect a student's predicted achievement.

REFERENCES

Balboni, G. & Pedrabissi, L. (2000). Attitudes of Italian support teachers and parents toward school inclusion of students with mental retardation: The role of experience. *Education and Training in Mental Retardation and Developmental Disabilities*, 35, 148-159.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29, 37-65.

Begeny, J.C. & Martens, B.K. (2007). Inclusionary education in Italy: A literature review and call for more empirical research. *Remedial and Special Education*, 28, 80-94.

Berliner, D.C. & Biddle, B.J. (1995). *The manufactured crisis: Myths, fraud, and the attack on America public schools.* Reading, MA: Addison-Wesley.

Boe, E.E., Shin, S., & Cook, L.H. (2007). Does teacher preparation matter for beginning teachers in either special education or general education? *The Journal of Special Education*, 41, 158-170.

Brownell, M.T., Ross, D.D., Colon, E.P., & McCallum, C.L. (2005). Critical features of special education teacher preparation: A comparison with general teacher education. *The Journal of Special Education*, 38, 242-252.

Brucker, P.O. (1994). The advantages of inclusion for students with learning disabilities. *Journal of Learning Disabilities*, 27, 581-582.

Casbarro, J. (2005). The politics of high-stakes testing. Retrieved on October 25, 2007, from www.eddigest.com.

Croninger, R.G., Rice, J.A., Rathbun, A. & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review*, 26, 312-324.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Journal of Education Policy Analysis*, 8(1).

Deno, E.N. (1970). Special education as developmental capital. *Exceptional Children*, 37, 229-237.

Ding, C. & Sherman, H. (2006). Teacher effectiveness and student achievement: Examining the relationship. *Educational Research Quarterly*, 29, 39-49.

Dunn, L. (1968). Special education for the mildly retarded: Is much of it justifiable? *Exceptional Children*, 35, 5-22.

Elliott, J.L., Erickson, R.N., Thurlow, M.L., & Shriner, J.G. (2000). State-level accountability for the performance of students with disabilities: Five years of change? *Journal of School Psychology*, 39, 3-24.

Englehard, G., JR. (2004). Review of the Iowa tests of basic skills, Forms A and B. Mental Measurement Yearbook, 17.

Fuchs, L.S. & Fuchs, D. (1999). Fair and unfair testing accommodations. *The School Administrator*, 56, 24-30.

Gartner, A. & Lipsky, D. (1989). *The yoke of special education: How to break it*. Rochester, NY: National Center on Education and the Economy.

Gersten, R., Chard, D., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities*, 33, 445-457.

Goldhaber, D.D. & Brewer, D.J. (1997). Evaluating the effect of teacher degree level on educational performance. In W.J. Fowler (Ed.), *Developments in School Finance, 1996*. Washington, DC: US Department of Educatoin.

Goldhaber, D.D. & Brewer, D.J. (1998). When should we reward degrees for teachers? *Phi Delta Kappan*, 80, 134-138.

Goldhaber, D.D. & Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22, 129-146.

Harris, D. & Sass, T.R. (2006). Value-added models and the measurement of teacher quality. Retrieved on November 15, 2007, from itp.wceruw.org

Herman, J.L. & Golan, S. (1993). The effects of standardized testing on teaching and schools. *Educational Measurement: Issues and Practice*, 12, 20-25.

Heubert, J.P. & Hauser, R.M. (Eds.). (1999). *High stakes: Testing for tracking, promotion and graduation*. Washington, DC: National Academics Press.

Hollenbeck, K., Tindal, G., & Almond, P. (1998). Teachers' knowledge of accommodations as a validity issue in high-stakes testing. *The Journal of Special Education*, 32, 175-183.

Hoover, H.D., Hieronymus, A.N., Frisbie, D.A., & Dunbar, S.B. (1996). *Iowa tests of basic skills, Form M*. Itasca, IL: Riverside.

Horn, R.A., Jr. (2002). *Understanding educational reform*. Santa Barbera, CA: ABC-CLIO.

Idol, L. (2006). Toward inclusion of special education students in general education. Remedial and Special Education, 27, 77-94.

Junkala, J. & Mooney, J.F. (1986). Special education students in regular classes: What happened to the pyramid? *Journal of Learning Disabilities*, 19, 218-221.

Karagiannis, A., Stainback, W., & Stainback, S. (1996). Rationale for inclusive schooling. In S. Stainback & W. Stainback (Eds.), *Inclusion: A guide for educators* (pp.3-16). Baltimore: Brookes.

Klingner, J., Arguelles, M., Hughes, M., & Vaughn, S. (2001). Examining the schoolwide "spread" of research-based practices. *Learning Disability Quarterly*, 24, 221-234.

Koretz, D.M. (1995). Sometimes a cigar is only a cigar, and often a test is only a test. In D. Ravitch (Ed.), *Debating the future of American education: Do we need national standards and assessment?* (pp. 154-166). Washington, DC: Brookings Institute.

Lane, S. (2004). Validity of high-stakes assessment: Are students engaged in complex thinking? *Educational Measurement: Issues and Practice*, 23, 6-14.

Lewis, Anne C. (2005). Endless ping-pong over math education. *Phi Delta Kappan*, *86*, 420-421.

Linn, R.L. (2000). Assessment and accountability. *Educational Researcher*, 29, 4-14.

Lockwood, J.R. & McCaffrey, D.F. (2007). Controlling for individual level heterogeneity in longitudinal models with applications to student achievement. *Electronic Journal of Statistics*, 1, 223-252.

Louisiana Department of Education. (2006a). Louisiana Education Assessment Program: LEAP 2006 Technical Summary. Retrieved November 5, 2007, from http://www.doe.state.la.us/lde/saa/2273.html.

Louisiana Department of Education. (2006b). Louisiana Education Assessment Program: *i*LEAP 2006 Technical Summary. Retrieved November 5, 2007, from http://www.doe.state.la.us/lde/saa/2273.html.

McCaffrey, D.F., Lockwood, J.R., Koretz, D.M., & Hamilton, L.S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation.

McDonnell, L.M., McLaughlin, M.J., & Morison, P. (Eds.). (1997). *Educating one and all: Students with disabilities and standards-based reform.* Washington, D.C.: National Academy Press.

McLeskey, J. (2004). Classic articles in special education: Articles that shaped the field, 1960 to 1996. *Remedial and Special Education*, 25, 79-87.

McLeskey, J. & Waldron, N.L. (2007). Making differences ordinary in inclusive classrooms. *Intervention in School and Clinic,* 42, p. 162-168.

Mitzel, H.C. & Borden, C.F. (2000). *LEAP for the 21st century: 1999 operational final technical report*. Monterey, CA: CTB/McGraw-Hill.

Moody, S.W., Vaughn, S.R., Hughes, M.T., & Fisher, M. (2000). Reading instruction in the resource room: Set up for failure. *Exceptional Children*, 66, 305-316.

Murnane, R.J. & Steele, J.L. (2007). What is the problem? The challenge of providing effective teachers for all children. *The Future of Children*, 17, 15-43.

National Association of State Boards of Education. (1992). *Winners all: A call for inclusive schools*. Alexandria, VA: Author.

National Education Association. (January 2001). *Where the stakes are high for students.* NEA Today.

Noell, George H. (2006). *Annual report of: Value-added assessment of teacher preparation*. Retrieved, November 1, 2007 from, http://asa.regents.state.la.us /TE/value_added_model.

Noell, G.H. & Burns, J.L. (2006). Value-added assessment of teacher preparation: An illustration of emerging technology. *Journal of Teacher Education*, 57, 37-50.

Noell, G.H., Porter, B.A., & Patt, R.M. (2007). *Value-added assessment of teacher preparation in Louisiana: 2004-2006*. Retrieved November 1, 2007, from http://asa.regents.state.la.us/TE/value_added_model.

Norwich, B. & Lewis, A. (2007). How specialized is teaching children with disabilities and difficulties? *Journal of Curriculum Studies*, 39, 127-150.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237-257.

Odden, A., Borman, G., & Fermanich, M. (2004). Assessing teacher, classroom, and school effects, including fiscal effects. *Peabody Journal of Education*, 79, 4-32.

Popham, W.J. (1999). Why standardized test scores don't measure educational quality. *Educational Leadership*, 56, 8-15.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd Ed.). London: Sage.

Reynolds, M.C. (1962). A framework for considering some issues in special education. *Exceptional Children*, 28, 367-370.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (1998). *Teachers, schools, and academic achievement.* National Bureau of Economics Research, Working Paper 6691.

Rogers, J. (1993). The inclusion revolution. *Research Bulletin, Phi Delta Kappa*, 11, 1-6.

Sanders, W.L. & Rivers, J.C. (1996). *Cumulative and residual effects of teachers on future student academic achievement.* Research Progress Report from the University of Tennessee Value-Added Research and Assessment Center. Retrieved on December 1, 2007, from http://www.heartland.org/pdf/21803a.pdf.

Schiller, L. (2000). Politics, pedagogy, and professional development in Michigan. In A.A. Glatthorn & J. Fontana (Eds.) *Coping with standards, tests, and accountability: Voices from the classroom* (pp. 95-107). Washington, DC: National Education Association.

Schulte, A.C., Osborne, S.S., & Erchul, W.P. (1998). Effective special education: A United States dilemma. *School Psychology Review*, 27, 66-76.

Schulte, A.C., Villwock, D.N., Whichard, S.M., & Stallings, C.F. (2001). High stakes testing and expected progress standards for students with learning disabilities: A five-year atudy of one district. *School Psychology Review*, 30, 487-506.

Strauss, R.P. & Sawyer, E.A. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review*, 5, 41-18.

Sykes, C.J. (1995). Dumbing down our kids: *Why American children feel good about themselves but can't read, write, or add.* New York: St. Martin Griffin.

Tekwe, C.D., Carter, R.L., Ma, C., Algina, J., Lucas, M.E., Roth, J., Ariet, M., Fisher, T, & Resnick, M.B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29, 11-37.

Thurlow, M.L. & Johnson, D.R. (2000). High-stakes testing of students with disabilities. *Journal of Teacher Education*, 51, 305-314.

Thurlow M.L., House, A., Scott, D., & Ysseldyke, J.E. (2000). Students with disabilities in large-scale assessments: State participation and accommodation policies. *Journal of Special Education*, 34, 154-163.

U.S. Department of Education, Office of Special Education and Rehabilitative Services. (1992). *Fourteenth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: Author.

U.S. Department of Education. (2000). *Twenty-second annual report to Congress on the implementation of the Individuals with Disabilities Education Act.* Washington, D.C.: Author.

U.S. Department of Education. (2002). *Meeting the highly qualified teachers challenge: The secretary's second annual report on teacher quality*. Washington, DC: Author.

Wang, L., Beckett, G.H., & Brown, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. *Applied Measurement in Education*, 19, 305-328.

Washburn-Moses, L. (2003). What every special educator should know about high-stakes testing. *Teaching Exceptional Children*, 35, 12-15.

Wayne, A.J. & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73, 89-122.

Zlatos, B. (1994). Don't test, don't tell: Is "academic redshirting" skewing the way we rank our schools? *American School Board Journal*, 181, 24-28.

VITA

Bethany A. Porter grew up in Sturgis, Michigan. She obtained her Bachelor of Science degree in psychology from Jacksonville State University in December 2003. She received her Master of Arts degree in school psychology from Louisiana State University in December 2006. She is currently a candidate for the degree of Doctor of Philosophy in school psychology at Louisiana State University, where she works under Dr. George H. Noell. In addition, she is currently completing her internship at the Louisiana School Psychology Internship Consortium under the supervision of Dr. George Hebert. She is working as a school psychologist in public elementary schools in Jefferson Parish. After receiving her doctorate, she hopes to pursue postdoctoral work with children and adolescents in either a school or outpatient clinical setting.