THE EFFECTIVENESS OF START TRAINING ON BUILDING LEVEL PERSONNEL

by

DEBORAH ANNE DURRENCE

(Under the Direction of William W. Swan)

ABSTRACT

The purpose of the study was to determine the effectiveness of the StART training to assist administrators and teachers in 18 elementary, middle, and high schools in accessing instructional achievement data regarding current students, accessing historical data about groups of students and individual students, linking information from a variety of data reports, and drawing accurate conclusions regarding the data to guide instructional decisions. The study was conducted across schools in the K-12 continuum of public education.

The Student Assessment Reporting Tool (StART) is a web-based application that has been designed to allow teachers, administrators, and other education personnel to view the performance of students on previous assessments based on the current enrolled student population, as well as historical information on previous groups of students. The training was based on a model in which the researcher trained a leadership team (5-20 people) within the school that then redelivered the training to the other personnel within the building. The training was approximately one and one-half hours at each school location and provided those in the training with hands-on experience using the StART program, as well as instruction on assessment score interpretation, appropriate uses for data, and data analysis methods.

The sample consisted of 175 participants who were either teachers, administrators or other certified personnel in 18 schools in which teacher level access to StART was piloted. The

independent variables were level of education, role in the school, years of experience, school level, Title I school status and leadership tenure. The dependent variables included three sub scores and a total score obtained from the StART pilot school survey.

The results showed that participants who were trained by the researcher had statistically significantly higher scores in all 4 scores. In addition, those who were in high schools had a statistically significant score for Accessing StART. Those who had a specialist degree had statistically significant differences for Accessing StART and in the Total Score.

Index words: Data-driven, Assessment, Technology, Professional Development, Leadership, Disaggregated, Accountability, Staff Development, Evaluation, Teaching, Education, No Child Left Behind, Adequate Yearly Progress, Web-based

THE EFFECTIVENESS OF START TRAINING ON BUILDING LEVEL PERSONNEL

By

DEBORAH ANNE DURRENCE

B.S.Ed., The University of Georgia, 1998

M.Ed., Brenau University, 2001

A Dissertation Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF EDUCATION

ATHENS, GEORGIA

2004

THE EFFECTIVENESS OF START TRAINING ON BUILDING LEVEL PERSONNEL

By

DEBORAH ANNE DURRENCE

Approved:

Major Professor: William W. Swan

Committee: C. Thomas Holmes
C. Kenneth Tanner

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2004

ACKNOWLEDGEMENTS

First I would like to acknowledge and thank Dr. William Swan, my major professor. I am thankful for the hours of guidance and assistance during the process of writing this dissertation.

I appreciate the assistance and support from my co-workers in Gwinnett County, specifically Dr. Carrie Booher and Dr. Colin Martin. Most especially thank you to Linda Mitchell for your encouragement to persevere through this process. Your wisdom and knowledge were invaluable to me.

Also thank you to my parents for believing in me and supporting my efforts to complete this degree. Your encouragement to reach for my dreams helped me to achieve this goal.

Finally, I would like to thank my husband for understanding the hours of classes, studying and writing that created this. Your love, support and encouragement helped me to make it through this long process.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

CHAPTER 1

INTRODUCTION

The No Child Left Behind Act (NCLB, 2001) and the A Plus Education Reform Act (A Plus, 2000) in the state of Georgia require that schools provide evidence of student achievement for all students and their disaggregated subgroups on a continuing basis. This is a paradigm shift in school accountability from a focus on aggregated group performance to disaggregated subgroup performance and the required use of standardized assessments. All states were required under NCLB to create their own accountability systems under a consolidated plan that conforms to the guidelines created under NCLB. Performance for students on the Georgia Criterion Referenced Competency Test (CRCT) in elementary and middle school and at the high school level performance on the Georgia High School Graduation Test (GHSGT) are used to measure the adequate yearly progress (AYP) for all schools within the state of Georgia (Georgia DOE, 2003). Initially, beginning in the school year 2002-2003, student performance was reviewed in the content areas of English/language arts, reading, and mathematics. The areas of science and social studies will be added in later years. The goal of the legislation is that 100% of students in all subgroups will be performing at or above the minimum level in all content areas by the year 2014. All data are disaggregated (sex, race-ethnicity, socio-economic status, limited English proficiency, and disability) to determine the impact of instruction on measured subgroups consistent with NCLB and A Plus. To accomplish this task, school personnel must analyze and use the outcome data from assessments given to the students in their schools to determine the strengths and weaknesses in instruction.

The results from these assessments are being used to both grade and rank schools. Additionally, schools that receive Title I assistance funding from the federal government have sanctions attached to their AYP. Sanctions can include the requirement to offer school choice, supplemental educational services, and even the possibility of complete school restructuring, with the removal of all leadership personnel within the building (Georgia Consolidated Plan, 2003).

The leadership in Gwinnett County Public Schools(GCPS) has been developing tools and strategies to address an observed achievement gap between groups of students prior to the introduction of this legislation. One of the first innovations was the development of the evaluation system used to measure the effectiveness of educators, as well as schools. The Results Based Evaluation System (RBES) was created to provide a framework for schools and individual staff members to conduct action research related to specific curricular needs. Each school leadership team develops a Local School Plan for Improvement (LSPI), which includes goals for the entire school. Teachers than develop individual goals within their classrooms that support the overall school goal. The outcomes of these plans are used in the evaluation at the end of the school year for staff members.

In addition, Gwinnett County leaders have also researched school improvement efforts throughout the country that have been successful in closing the achievement gaps between groups of students. The model selected as most appropriate for the needs of the school district has been adapted from the model used in Brazosport, Texas. In Closing the Achievement Gap, Davenport and Anderson (2002) described an eight step method for identifying the weaknesses of students and then addressing those needs. GCPS has adapted the model to fit the curriculum of the system and has created a process called the AKS/CQI model. Academic Knowledge and

Skills (AKS) is the curriculum that is taught in Gwinnett County Public schools which is based on the Georgia Quality Core Curriculum (QCC). Continuous Quality Improvement (CQI) for all schools and students is one of the overall goals in the school system. Several schools began implementing this school improvement model during the 2002-2003 school year.

The steps for school improvement in the AKS/CQI model require that schools have access to data in addition to understanding the data for their students. The first step in the process is to disaggregate test scores and determine the strengths and weaknesses for the groups of students in a school by grade level. Once the areas of weakness are identified, a target calendar is created that provides time for additional instruction in areas of weakness, as well as enrichment for those students who have strengths in the area being reviewed. Target lessons are given in one or two week time periods on the focus objectives. At the end of the target time period, students are given mini-assessments covering the focus objective. Those students who are still struggling with the concept then receive additional instruction while those students who have mastered the concept receive enrichment. Throughout the year students are provided with practice materials in the target areas to maintain their knowledge-base and the overall performance is monitored throughout the school year.

The process for examining all of these data has been both new and overwhelming to many educators. Although the concept for using assessment data for improving student achievement was certainly not invented by NCLB, the accountability system that accompanies this information is new (Johnson, 1997). Mertler (2002) explained that understanding and making use of the large amount of information available from assessments is often a daunting task for classroom teachers. Many leaders in education are beginning to see the potential of

using the standardized assessment data to improve student achievement and not just measure school accountability (Educational Research Service, 2001).

In order to address this concern in Gwinnett County Public Schools, the Student Assessment Reporting Tool (StART) was developed to provide school administrators and teachers with access to test score data for the students currently in their schools that can be used to guide their instructional focus. In addition, in order for schools to accurately construct their AKS/CQI objectives for the school year, accessible assessment data were needed. StART is a web-based application that has been designed to allow teachers, administrators, and other education personnel to view the performance of students on previous assessments based on the current enrolled student population, as well as historical information on previous groups of students. The program was constructed utilizing input from educators within the classroom, central office administrators, and technology personnel in order to best customize the program for the needs of those using the information. The reports in the StART program provide both aggregated and disaggregated information based on where the students were enrolled when they took the assessment and where the students are currently enrolled. There are some schools within GCPS in which almost 40 % of the students who are enrolled in their schools during the spring will not be enrolled the following year. By creating reports several times a year based on the current enrollment of students, StART provides more current information that can be used to make more informed decisions regarding appropriate instruction for students.

In order for administrators and teachers at schools to effectively use the StART program, training must be provided that addresses both the technological use of the tool and training related to using data for instruction. The training was based on a redelivery model in which a trainer worked with a leadership team (5-20 people) within the school that then redelivered the

training to the other personnel within the building. The training was approximately one and one-half hours at each school location and provided those in the training with hands-on experience using the StART program, as well as instruction on assessment score interpretation, appropriate uses for data, and data analysis methods. Each leadership team was given a CD-ROM which contained the PowerPoint presentations that could be used with the teachers for training, instructional videos that could be shared with the staff on how to access reports, a data guide on which reports to view in order to find the assessment data most useful to the staff, and guides for using the StART information to create the RBES goals for the teachers. Without the ability to analyze and use the data provided in the program, teachers would be unable to understand how to link the data to the instruction in their classrooms.

Currently in GCPS, 18 elementary, middle, and high schools have teacher level access to the StART program. Data from this study can be used for evaluating the successes and failures of the training which would allow for retraining when necessary and allow for future refining of the training program in order for participants to gain greater understanding in the use of data. The goal is to eventually have access to the StART program for all teachers in the 59 elementary, 16 middle, and 15 high schools. Also, there are seven special entities within the school system including two alternative schools, a vocational and technical school, a special education pre-kindergarten center, a special education center, an online campus, and a psychoeducational facility. The school system will add an additional four elementary schools, four middle schools, one high school, and an additional vocational and technical school for the 2004-2005 school year.

Importance of This Study

Without a clear understanding of data and the use of the StART program, the eventual link to increasing student achievement may not take place. The increasing expectations of society, business, and legislation (NCLB & A Plus) require that schools examine the effectiveness of programs concerning student achievement. This research should provide valuable information regarding the effectiveness of the training program that is being used with the pilot schools and will eventually be used with all schools in the system. The information can be combined with the other activities in Gwinnett County Public Schools to further enhance the continuous quality improvement process.

Problem

With an increased emphasis on the need to use data-drive decision making, the problem of the study was to assess the perceived gap in information regarding analysis of disaggregated student data using the StART program. Participant performance will be analyzed in the areas of ability to access the StART program, to analyze assessment data, to draw conclusions based on data presented, and to gain an overall understanding of the training program.

Purpose of the Study and Research Questions

The purpose of the study was to determine the effectiveness of the StART training to assist administrators and teachers in 18 elementary, middle, and high schools in accessing instructional achievement data regarding current students, accessing historical data about groups of students and individual students, linking information from a variety of data reports, and drawing accurate conclusions regarding the data to guide instructional decisions. The study was conducted across schools in the K-12 continuum of public education.

The research questions addressed in this study were as follows:

1. Are there significant differences in understanding and use of the information based on the education level (Bachelors, Masters, Specialist or beyond) of the participants, as determined by performance on the StART Pilot School Survey?

2. Are there significant differences in understanding and use of the information based on the role of the participants within the school (administrator, teacher, etc.), as determined by performance on the StART evaluation survey?

3. Are their significant differences in understanding and use of information based on the years of experience in education of the participants, as determined by performance on the StART Pilot School Survey?

4. Are their significant differences in understanding and use of information based on the school level (elementary, middle or high school) of the participants, as determined by performance on the StART Pilot School Survey?

5. Are there significant differences in understanding and use of the information based on the Title I status of the school, as determined by performance on the StART Pilot School Survey?

6. Are there significant differences in understanding and use of the information with a new leadership team as compared with an established leadership team, as determined by performance on the StART Pilot School Survey?

Limitations of the Study

1. The trainer had no control over the exact format of the re-delivery of the training for the teachers.

2. Only some of the schools in the pilot program were implementing the AKS/CQI improvement model.

3. Technological expertise varied among the schools based on the level (elementary, middle, or high school) and the age of the school.

4. Only schools in Gwinnett County Public Schools were used in the study.

5. A pre-test was not used with the StART Pilot School Survey so information gained from analysis of results is limited due to lack of historical knowledge.

Organization of the Remainder of the Study

Chapter 1 provided an introduction to the StART program and training model, the importance of the study and the purpose of study and research questions. Chapter 2 contains a review of research which includes the use of data to drive instructional decisions, accountability systems, delivery of technology staff development, and effective evaluation of staff development. Chapter 3 contains the research design, instrumentation, data collection, and statistical analysis. Chapter 4 contains the results and findings for the study. In Chapter 5, a summary, conclusions, implications and recommendations for further research are included.

CHAPTER 2

REVIEW OF RELATED LITERATURE

Introduction

Using a technology tool to enhance the knowledge and skills of education personnel has a variety of challenges that must be reviewed to determine the effectiveness of the implementation. Components in the evaluation of the program training include the use of data to drive instructional decisions, the use of accountability systems to chart education reform, the implementation of effective technology-based staff development, and research-based evaluation of staff development programs. Due to the novelty of the concept of combining the elements within the training program, this review revealed there is little research that exists combining the various elements of the program.

Use of Data to Drive Instructional Decisions

Studies have been conducted over the last decade or more that target the concept of using assessment results to drive instructional decisions. Vitali (1994) found that teachers more often would teach to the type of question that was used on the test rather than the information regarding the content measured on the assessment. Teachers had a variety of perceptions related to the use of those assessments and the researcher determined that the concept was rather complicated.

Later research indicated that there were positive effects that came from teachers utilizing assessment data to determine the effectiveness of their instruction and the impact on student learning. Johnson (1997) found that schools and school systems had such a large amalgamation of data that the key to using data effectively was to organize the information and determine

which pieces of data could best be used to improve the learning of students. An emphasis is placed on teachers becoming researchers or analyzing the data about the students that they teach so that decisions can more directly reflect the needs of their students. The key is to find methods for analyzing data by the school administrators and teachers and helping them to make meaning out of the massive amounts of available information.

Kinder (2000) provided a guide for districts to use when implementing a process of data-driven decision making. The use of data should provide a support system for making decisions about students. The data used in the process should include indicators that reflect whether students are learning the curriculum presented in the classroom. Consistently using the same indicators is a key element in effectively using data to make instructional decisions. Although specific methods for storing the data were not discussed, the researcher indicated that creating a collection of data that could be later reviewed was a vitally important step.

In a brief from Education Research Service (2001), successful uses of data to make instructional decisions were reviewed. In the Twin Falls School District in Idaho, the superintendent stated that each successive group of students was doing better because teachers were using the information from previous student achievement data to improve their instruction. The Pine Springs Elementary School staff in Fairfax, Virginia, analyzed the past performances of their current students and were able to improve instruction resulting in student achievement gains in just one year. The Umatilla-Morrow Education Service District in Oregon developed a technology system to collect and analyze student assessment data that were linked to specific curriculum standards. The data were kept on a student-by-student basis so teachers were able to intervene with small problems before issues escalated with gaps in student learning.

As the concept of using data has become more wide-spread, success stories from schools that used data to make decisions began to surface. Liddle (2000) found that an elementary school faculty in Colorado was able to use their data to determine what areas needed improvement and a significant increase in student scores on the Colorado State Assessment was observed. Cooper and Cromey (2000) reviewed a pilot program in which both teachers and students tracked achievement data daily to determine the extent of learning that was taking place related to specific content areas. The researchers found that one of the key elements was to make sure that the data being used when making decisions were the most appropriate data for the situation. Davenport and Anderson (2002) found that the Brazosport Independent School District in Texas was able to show such dramatic results in increased achievement that their model of data analysis has been published in a book and is being used by districts throughout the country to guide their efforts.

Wade (2001) found that student assessment data was one of the four types of data that should be included in the profile of a school. He found that data from assessments could reveal the strengths and weaknesses of programs and students to determine changes that could best allow schools to meet their goals of increased student achievement. Guidance can be given related to the areas in teaching that should be changed to maximize student learning.

Levine (2002) reviewed the process that school districts followed when they begin to want to assemble their data sources into one centralized source. Many school districts have discovered that technology could be used to more easily analyze their data. Most districts find that 6 to 10 database sources are combined when constructing a district-wide data warehouse. One of the processes when using a data warehouse that is challenging is cleansing the data. This process involves removing duplicate information and inconsistent information that could

possibly corrupt reports built from the data warehouse. Depending on the size of the district, some districts may find software created by a vendor that works well for their system while larger districts may need to work directly with a developer due to the large volume of data within their system.

Other studies have shown that test data from standardized assessments can provide information that can be used to make instructional decisions for both groups of students and individual students within a classroom. Mertler (2002) found that there were several areas in which information could be obtained by reviewing assessment data. Teachers can use subtest performance to determine if problems exist between the assessment and alignment with the curriculum, if the teaching method for a particular concept is not enabling students to learn the intended skill, or if the students are not being assessed in the classroom in the same manner as the standardized assessment. Teachers could prioritize the areas of concern based on subtests in which the student performance did not meet expectations. For individual students, interventions can be designed that address the needs as indicated on the test reports for those students.

Decker (2003) reviewed information from the Lead Mine Elementary School in Raleigh, North Carolina which began using data in 1998 to determine the links between the current curriculum and student performance on assessments. The school used a four step process, including curriculum alignment, curriculum mapping, benchmarking, and differentiation, to examine the curriculum and determine the areas of change that were needed to increase student achievement. The school then chose to use a program called SuccessMaker to forecast the student performance on the upcoming assessments and provide additional assistance to those students who were in need. Teachers were able to use the results to provide ongoing intervention to students throughout the school year.

Although many educators view the use of data as only an accountability measure, the use of data can help educators better serve the needs of their students. Using data is often seen as a negative because the data may not be viewed as timely or accurate and therefore useless. Doyle (2003) advocated a three dimensional approach to data analysis with the first component being the construction and maintenance of a data warehouse with "decision-support tools" for educators to use to effectively make sense of the data. Hardy (2003) found that many software companies were working to create programs that would address the data needs that school systems have based on the requirements of NCLB. The need to disaggregate student results and match those results to where the students are currently enrolled provided another challenge to the software companies. Providing accurate information that can be used to assist student achievement in elementary school before problems grow in later years was seen as a primary focus of data analysis.

Sharkey and Murnane (2003) provided recommendations and noted challenges associated with using student assessment data. While working with the teachers in the Boston Public Schools, these researchers found that often teachers are trained on how to access the assessment data, but are not trained on what the results mean and how to connect that information to instruction. There was also wide variation in the teachers' perceptions regarding the use of the data. Those teachers who did work together to make meaning of the students' data often began comparing results and decided to work interdependently. Through the knowledge gained in their research, several recommendations surfaced regarding the use of technology to access student data. The data should be stored on some type of district database that is updated frequently to compensate for student mobility. In addition, there should be discretion used when balancing access to student information and the privacy of both student and teacher data. Finally, the

system should be easy to manipulate with information available for the students who are currently in the teachers' classes and scores from the previous year.

The Education Commission of the States (2002) released an issue brief on data-driven decision making in relationship to NCLB. A review was completed of six school districts across the country and the use of data in those districts. One of the exceptional districts used web-based assessments to measure student achievement and the results were available almost immediately. This access to data allowed decisions to be made quickly. The results were used to provide intervention and assign students to appropriate classes. The districts that were most successful in providing effective interventions had significant central office support in using the data from student assessments. Intervention impact was measured using student achievement from the students who participated in the intervention activities. School improvement plans that were driven by data also led to greater student gains. As technology allows for greater and more efficient access to student data, the power of decisions based on the student data will continue to increase.

Accountability Systems and School Reform

Accountability movements have come and gone through the 20th and into the 21st centuries in education. Pearson, Vyas, Sensale, and Kim (2001) reviewed the progression of movements. The first major movement came in the 1920s with the push to use the scientific method and the implementation of multiple-choice format tests for the first time. In 1965 with the passage of the Elementary and Secondary Education Act (ESEA, 1965), the federal government began to require schools that received Title I assistance funding to provide data to document student achievement. In the 1980s the government began to assert that schools were failing students and new goals should be set to improve student achievement. The standards

movement forced the development of more assessments. By the early 1990s many in education still viewed student testing as a method of teacher accountability and nothing more. Several states had begun to push reform movements using assessments including Texas, Michigan, and Illinois. The end result for the researchers in this study was that the link between assessment and accountability systems should be carefully examined to determine if the intended outcomes are truly the experienced outcomes.

Accountability systems have existed for many years in states across the nation. The new emphasis on accountability is coming from a federal mandate (NCLB) to create structured accountability systems that meet federal guidelines in each state. Blum (2000) examined the school improvement efforts in Reynolds School District in Oregon. The data collected were mostly qualitative with interviews of the central office staff regarding quantitative data used in the school improvement process. The state of Oregon began using an accountability system based on standards during their Competency-Based Education (CBE) movement in the 1970s. In the mid-1980s, Oregon began to implement the Oregon Action Plan for Excellence that set a common set of curriculum goals and learning standards for the state. In the late 1980s Oregon passed the 21st Century Schools program which allowed schools to apply for waivers from the standards if the schools were experimenting with innovative instructional strategies. State assessments measuring mathematics, reading and literature, writing, and science were implemented in the late 1990s in order raise academic standards and a timeline was established for creation of assessments for the other areas of the curriculum. In addition to these statewide reform efforts, Reynolds school district has participated in the Onward to Excellence program that helps schools create plans for continuous improvement for all students. There were several impacts of the joint reform elements used in the Reynolds school district. One of the reforms

implemented was teams of language arts, science, math and social studies teachers at the 9th and 10th grade levels to allow for more flexible planning and cross-discipline assignments. From 1996 to 1999, the percentage of 10th grade students meeting or exceeding the standards in reading increased from 36% to 61% and the percentage meeting or exceeding standards in writing increased from 62% to 82%. A passing score on the state assessments in the 10th grade was required for graduation and the students in the Reynolds school district have made significant progress. While these changes occurred, the percentage of students on free and reduced lunch increased by 3%, the number of students in the ESL population increased by 85%, and the percentage of students who lived in poverty increase to near 37%.

Many researchers have reviewed the success of accountability systems in terms of student achievement. Elmore and Fuhrman (2001) focused on the accountability systems in 20 schools including public, charter, and independent schools. Also, there was a follow-up examination of 12 of the schools that were in 4 states with varying accountability systems, and an examination of how various accountability structures worked simultaneously from the local level through federal mandates. Finally, there was a 50 state survey regarding accountability, assessments and performance reporting. As of 1999-2000, 12 states tested students in grades 2 or 3 until grade 8 every year in the same subject areas (Goertz & Duffy, 2001). In Kentucky and Charlotte-Mecklenberg schools, the teachers reported satisfaction in the recognition they gained for achieving high test scores. In Vermont, New York, and Kentucky, the accountability system was credited by teachers for helping to expand the curriculum in the areas of mathematics and problem-solving. Some schools in San Francisco and Chicago that were under the most severe sanctions due to low test scores, were not making fundamental changes to their schools and the results have not changed. In the eight state in-depth studies, the teachers reported in all eight

states that they "felt that schools were held more accountable (than teachers)" (p. 70). Teachers stated that there was "little influence" (p.70) that they had over students in terms of higher achievement when the rewards or sanctions were based on the entire school. However, the principals in that same study indicated that they held the teachers most accountable for student performance and not the school as a whole (Elmore & Fuhrman, 2001). With the implementation of NCLB this information has become much more critical, specifically for those schools with students who have achieved lower test scores.

Some accountability systems require that schools that do not meet the student achievement benchmark standards write school improvement plans with specific areas in which the school teachers will work to increase student achievement. Mintrop and MacLellan (2002) examined mandatory school improvement plans that were created due to low performance on high stakes assessments in Maryland. The study examined the content of the school plans as well as the role that the plans played in the school improvement process. Plans were examined from 46 schools and an in-depth qualitative investigation was completed in three elementary schools and four middle schools. The information from the seven schools was collected through interviews with teachers and administrators in the schools. The researchers found that a small group of administrators and a few teachers wrote the plans and the district and state officials oversaw the final version of the plan. The research data indicated that the plans showed significant alignment with the state goals. The faculty and staff within the school who were given the responsibility of implementing the plan knew the minute details, yet most classroom teachers knew little of the plans. In the schools where the principal supported the plan, the faculty was forced to comply. However, in most schools the plan had little impact on the workings of the schools.

Finn (2002) described three basic accountability ideas that have been used in the United States with the intention of improving education. The basis for the use of these accountability systems, in his opinion, came from the information released in the "Nation at Risk" report from the National Commission on Excellence in Education in 1983 (as cited in Finn, 2002), the information from highly successful schools, and the increased efficiency in business. The first accountability system described was based on the concept of trusting the educational professionals within the schools to make decisions that were in the best interest of children. The second method was dubbed the trust, but verify method. This is the basis for NCLB, in that schools are required to verify that students are achieving at a certain level, but have the flexibility to make decisions as long as the students continue to achieve. The final accountability system is based on the idea of market-style accountability in that the customers (parents) will chose the best product (school) and those schools that are not successful will cease to exist. None of the strategies has proven to have greater results than the others and accountability systems will probably continue to evolve and change over time.

Linn, Baker, and Betebenner (2002) reviewed the implications for school systems under the passage of NCLB. First, all states were required to develop content standards and assessments for reading and mathematics in grades three through eight. States were given until the 2005-2006 school year to fully meet this requirement. Adequate Yearly Progress (AYP) objectives must be set for each state with the goal that all students meet the proficient or higher categories by the 2013-1014 school year. Progress must be reported by school for all subgroups of students and schools that do not make AYP for two consecutive years will be placed in school improvement. States can aggregate up to three years of data in the determination of AYP. States are able to set their own benchmarks and standards for performance levels, which vary greatly

among the states. For example, in 2001 on the grade three mathematics assessment only 39% of students in Mississippi scored proficient or higher, while the percentage was 7% in Louisiana, and was 92% in Texas. Based on this information, the extent of improvement needed to meet the federal accountability guidelines within 12 years varies greatly in these three states. School level results can also show large variability due to the use of different cohorts of students to measure student achievement annually. In order for states to meet the overall assessment goal of NCLB, some states may choose to lower their proficiency levels due to the large gap between current achievement and the goal. Although the goals of NCLB are commendable, the specifics of the law and the accompanying timeline may produce some undesirable results or unanticipated consequences.

An additional component that adds to the complexity of compliance with NCLB is the new technology requirement. Lohr (2003) reviewed the process that Wisconsin has used to ensure that schools are effectively using their technology resources. A process was created to ensure that the technology purchased through the Title II funds was used to increase student achievement. The initial results indicated that the process used is assisting in school improvement efforts through the use of technology. Levine (2003) found that the national average for collecting, processing, and reporting data per student, as required by NCLB, was 40 minutes. If a district has just 2500 students, that equates to 1500 hours. The US Department of Education plans to decrease the time by improving the technology that states use for reporting and in turn helping the individual school districts. The goal of the use of technology would be to provide teachers with data that could be used to help all the individual students in their classroom rather than simply complying with the NCLB accountability system.

Goertz and Duffy (2003) described the various accountability systems that were used in all 50 states prior to the enactment of NCLB. There were 48 states prior to 2001 that had utilized some type of state-wide assessment in both reading and mathematics and the two additional states, Iowa and Nebraska, required local districts to measure both of those areas. Only thirteen states and the District of Columbia were testing consecutive grade levels using the same assessments as NCLB would require. All 50 states had some type of public report card that was produced either at the local or state level that listed the performance of students in the school. There were 13 states that used the report cards as the only form of accountability, while others used a variety of measures including absolute target goals, annual growth goals, or a reduction in the number of students in the lowest achievement levels. All states had to re-examine the systems in place and determine the changes in both assessments and accountability that are necessary to meet the requirements of NCLB.

Administrative Technology Use

Carrying out the administrative tasks within a school has been changed by the increased use of technology in schools. Telem (1999) reviewed the impact of the increased use of technology and the changing dynamics in the school leadership. The study was based on a case study of a high school and the change in the roles of the department heads in that school. The basis of discussions and interactions was increasingly more dependent on the reports created using the new technology system. The feedback provided was thought to be more cohesive and the issues were more thoroughly reviewed than prior to the implementation of the technology system. There were more meetings held among all groups in the school due to the increased amount of data provided by the technology system. While, the researcher indicated that although

these results are based on only one high school, the example indicated a changing role of all administrators and educational leaders in the school.

Ausbrooks (2000) reviewed how technology has changed the format of schools in addition to the changes that administrators should make in their activities using technology. Schools are now implementing virtual schools and additional distance education opportunities for students. Administrators would be able to have more time to interact with the teachers in their buildings if administrative work is more efficient through the use of technology. Using technology to review data used in accountability systems can assist administrators in making decisions. Connectivity and communication can be increased through the efficient use of technology tools. More individuals will be involved in the decision-making process due to the increased spread of data. The focus is shifted toward making decisions based on the data, rather than collection or management of the data.

Technology leadership within a school is essential for successful technology innovation implementation. Geer (2002) stated that there were eight essential areas of knowledge that school leaders need to be proficient in order to serve as the technology leaders within their building. Participating in teacher training is a crucial component in the implementation of a successful use of technology. Often leaders can be used as the trainers or mentors in a training program. Bradshaw (2002a) examined how leaders' technology use can impact school reform. The researcher indicated that a large knowledge and skill base must be developed in order for technology to be used effectively in schools.

Information systems are an essential component of decision-making in schools. Kalay and Chen (2002) reviewed the use of student information systems to increase the effectiveness of school management. The study was based on a two year review of a high school in Israel that

implemented a student information system to make decisions related to student assignments and learning activities. The results of the study indicated that there were improvements in failure rates, dropout rates, and student achievement with the use of the technology system to collect data regarding students. The tools assisted the administrators in creating a more complete student achievement picture by using the technology tools rather than individual data components available prior to the system.

Dawson and Rakes (2003) reported on the effect of principal technology training on the use of technology by teachers. As the instructional leader of the school, the principal should have a commitment to learning new skills that can impact instruction. A total of 398 principals surveys were included in the study and were chosen from those principals who used the internet across the United States. The results indicated that principals received more training on how to integrate technology into the curriculum than any other type of training. The age of the principal was a significant influence on the effectiveness of technology integration with the younger principals having higher rates of technology integration. Principals were more effective in using technology when their training was sustained over time and support was provided.

Professional Development

Guidelines have been provided by numerous organizations with the components of effective professional development. Guskey (1995) outlined six guidelines that were essential components for a successful professional development endeavor. Among the guidelines is the concept of follow-up with support provided and encouragement given to continue applying the ideas learned during the staff development activity.

Increasing student achievement is one of the ultimate outcomes of any educational professional development endeavor. Fisher (2001) focused on a professional development

system that was used at a high school in San Diego, California, that targeted specific instructional techniques and areas of concern. The plan included targeting specific instructional strategies that could improve student achievement, giving teachers a choice of formats for the professional development, creating teacher peer groups that work together to implement the changes, and including administrative participation in the project. The school had 1900 students and 100% of the students received free or reduced lunch. In addition, 4% of the school was white with 54% Latino, 21% African-American, and 20% Asian. The master teachers worked as mentor teachers with small groups of teachers and gave support during the implementation of the new strategies. All administrators participated in the monthly seminars and worked with specific departments on implementation. After two years of implementation, student reading scores had increased by 12% and the school was meeting the accountability standards for the first time in 15 years. Having leadership within the building support the professional development was cited as a key component to the success of the staff development.

In 2001 the National Staff Development Council (NSCD) revised their standards for staff development with a new focus on students. The new standards are categorized as context standards, process standards, and content standards. A focus was placed on using disaggregated student data to assist in continuous improvement in student achievement. Prior to this revision, much of the literature on staff development focused on participant perceptions related to staff development effectiveness, rather than on the impact on student achievement.

Holland (2001) explored the effect of professional development in technology on school reform efforts. Professional development for technology should not be a one time event with no follow-up training or discussion because are likely not to apply the skills they learned during the in-service training. In addition, this study examined the impact of technology training on

teachers' willingness to participate and lead reform efforts within their schools as they reviewed how to best utilize the technology in their school. The results indicated that teachers needed human support and frequent contact when integrating technology into their daily activities. In addition, teachers were willing to take more risks with technology when frequent support and follow-up was provided.

Cooley (2001) analyzed the use of the teachers as trainers model to implement the use of technology for two programs in Westfield Washington Schools in Westfield, Indiana. There were four steps in the model used for instruction for these two programs. A needs assessment was completed which was followed by selection and planning for the core group of instructors. The third phase was the actual delivery of the technology training, and finally an evaluation was completed. Each phase had several components in which the groups involved could personalize the activities to fit the needs of the group being trained. The findings indicate that there were significant time and cost savings associated with the use of this model. In addition, the participants provided evidence during the evaluation process that the skills were implemented in their practice. The ongoing support and resources were cited as key factors to the success of the program versus the use of a single event training session.

Whittaker and Young (2002) examined the effects on teachers who created their own curriculum-based assessments to discover if there was an impact on teacher understanding of the use of assessment for instruction when compared to the current accountability system in place. The research centered on a case study of a school district in the San Francisco Bay area that had a student population of over 3,800 students k-12, which was described as a large urban school district. The school district formed a partnership with local businesses to create a hands-on learning environment with the goal of improving math, science, and technology knowledge of

students. A group of teachers was chosen to create a new curriculum for this activity that matched national and state standards. In addition, the group was also told to create an assessment to measure the students' skills and a rubric for scoring the assessment. While these teachers were working on this project, the state of California was also implementing a state accountability system using nationally and state norm-referenced assessments. Much of the information used to evaluate the research questions was obtained from qualitative interviews conducted with the teachers involved in the project. Three of the teachers stated that they learned that students could create evidence of knowledge gained through authentic assessments and yet the teachers were also frustrated with knowing that the results of the norm-referenced tests would be used to measure student achievement. The results of the study indicated that the teachers' impressions were that the low-level skills required to be successful on the norm-referenced tests were often in conflict with encouraging students to perform higher-level thinking tasks. However, the teachers also stated that they gained more knowledge in accurately assessing student knowledge and encouraging higher-level thinking skills by creating curriculum-based assessments.

Jenson, Lewis, and Smith (2002) reviewed the use of three methods of technology training implemented in Canada. The methods included use of technology training provided by a university, a district-based training program during the summer, and school-based technology instruction. The results indicated that there are several key factors to consider when implementing a technology training program. Providing incentives in the form of time or money was one of the key success elements. Also, the flexibility for variations in ability levels and time to learn independently were important components of the successful programs. Ongoing support, as well as, school site learning added to the utility of the training. Teachers were not

successful in implementing knowledge and skills provided in the single session stand-alone training programs.

The success of professional development activities within an organization is dependent on the role and attitude of the leadership. Specifically within a school, the role of the principal is critical in the success and teacher buy-in achieved in the effort. Zimmerman and May (2003) examined some factors that can inhibit professional development from becoming an effective activity. The researchers found that the various demands on the principals do make participation in staff development activities difficult. In addition, staff development that takes place in a one time event with no follow-up or evaluation is often ineffective. Principals cited a lack of funds in combination with the time demands of teachers as the two primary factors that inhibited successful staff development. Professional development can play a crucial role in the effort to increase student achievement; however, a concerted effort must be put forth in terms of making the staff development an effective experience.

Often, for the reasons stated by previous studies, teachers do not implement the strategies learned in professional development activities in their classrooms and make transfer of those skills into practice. Spencer and Logan (2003) examined a Research Lead Teacher (RLT) model that provided ongoing support for implementation of staff development ideas to determine if there was a measurable change in practice for this model. The results indicated that there was a greater positive impact of instructional change for the teachers that had follow-up with the RLT model than those who participated in a traditional one sitting staff development activity. Although the study had numerous limitations, including a small sample and self-selection for the intervention, the results support the concept of providing follow-up and examination of training effectiveness.

Providing staff development to groups of education professionals that includes the use of new or different technology tools presents a set of challenges. Poplin (2003) reviewed models for the Technology Innovation Challenge Grant programs. The train-the-trainer model was used most effectively for reaching large groups of participants. The Alliance+ program was a group of school systems in three states which was able to reach 6,500 teachers. Another program selected two master teachers from each school that would then become trainers for the rest of the school.

Staff development for technology areas is a relatively new arena for staff development and districts are still working their way through developing effective technology plans that address this need. Bradshaw (2002b) reviewed the technology plans for 27 schools districts and discussed the positive and negative aspects of the various plans. The results showed that 20 of the districts had plans which included follow-up strategies, although only six of those districts explained their specific plans. One of the areas with the greatest variation in the plans was that of evaluation. There were some observed inconsistencies between the training aspects of the plans and the evaluation that would follow those plans.

Staff Development Evaluation

Much the same as teaching and learning, a concept may be presented in a staff development session; however, without follow-up and evaluation the outcomes of the activity will never be known. Professional development should be seen as a process and not a single event (Guskey, 1995). As the push for increased information on student learning has evolved, so has the need for more thorough information regarding the outcome of staff development activities. A component of the Goals 2000 (US DOE, n.d.) plan created under the Bush administration in the 1980s was professional development for teachers. One of the ten key

principles of high-quality professional development was evaluating the impact on teacher effectiveness and linking that information to student achievement. The information should then be used to guide further staff development activities (US DOE, n.d.). Currently the emphasis for staff development evaluation is coming from the National Staff Development Council's standards, NCLB Title II, and many state accountability plans (Killion, 2003).

Feedback gained from evaluations of professional development activities is critical for assessing the success of the training and the needs for follow-up support (Bradshaw, 2000b; Mizell, 2003). In the past, the evaluations of activities were based on the perceptions or the contentment of the participants. The focus was on the actual training including the length of the activity and the temperature of the room, rather than on the knowledge and skills that the participants were able to later put into practice. The focus has now shifted toward measuring the skills gained and the eventual link to student achievement. In the study by Bradshaw, the results indicated that only 12 of the 27 districts in the study had included activities for measuring the change in the teachers' knowledge and skills following the training. However, 20 of those districts had plans for follow-up support meaning that a number of those districts would not have accurate information to use in the planning of future staff development activities (Bradshaw, 2002b). Guskey (1995) stated that often new practices would be abandoned by teachers if they did not receive feedback on their activities. The successful practices are those in which there is positive reinforcement of the activity and therefore the individuals continue to practice that activity. There must be a balance between time to acquire new skills and feedback given regarding the teachers' successful use of the new knowledge and skills.

In order for evaluation to have a purpose, the goals for the program must first be set so that the evaluation has a framework for creation. A master plan for the program can help

determine those goals (Champion, 2002). From the goals, the benchmarks or standards can be created that will determine the success of the program (Killion, 2003). Once the benchmarks have been set, a method for gathering data that best assesses those skills must be determined (Champion, 2002; Guskey, 2003). Multiple measures may be needed to gather a complete and thorough evaluation of the program (Guskey, 2003; Heller, Daehler, & Shinohara, 2003). In some cases a sample may be a more manageable evaluation group rather than trying to assess the goals based on all the participants in the staff development (Champion, 2002).

Killion (2003) described an eight step process for evaluating staff development activities that she categorized in three phases. The first phase is the planning phase which includes determining if the program is ready to be evaluated, creating the evaluation questions, and constructing the method for conducting the evaluation. The second phase included the data collection, analysis, and interpretation. The final phase, the reporting phase, which included communication of the results and reflection on the evaluation process. The process of effective evaluation of staff development activities should be a structured process so that the results gained from the evaluation can be used to improve on the activities in the future.

Although teachers are often apprehensive about being evaluated following a staff development activity, frequently the most useful information is gained from measuring the skills gained by the teachers. Champion (2003) made a number of suggestions for successful evaluation of staff development activities. Participants should be told at the beginning of the activity that they will have their skills measured following the activity. In addition, evaluations should not make the personal performance of individuals known, should assess only the skills that are most critical, should assess the use of higher level skills, and should use the information

from the evaluation immediately to improve the program. These ideas can help better determine the actual learning of the participants.

Kellaher and Maher (2003) studied a mentor program in Prince George's County Schools that used an evaluation component to determine whether the program goals were being met by the mentors. The goals of the program were to reduce the attrition rate of new teachers, reduce the number of provisionally certified teachers, and to increase the academic achievement of students. Mentors were required to record information in journals that were reviewed monthly, report the time spent working with mentees, conduct action research projects within their schools and create an annual portfolio. By using the information gained from the data collection, the evaluation process was expanded and the district began working with the University of Maryland to review their data. The evaluation process led to further improvements in the mentor program which continued to change as the data was reviewed. While in another study, Heller, et al. (2003) found that, by creating an evaluation framework for reviewing a staff development for new science curriculum, evidence could easily be gathered for determining the program's success. The study found changes in both teacher and student behaviors in relation to the study.

Summary

Many different components are part of the complex evaluation of a technology-based staff development program using data in the current accountability system. The literature reported that evaluation has become a key component to determining the usefulness of a staff development activity. That evaluation should be substantive and should focus on the skills and learning attained by the participants. In addition, under the current accountability system, the understanding of student achievement data is vital and tools that can make that task more manageable should be utilized whenever possible. Because the federal and state structures for

school accountability can only be changed through legislative action, the systems are unlikely to disappear in the near future.

Given those facts, information regarding teacher understanding of the use of the Student Assessment Reporting Tool (StART) based on the train-the-trainer model of training that was used is vitally important. Based on the data gained from this evaluation the success of the current model can be assessed as well as determining the future needs to better support the teachers in their use of data to make instructional decisions for their students.

CHAPTER 3

METHODOLOGY

This chapter describes the research procedures of this study. The following sections are included: population and sample, treatment, instrumentation, data collection procedures and statistical analysis.

Research Design

The design of the study was an action research one-shot case study design (Campbell & Stanley, 1963) of the teacher level access pilot program. The sample for the study includes teachers, administrators, and other school personnel from the pilot schools. The participants were randomly chosen from the schools.

Null Hypotheses

The following null hypotheses were tested in this study:

Ho1: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on participants' level of education (Bachelors, Masters, Specialist or Beyond).

Ho1a: Subtest scores on accessing StART

Ho1b: Subtest scores on interpreting reports

Ho1c: Subtest scores on instruction

Ho1d: Overall scores

Ho2: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on participants' position (administrator, teacher, etc.) within the school.

Ho2a: Subtest scores on accessing StART

Ho2b: Subtest scores on interpreting reports

Ho2c: Subtest scores on instruction

Ho2d: Overall scores

Ho3: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on participants' years of experience in education.

Ho3a: Subtest scores on accessing StART

Ho3b: Subtest scores on interpreting reports

Ho3c: Subtest scores on instruction

Ho3d: Overall scores

Ho4: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on participants' school level (elementary, middle or high school).

Ho4a: Subtest scores on accessing StART

Ho4b: Subtest scores on interpreting reports

Ho4c: Subtest scores on instruction

Ho4d: Overall scores

Ho5: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on the Title I status of the participants' school.

Ho5a: Subtest scores on accessing StART

Ho5b: Subtest scores on interpreting reports

Ho5c: Subtest scores on instruction

Ho5d: Overall scores

Ho6: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on whether the leadership team in the school is new or established.

Ho6a: Subtest scores on accessing StART

Ho6b: Subtest scores on interpreting reports

Ho6c: Subtest scores on instruction

Ho6d: Overall scores

Population and Sample

The population for this study consisted of certificated personnel in the 18 pilot elementary, middle, and high schools. The schools were given priority as pilot schools were Title schools, new schools and schools with new principals. The principals were given the option of becoming a pilot school and 7 of the original list of 25 principals declined the offer. Attendance at the initial training session at each school varied based on the school level (elementary, middle, or high school). The elementary schools had all administrators (average of three people), the local school technology coordinator and a representative teacher from grades one through five. The middle schools had just the administrators (average of six people) and the local school technology coordinator. The high schools had the administrators (average of eight people), the department heads (average of eight people), and the local school technology coordinator. Each school had a total staff ranging from 80 to 200, based on the size of the school. Twenty participants were chosen at random from each of the schools including five non-direct instruction staff members, building level administrators, media specialists, local school technology coordinators, and school counselors. The decision was made to randomly select a representative sample of participants due to the time constraints of the school personnel as well as the variation in total school size ranging from 80 to 200. The total representative sample was 360 participants.

Variables

The independent variables were the years of experience in education, position, education level, previous data analysis experience, Title I status for the school, school level, and the tenure of the leadership team at the school for the participants.

The dependent variables were the demonstrated knowledge and skills when accessing the StART program as determined by the overall score and three subtest scores on the StART Pilot School Survey.

Data Collection Procedures

Data were collected from a random sample of participants in the pilot school program through the 2003-2004 school year, using the StART Pilot School Survey. The instrument was administered to the randomly chosen group at each pilot school by the researcher. Each participant was given a maximum of 30 minutes to complete the evaluation. The researcher carefully observed completion of the instrument in order ensure independence of answers by the participants. All instrument responses were scored by the researcher using the rubric reviewed by the peer review committee.

Instrument

The StART Pilot School Survey was a criterion-referenced instrument developed by the researcher with the work of an expert peer review team. Prior to the creation of the StART Pilot School Survey, all information related to the StART tool was based on perception surveys completed at the pilot schools the previous year. The team working on the creation and training for StART decided that application based data was required to fully evaluate the knowledge and skills of the users of StART. The initial draft of the evaluation instrument included the demographic questions and five questions that required the respondent to use StART to find the

answers to the questions. The committee agreed that the instrument did not contain enough items to evaluate the knowledge of participants in a valid or reliable manner. The Gwinnett County Public Schools research review committee, including the Director of Research and Accountability, reviewed and modified the instrument before approval. The modifications included the addition of items with screenshots from StART to determine the knowledge of participants related to the data without the need to access the actual tool. In addition, questions were added that required the respondent to connect assessment results to instructional decisions. The Executive Director for Student Accountability, Assessment, and Advisement also participated in the creation of the StART Pilot School Survey, ensuring that the items used in the instrument would provide valid and reliable information.

The StART Pilot School Survey (See Appendix A) consists of five demographic questions followed by items related to accessing the StART program, interpreting reports and drawing conclusions from data presented. Performance scores on the instrument were determined using a rubric created with the assistance of the committee in which each answer on the instrument was rated on a scale of one to four, with the exception of the demographic questions. All answers were specific to the data from the schools in which the participants were employed.

Statistical Analysis

Since pre-post measures were not available for the participants, a one-way analysis of variance (ANOVA) was used to compare participant results based on the role of the participant in the school, the education level of the participant, the school level and the years of experience in education of the participant. The participant performance based on the Title I status of the school, as well as the tenure of leadership team was analyzed using a t-test for all analyses.

The level of statistical significance for all statistical analyses was .05. The level of significance is the probability of making a Type I error when the null hypothesis is rejected. There may be an incorrect decision made 5 % of the time when a true hypothesis is rejected.

CHAPTER 4

ANALYSIS OF DATA

This chapter reports the findings which emerged from the data analyzed in this study. The chapter is organized into as follows: (a) Sampling and data collection, (b) results of null hypotheses, and (c) summary.

Sample and Data Collection

At the beginning of this study, 360 building level personnel were identified to participate in the administration of the StART Pilot School Survey. There were 20 individuals identified at each of the 18 participating schools. Participants were either teachers, administrators, or other certified personnel. Of the selected participants, 175 completed the instrument. The decrease in the total sample can be attributed to the following factors: scheduling conflicts prevented personnel from attending during the site visit by the researcher, some school personnel had left the assigned school during the school year and were no longer employed at that facility, and several schools did not complete the training of teachers and chose not to participate.

Data collection took place during the Spring of 2004. The researcher visited each school participating in the evaluation instrument over a two month period. All instruments were completed during a 30 minute session with the researcher in a computer lab. All participants worked independently as instructed to by the researcher.

Descriptive and Inferential Results

The findings in the section were the results of statistical analysis of the independent variables of level of education, role in the school, years of experience, school level, Title I school status, and leadership tenure. The dependent variables included three sub scores in addition to

the Total Score achieved by each participant. For each independent variable, descriptive statistics are provided followed by the statistical analysis, post hoc analysis and discussion for those scores that are statistically significant. Table 1 provides an overview of statistical results obtained to assist the reader with the following results.

For level of education, role in the school, years of experience and school level an analysis of variance was completed because there were three choices, while a t-test was used to analyze the results for Title I status and leadership tenure because there were only two choices.

Ho1: There is no statistically significant difference in the participants' mean scores on the StART Pilot School Survey, based on participants' level of education (Bachelors, Masters, Specialist or Beyond).

Ho1a: Subtest scores on accessing StART: Table 2 indicates the N, range, mean, and standard deviation for each of the levels of education. The mean for those participants with a specialist degree or higher was larger than that of those with lesser degrees. However, the mean was still below 50% correct. Table 3 shows the analysis of variance for the subtest score for accessing StART. As Table 3 indicates there was a statistically significant difference ($F= 14.190$, $p \leq .000$) in the scores based on education level. Therefore the null hypothesis was rejected. Table 4 displays the post hoc analysis. The scores indicate that those with a specialist or higher education degree were more successful in accessing StART than those with a masters or bachelors degree.

The difference in the scores may be related to the fact that those with higher degrees typically have more experience and therefore are more knowledgeable in using tools within the school system. In addition, those with a masters' or specialist degree have often received more training related to using data in upper level graduate courses.

Table 1

Analysis Summary

	Education Level		Role in School		Years of Experience		School Level		Title I Status		Leadership Tenure	
	F	sig	F	sig	F	sig	F	sig	t	sig	t	sig
Accessing StART Score	14.190	.000	25.375	.000	2.301	.103	15.042	.000	.062	.806	3.222	.076
Interpreting Reports Score	1.578	.209	3.468	.033	.874	.419	.497	.609	.922	.346	1.153	.286
Instruction Score	1.794	.169	4.479	.013	.467	.627	.448	.640	2.752	.110	.774	.381
Total Score	8.173	.000	15.438	.000	.257	.774	1.993	.139	.013	.910	.088	.768

Table 2

Descriptive Statistics for Accessing StART Score and Education Level

Access Score

Education Level	N	Range	Mean	Std. Deviation
Bachelors Degree	57	0 - 8	1.88	2.605
Masters Degree	83	0 -16	3.57	4.675
Specialist Degree or Higher	35	0 -16	6.63	4.845
Total	175	0 -16	3.63	4.460

Table 3

Analysis of Variance for Accessing StART Score and Education Level

Access Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	490.160	2	245.080	14.190	.000
Within Groups	2970.697	172	17.271		
Total	3460.857	174			

Table 4

Scheffe Post Hoc for Accessing StART Score and Education Level

Dependent Variable: Access Score

Scheffe

(I) Education Level	(J) Education Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Bachelors Degree	Masters Degree	-1.689	.715	.064	-3.45	.08
	Specialist Degree or Higher	-4.751*	.892	.000	-6.96	-2.55
Masters Degree	Bachelors Degree	1.689	.715	.064	-.08	3.45
	Specialist Degree or Higher	-3.062*	.838	.002	-5.13	-.99
Specialist Degree or Higher	Bachelors Degree	4.751*	.892	.000	2.55	6.96
	Masters Degree	3.062*	.838	.002	.99	5.13

*. The mean difference is significant at the .05 level.

Ho1b: Subtest scores on Interpreting Reports: Table 5 indicates the N, range, mean, and standard deviation for each of the levels of education. The means were similar for all three groups, although the standard deviation was slightly higher for the participants with a masters' degree. Table 41 (Appendix D) shows the analysis of variance for the subtest score for Interpreting Reports. There was no statistically significant difference (F=1.578) in these scores. Therefore the null hypothesis was accepted.

Ho1c: Subtest scores on Instruction: Table 6 indicates the N, range, mean, and standard deviation for each of the levels of education. The scores indicate a better understanding of connecting Instruction with the assessment data for those with higher degrees. However, Table 42 (Appendix D) shows the analysis of variance for the subtest score for Instruction. There was no statistically significant difference (F=1.794) in these scores. Therefore the null hypothesis was accepted.

Ho1d: Total Score: Table 7 indicates the N, range, mean, and standard deviation for each of the levels of education. When the aggregated scores were tabulated, the difference between those with the highest level of education and the two lower levels was greater. Table 8 shows the analysis of variance for the Total Score. As Table 8 indicates there was a statistically significant difference ($F= 8.173$, $p \leq .000$) in the scores based on education level. Table 9 displays the post hoc analysis. Therefore the null hypothesis was rejected. The scores indicate that those with specialist degrees or higher were more successful in accessing StART.

Similar to the Accessing StART Score the difference in the scores may be due to those with higher degrees typically having more experience in using tools within the school system. Although, the overall hypothesis was sustained because the rule of having two or more subscores with significant differences was not met.

Table 5

Descriptive Statistics for Interpreting Reports Score and Education Level

Report Score

Education Level	N	Range	Mean	Std. Deviation
Bachelors Degree	57	0 - 16	11.60	3.256
Masters Degree	83	0 - 18	11.39	4.161
Specialist Degree or Higher	35	0 - 16	12.71	3.536
Total	175	0 - 18	11.72	3.778

Table 6

Descriptive Statistics for Instruction Score and Education Level

Instruction Score

Education Level	N	Range	Mean	Std. Deviation
Bachelors Degree	57	0 - 16	5.75	3.646
Masters Degree	83	0 - 16	6.94	3.905
Specialist Degree or Higher	35	0 - 16	7.09	4.883
Total	175	0 - 16	6.58	4.057

Table 7

Descriptive Statistics for Total Score and Education Level

Total Score

Education Level	N	Range	Mean	Std. Deviation
Bachelors Degree	57	0 - 32	19.05	6.160
Masters Degree	83	4 - 47	21.76	9.024
Specialist Degree or Higher	35	8 - 46	26.31	9.749
Total	175	0 - 47	21.79	8.702

Table 8

Analysis of Variance for Total Score and Education Level

Total Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1143.611	2	571.806	8.173	.000
Within Groups	12033.566	172	69.963		
Total	13177.177	174			

Table 9

Scheffe Post Hoc for Total Score and Education Level

Dependent Variable: Total Score

Scheffe

(I) Education Level	(J) Education Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Bachelors Degree	Masters Degree	-2.706	1.439	.174	-6.26	.85
	Specialist Degree or Higher	-7.262*	1.796	.000	-11.70	-2.83
Masters Degree	Bachelors Degree	2.706	1.439	.174	-.85	6.26
	Specialist Degree or Higher	-4.555*	1.686	.028	-8.72	-.39
Specialist Degree or Higher	Bachelors Degree	7.262*	1.796	.000	2.83	11.70
	Masters Degree	4.555*	1.686	.028	.39	8.72

*. The mean difference is significant at the .05 level.

Ho2: There is no statistically significant difference in the participants' mean scores on the StART Pilot School Survey, based on participants' position (administrator, teacher, etc.) within the school.

Ho2a: Subtest scores on accessing StART: Table 10 indicates the N, range, mean, and standard deviation for each of the school roles. Those who were administrators had more than double the score of any other group of participants for accessing StART. Table 11 shows the analysis of variance for the subtest score for accessing StART. As Table 11 indicates there was a statistically significant difference ($F= 25.375$, $p \leq .000$) in the scores based on school role. Therefore the null hypothesis was rejected. Table 12 displays the post hoc analysis. The scores indicate that those who have the role as administrator in the school are more successful in accessing the StART program than those in other roles.

The difference in the scores may be due to the fact that the administrators were all trained by the researcher on how to access StART. All other school personnel were given the training in the redelivery model. Also, administrators had access to optional additional training sessions offered centrally. In some instances, other personnel were never given the training that the administrators received.

Ho2b: Subtest scores on Interpreting Reports: Table 13 indicates the N, range, mean, and standard deviation for each of the school roles. Those who were administrators were also more proficient at Interpreting Reports based on the scores. Table 14 shows the analysis of variance for the subtest score for Interpreting Reports. As Table 14 indicates there was a statistically significant difference ($F= 3.468$, $p \leq .033$) in the scores based on school role. Therefore the null hypothesis was rejected. Table 15 displays the post hoc analysis. The scores indicate that those

Table 10

Descriptive Statistics for Accessing StART Score and School Role

Access Score

School Role	N	Range	Mean	Std. Deviation
Teacher	139	0 - 16	2.80	4.000
Administrator	17	4 - 16	10.00	4.359
LSTC, TST or other	19	0 - 8	4.00	2.981
Total	175	0 - 16	3.63	4.460

Table 11

Analysis of Variance for Accessing StART Score and School Role

Access Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	788.497	2	394.249	25.375	.000
Within Groups	2672.360	172	15.537		
Total	3460.857	174			

Table 12

Scheffe Post Hoc for Accessing StART Score and School Role

Dependent Variable: Access Score

Scheffe

(I) School Role	(J) School Role	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-7.201*	1.013	.000	-9.70	-4.70
	LSTC, TST or other	-1.201	.964	.462	-3.58	1.18
Administrator	Teacher	7.201*	1.013	.000	4.70	9.70
	LSTC, TST or other	6.000*	1.316	.000	2.75	9.25
LSTC, TST or other	Teacher	1.201	.964	.462	-1.18	3.58
	Administrator	-6.000*	1.316	.000	-9.25	-2.75

*. The mean difference is significant at the .05 level.

Table 13

Descriptive Statistics for Interpreting Reports Score and School Role

Report Score

School Role	N	Range	Mean	Std. Deviation
Teacher	139	0 - 18	11.82	3.586
Administrator	17	6 - 16	13.00	2.550
LSTC, TST or other	19	0 - 14	9.84	5.315
Total	175	0 - 18	11.72	3.778

Table 14

Analysis of Variance for Interpreting Reports Score and School Role

Report Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	96.250	2	48.125	3.468	.033
Within Groups	2387.030	172	13.878		
Total	2483.280	174			

Table 15

Scheffe Post Hoc for Interpreting Reports Score and School Role

Dependent Variable: Report Score

Scheffe

(I) School Role	(J) School Role	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-1.180	.957	.469	-3.54	1.18
	LSTC, TST or other	1.978	.911	.098	-.27	4.23
Administrator	Teacher	1.180	.957	.469	-1.18	3.54
	LSTC, TST or other	3.158*	1.244	.042	.09	6.23
LSTC, TST or other	Teacher	-1.978	.911	.098	-4.23	.27
	Administrator	-3.158*	1.244	.042	-6.23	-.09

*. The mean difference is significant at the .05 level.

who have the role as administrator in the school are more skilled with Interpreting Reports than those in other roles.

The difference in the scores may be due to the fact that the administrators were all trained by the researcher on how to interpret StART reports. Also, administrators have had access to the StART tool for a longer period of time. In addition, administrators often have significantly more experience in reviewing assessment data than others within the school. Their prior knowledge with assessments may have impacted their performance.

Ho2c: Subtest scores on Instruction: Table 16 indicates the N, range, mean, and standard deviation for each of the school roles. Again, the administrators had a higher mean score followed by those who were identified as teachers. Table 17 shows the analysis of variance for the subtest score for Instruction. As Table 17 indicates there was a statistically significant difference ($F= 4.479$, $p \leq .013$) in the scores based on school role. Therefore the null hypothesis was rejected. Table 18 contains the post hoc analysis. The scores indicate that those who have the role as administrator in the school are more skilled with connecting assessment report data to instructional decisions.

The difference in the scores may be due to the fact that part of the job responsibility of an administrator is to work on improving student achievement through full school data analysis. A major component of that responsibility entails finding connections between student achievement data and instruction. Teachers are just beginning to have experience with StART and this may have had an impact on their performance. In addition, the technology personnel (TST and LSTC) have the responsibility with assisting with accessing the program, but may not have had any training in data interpretation.

Table 16

Descriptive Statistics for Instruction Score and School Role

Instruction Score

School Role	N	Range	Mean	Std. Deviation
Teacher	139	0 - 16	6.35	3.776
Administrator	17	4 - 16	9.29	4.524
LSTC, TST or other	19	0 - 16	5.89	4.875
Total	175	0 - 16	6.58	4.057

Table 17

Analysis of Variance for Instruction Score and School Role

Instruction Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	141.805	2	70.903	4.479	.013
Within Groups	2722.743	172	15.830		
Total	2864.549	174			

Table 18

Scheffe Post Hoc for Instruction Score and School Role

Dependent Variable: Instruction Score

Scheffe

(I) School Role	(J) School Role	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-2.949*	1.022	.017	-5.47	-.42
	LSTC, TST or other	.451	.973	.898	-1.95	2.85
Administrator	Teacher	2.949*	1.022	.017	.42	5.47
	LSTC, TST or other	3.399*	1.328	.040	.12	6.68
LSTC, TST or other	Teacher	-.451	.973	.898	-2.85	1.95
	Administrator	-3.399*	1.328	.040	-6.68	-.12

*. The mean difference is significant at the .05 level.

Ho2d: Total Score: Table 19 indicates the N, range, mean, and standard deviation for each of the school roles. The aggregated effect of the Total Score gave the administrators a higher mean for this measure also. Table 20 shows the analysis of variance for the Total Score. As Table 20 indicates there was a statistically significant difference ($F= 15.438$, $p \leq .000$) in the scores based on education level. Therefore the null hypothesis was rejected. Table 21 contains the post hoc analysis. The scores indicate that those who have the role as administrator in the school were more successful in total on the evaluation instrument.

The difference in the scores may be due to the fact that administrators were more successful on in each individual subtest and the cumulative effect impacted the administrators' Total Score. Also, administrators had a better overall understanding of the tool. Therefore, the overall hypothesis that there is no statistically significant difference in the participants' mean score on the StART Pilot School Survey was rejected.

Ho3: There is no statistically significant difference in the participants' mean scores on the StART Pilot School Survey, based on participants' years of experience in education.

Ho3a: Subtest scores on Accessing StART: Table 22 indicates the N, range, mean, and standard deviation for each of the years of experience. There is little variation among the means and standard deviations for the three groups. Table 43 (Appendix D) shows the analysis of variance for the subtest score for accessing StART related to years of experience. There was no statistically significant difference ($F=2.301$, $p \leq .103$) in these scores. Therefore the null hypothesis was accepted.

Ho3b: Subtest scores on Interpreting Reports: Table 23 indicates the N, range, mean, and standard deviation for each of the years of experience. The largest difference in scores for this measure is the range difference of a starting score of 6 for those with less experience and starting

Table 19

Descriptive Statistics for Total Score and School Role

Total Score

School Role	N	Range	Mean	Std. Deviation
Teacher	139	0 - 47	20.81	7.783
Administrator	17	20 - 46	32.06	9.250
LSTC, TST or other	19	4 - 32	19.74	8.956
Total	175	0 - 47	21.79	8.702

Table 20

Analysis of Variance for Total Score and School Role

Total Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2005.415	2	1002.708	15.438	.000
Within Groups	11171.762	172	64.952		
Total	13177.177	174			

Table 21

Scheffe Post Hoc for Total Score and School Role

Dependent Variable: Total Score

Scheffe

(I) School Role	(J) School Role	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-11.246*	2.071	.000	-16.36	-6.13
	LSTC, TST or other	1.076	1.971	.862	-3.79	5.94
Administrator	Teacher	11.246*	2.071	.000	6.13	16.36
	LSTC, TST or other	12.322*	2.691	.000	5.68	18.97
LSTC, TST or other	Teacher	-1.076	1.971	.862	-5.94	3.79
	Administrator	-12.322*	2.691	.000	-18.97	-5.68

*. The mean difference is significant at the .05 level.

Table 22

Descriptive Statistics for Accessing StART Score and Years of Experience

Access Score

Years of Experience	N	Range	Mean	Std. Deviation
0-5	48	0 - 16	2.58	4.083
6-16	71	0 - 16	3.69	4.364
17 or more	56	0 -16	4.45	4.775
Total	175	0 - 16	3.63	4.460

Table 23

Descriptive Statistics for Interpreting Reports Score and Years of Experience

Report Score

Years of Experience	N	Range	Mean	Std. Deviation
0-5	48	6 - 18	12.21	2.760
6-16	71	0 - 18	11.77	3.642
17 or more	56	0 - 16	11.23	4.616
Total	175	0 - 18	11.72	3.778

scores of zero for the other groups. Table 44 (Appendix D) shows the analysis of variance for the subtest score for Interpreting Reports related to years of experience. There was no statistically significant difference ($F=.874$, $p\leq.419$) in these scores. Therefore the null hypothesis was accepted.

Ho3c: Subtest scores on Instruction: Table 24 indicates the N, range, mean, and standard deviation for each of the years of experience. The scores are almost identical for all three groups for this measure. Table 45 (Appendix D) shows the analysis of variance for the subtest score for Instruction related to years of experience. There was no statistically significant difference ($F=.467$, $p\leq.627$) in these scores. Therefore the null hypothesis was accepted.

Ho3d: Total Score: Table 25 indicates the N, range, mean, and standard deviation for each of the years of experience. Once more, the scores are very similar among the groups based on years of experience Table 46 (Appendix D) shows the analysis of variance for the subtest score for Total Score related to years of experience. There was no statistically significant difference ($F=.257$, $p\leq.774$) in these scores. Therefore the null hypothesis was accepted.

The overall hypothesis that there is no statistically significant difference in the participants' mean score based on years of experience is sustained. The results did not meet the rule of two or more subscores that were statistically significant.

Ho4: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on participants' school level (elementary, middle or high school).

Table 24
Descriptive Statistics for Instruction Score and Years of Experience

Instruction Score

Years of Experience	N	Range	Mean	Std. Deviation
0-5	48	0 - 16	6.46	3.307
6-16	71	0 - 16	6.93	4.054
17 or more	56	0 - 16	6.25	4.641
Total	175	0 - 16	6.58	4.057

Table 25
Descriptive Statistics for Total Score and Years of Experience

Total Score

Years of Experience	N	Range	Mean	Std. Deviation
0-5	48	11 - 46	21.08	7.222
6-16	71	4 - 47	22.25	8.535
17 or more	56	0 - 44	21.80	10.079
Total	175	0 - 47	21.79	8.702

Ho4a: Subtest scores on Accessing StART: Table 26 indicates the N, range, mean, and standard deviation for each school level. The mean score for high school was higher than those in the other school levels. Table 27 shows the analysis of variance for the subtest score for accessing StART. As Table 27 indicates there was a statistically significant difference ($F= 15.042$, $p \leq .000$) in the scores based on school level. Therefore the null hypothesis was rejected. Table 28 displays the post hoc analysis. The scores indicate that those who are in high schools were more successful in accessing the StART program than those in other levels.

The difference in the scores may be due to the fact that the school structure for high schools allowed for more assistance in the training and use of StART. Department chairpersons were used as an additional resource for teachers at the pilot schools that did not exist at the elementary and middle schools. In addition, all high school leaders indicated that their teachers were trained when that was not the case at all the elementary and middle schools.

Ho4b: Subtest scores on Interpreting Reports: Table 29 indicates the N, range, mean, and standard deviation for each school level. The major difference is the higher standard deviation at the high school level. Table 47 (Appendix D) shows the analysis of variance for the subtest score for Interpreting Reports related to school level. There was no statistically significant difference ($F=.497$, $p\leq.609$) in these scores. Therefore the null hypothesis was accepted.

Ho4c: Subtest scores on Instruction: Table 30 indicates the N, range, mean, and standard deviation for each school level. The elementary level mean is slightly higher than the other two levels. However, Table 48 (Appendix D) shows the analysis of variance for the subtest score for Instruction related school level. There was no statistically significant difference ($F=.448$, $p\leq.640$) in these scores. Therefore the null hypothesis was accepted.

Table 26

Descriptive Statistics for Accessing StART Score and School Level

Access Score

School Level	N	Range	Mean	Std. Deviation
Elementary School	89	0 - 16	2.18	3.588
Middle School	14	0 - 12	2.29	3.407
High School	72	0 - 16	5.68	4.838
Total	175	0 - 16	3.63	4.460

Table 27

Analysis of Variance for Accessing StART Score and School Level

Access Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	515.224	2	257.612	15.042	.000
Within Groups	2945.634	172	17.126		
Total	3460.857	174			

Table 28

Scheffe Post Hoc for Accessing StART Score and School Level

Dependent Variable: Access Score

Scheffe

(I) School Level	(J) School Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.106	1.190	.996	-3.04	2.83
	High School	-3.501*	.656	.000	-5.12	-1.88
Middle School	Elementary School	.106	1.190	.996	-2.83	3.04
	High School	-3.395*	1.209	.021	-6.38	-.41
High School	Elementary School	3.501*	.656	.000	1.88	5.12
	Middle School	3.395*	1.209	.021	.41	6.38

*. The mean difference is significant at the .05 level.

Table 29
Descriptive Statistics for Interpreting Reports Score and School Level

Report Score

School Level	N	Range	Mean	Std. Deviation
Elementary School	89	2 - 16	12.00	2.896
Middle School	14	0 - 16	11.50	3.818
High School	72	0 - 18	11.42	4.660
Total	175	0 - 18	11.72	3.778

Table 30
Descriptive Statistics for Instruction Score and School Level

Instruction Score

School Level	N	Range	Mean	Std. Deviation
Elementary School	89	0 - 16	6.85	3.541
Middle School	14	0 - 16	6.00	4.076
High School	72	0 - 16	6.36	4.643
Total	175	0 - 16	6.58	4.057

Ho4d: Total Score: Table 31 indicates the N, range, mean, and standard deviation for each school level. The range, mean and standard deviation had slight variations among the school levels. Table 49 (Appendix D) shows the analysis of variance for the subtest score for Total Score related to school level. There was no statistically significant difference ($F=1.993$, $p\leq.139$) in these scores. Therefore the null hypothesis was accepted.

The overall hypothesis that there is no statistically significant difference in the participants' mean score based on school level is sustained. The results did not meet the rule of two or more subscores that were statistically significant.

Ho5: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on the Title I status of the participants' school.

Ho5a: Subtest scores on Accessing StART: Table 32 indicates the N, range, mean, and standard deviation for Title I status. There were few differences between Title I and Non-Title I schools for accessing StART. Table 50 (Appendix D) shows the t-test for the subtest score for accessing StART related to Title I status. There was no statistically significant difference ($t=.062$, $p\leq.806$) in these scores. Therefore the null hypothesis was accepted.

Ho5b: Subtest scores on Interpreting Reports: Table 33 indicates the N, range, mean, and standard deviation for Title I status. The means and standard deviations were also similar for interpreting the reports. Table 51 (Appendix D) shows the t-test for the subtest score for Interpreting Reports related to Title I status. There was no statistically significant difference ($t=.992$, $p\leq.346$) in these scores. Therefore the null hypothesis was accepted.

Ho5c: Subtest scores on Instruction: Table 34 indicates the N, range, mean, and standard deviation for Title I status. Those participants in Title I schools had slightly higher scores for connecting data to Instruction. Table 52 (Appendix D) shows the t-test for the subtest score for

Table 31
Descriptive Statistics for Total Score and School Level

Total Score

School Level	N	Range	Mean	Std. Deviation
Elementary School	89	6 - 46	20.92	6.490
Middle School	14	4 - 44	19.57	8.706
High School	72	0 - 47	23.29	10.738
Total	175	0 - 47	21.79	8.702

Table 32
Descriptive Statistics for Accessing StART Score and Title I Status

Access Score

Title I Status	N	Range	Mean	Std. Deviation
Non-Title	154	0 - 16	3.66	4.547
Title	21	0 - 12	3.43	3.854
Total	175	0 - 16	3.63	4.460

Table 33

Descriptive Statistics for Interpreting Reports Score and Title I Status

Report Score

Title I Status	N	Range	Mean	Std. Deviation
Non-Title	154	0 - 18	11.82	3.796
Title	21	2 - 16	11.00	3.647
Total	175	0 - 18	11.72	3.778

Table 34

Descriptive Statistics for Instruction Score and Title I Status

Instruction Score

Title I Status	N	Range	Mean	Std. Deviation
Non-Title	154	0 - 16	6.39	4.014
Title	21	0 - 16	8.00	4.195
Total	175	0 - 16	6.58	4.057

Instruction related to Title I status. There was no statistically significant difference (t=2.752, p≤.110) in these scores. Therefore the null hypothesis was accepted.

Ho5d: Total Score: Table 35 indicates the N, range, mean, and standard deviation for Title I status. Although the means were similar, the range of scores for participants in Title I schools was smaller than that of participants in Non-Title I schools. Table 53 (Appendix D) shows the t-test for the subtest score for Total Score related to Title I status. There was no statistically significant difference (t=.013, p≤.910) in these scores. Therefore the null hypothesis was accepted.

The overall hypothesis that there is no statistically significant difference in the participants' mean score based on Title I status is sustained. The results did not meet the rule of two or more subscores that were statistically significant.

Ho6: There is no statistically significant difference in the participants' mean scores on the StART evaluation, based on whether the leadership team in the school is new or established.

Ho6a: Subtest scores on accessing StART: Table 36 indicates the N, range, mean, and standard deviation for Leadership Tenure. There was little difference in accessing StART when scores were compared between participants in schools with new or established leadership. Table 54 (Appendix D) shows the t-test for the subtest score for accessing StART related to Leadership Tenure. There was no statistically significant difference (t=3.222, p≤.076) in these scores. Therefore the null hypothesis was accepted.

Ho6b: Subtest scores on Interpreting Reports: Table 37 indicates the N, range, mean, and standard deviation for Leadership Tenure. The Interpreting Reports Scores were also quite similar. Table 54 (Appendix D) shows the t-test for the subtest score for Interpreting Reports related to Leadership Tenure.

Table 35

Descriptive Statistics for Total Score and Title I Status

Total Score

Title I Status	N	Range	Mean	Std. Deviation
Non-Title	154	0 - 47	21.77	8.955
Title	21	12 - 42	21.95	6.734
Total	175	0 - 47	21.79	8.702

Table 36

Descriptive Statistics for Accessing StART Score and Leadership Tenure

Access Score

Leadership Tenure	N	Range	Mean	Std. Deviation
New	121	0 - 16	3.21	4.239
Established	54	0 - 16	4.57	4.828
Total	175	0 - 16	3.63	4.460

Table 37

Descriptive Statistics for Interpreting Reports Score and Leadership Tenure

Report Score

Leadership Tenure	N	Range	Mean	Std. Deviation
New	121	0 - 18	11.95	3.346
Established	54	0 - 18	11.20	4.594
Total	175	0 - 18	11.72	3.778

There was no statistically significant difference (t=1.153., p≤.286) in these scores. Therefore the null hypothesis was accepted.

Ho6c: Subtest scores on Instruction: Table 38 indicates the N, range, mean, and standard deviation for Leadership Tenure. There was little differentiation between the scores for connecting data to instructional changes. Table 56 (Appendix D) shows the t-test for the subtest score for Instruction related to Leadership Tenure. There was no statistically significant difference (t=.774, p≤.381) in these scores. Therefore the null hypothesis was accepted.

Ho6d: Total Score: Table 39 indicates the N, range, mean, and standard deviation for Leadership Tenure. The most noticeable disparity was in the larger standard deviation of the established leadership. Table 57 (Appendix D) shows the t-test for the subtest score for Total Score related to Leadership Tenure. There was no statistically significant difference (t=.088, p≤.768) in these scores. Therefore the null hypothesis was accepted.

Summary of Data

A summary of the areas of significance are noted in Table 50. Statistically significant differences occurred for all scores related to the role in the school. Those who were administrators performed significantly better than those in other positions within the school. In addition, those who were in high schools and had a higher education level had scores that were statistically significantly higher for accessing StART.

Table 38
Descriptive Statistics for Instruction Score and Leadership Tenure

	Instruction Score			
Leadership Tenure	N	Range	Mean	Std. Deviation
New	121	0 - 16	6.78	3.798
Established	54	0 - 16	6.15	4.595
Total	175	0 - 16	6.58	4.057

Table 39
Descriptive Statistics for Total Score and Leadership Tenure

	Total Score			
Leadership Tenure	N	Range	Mean	Std. Deviation
New	121	4 - 47	21.93	7.842
Established	54	0 - 46	21.46	10.452
Total	175	0 - 47	21.79	8.702

Table 40

Areas of Significance

	Education Level	Role in School	Years of Experience	School Level	Title I Status	Leadership Tenure
Accessing StART Score	**	**	NS	**	NS	NS
Interpreting Reports Score	NS	*	NS	NS	NS	NS
Instruction Score	NS	*	NS	NS	NS	NS
Total Score	**	**	NS	NS	NS	NS

* Statistically Significant p <.05

** Statistically Significant p<.01

CHAPTER 5

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Chapter 5 contains the summary of the study. Conclusions are given and followed by recommendations for further study.

Summary

The purpose of this study was to determine the effectiveness of StART training for building level personnel. The participants were all faculty members at the 18 elementary, middle, and high schools pilot schools for teacher level access to StART. Success of the training was measured using the StART Pilot School Survey which had scores for accessing StART, report interpretation, instructional focus and an overall score.

Six variables were identified that might have an impact on the success of personnel when using the tool. Education level, role in school, years of experience, school level, Title I status of the school, and leadership tenure were all compared to determine areas of significance related to performance.

Education level of the participants had a statistically significant impact on both the Accessing StART Score and the Total Score. Those individuals with more advanced education levels were more successful than those with on a Bachelors degree. This result was expected because those with higher degrees often have had more coursework related to using and understanding assessment data than those with a Bachelors' degree.

Significant differences were found in all four scores related to role in the school. Those who were identified as school administrators had significantly higher scores than all others. This group was in almost all instances the group of individuals that were trained by the researcher.

Therefore, they did not receive the redelivery of training. Unlike the study by Poplin (2003), this study found in large part that the train the trainer model for technology staff development was unsuccessful. Although this model was efficient in training a large number of participants in the Poplin study, the same result was not found here.

The StART training was not just technology training, but included training related to student assessment and data interpretation skills. In addition, many school administrators admitted that the teachers in their schools were never trained on how to access StART and were just given copies of reports to use. Therefore, teachers were able to understand the information provided in the reports, but were unable to access the reports themselves. Local school administrators were also present in numerous centrally held meetings in which the expectation of using this tool was given by central office personnel. Finally, administrators have had access to the StART program prior to the pilot phase and have had more experience in using this tool.

School level only impacted the Accessing StART Score and the individuals who were in high schools were more able to use the tool. The department chairs were used to train the teachers in the high schools and this additional resource was not available at the other school levels. In the research by Cooley (2001), the results also indicated that using teachers to train other teachers was successful with the implementation of a new technology tool. In most of the other schools in the study the leadership teams did not train a group of teacher leaders in the school. Therefore, many of the teachers in the elementary and middle schools were not trained at all on how to access the tool. Therefore, the scores for many individuals were near zero for accessing StART.

Conclusions

The StART training for administrators that was done all by the researcher was successful based on the results of the surveys. In addition, regardless of the training provided to individuals in the schools, the participants were able to gain information from the sample reports provided. Unlike the study by Sharkey and Murnane (2003), this study found that participants had more knowledge about the meaning of the data although they were less likely to be able to access the data using StART themselves. Also, most participants were able to provide some information regarding instructional changes that could improve student achievement. Additionally, there was no statistically significant difference found among participants based on the Title I status of the school or the Leadership Tenure within the school.

As Kellaher and Maher (2003) state the purpose of evaluation should be program improvement. Consequently, changes need to be made to the train the trainer model of implementation used in the pilot schools. Although training was provided on how to use StART, the results indicate that basic data and assessment training should be given prior to any StART training. Therefore a reworking of the training materials will take place prior to the next phase of StART teacher level access implementation. The PowerPoint presentation will be redone to include specific school examples for each school that is trained. Also, slides will be provided that explain how assessments are organized prior to explaining how StART can be accessed and used. A document explaining exactly which test information is available for students at each grade level will be created to assist teachers in understanding the information that is included in the StART reports. More scaffolding will be provided to assist users of StART in the understanding of information provided within the tool related to assessments and data. Also,

centrally we may consider providing some data analysis training sessions that would allow administrators to bring that training back to their local schools.

The number of participants in the survey was lower than the anticipated number due to several factors including busy school schedules, conflicts with other activities, and reluctance of leadership personnel to participate. The StART program was in a pilot phase during the research period and a need existed to encourage a positive view of StART. Subsequently when someone was not willing to participate in the survey, the issue was not pressed. Furthermore, surveys could only be completed while the researcher was at the school site to ensure that the participants completed the survey under similar circumstances. Moving forward, schools will be required to set follow-up survey sessions at the time of initial training to increase the participation.

Another major issue was the lack of teacher training. The administrators within the surveyed schools were trained, but a large percentage of the teachers were not trained. Zimmerman and May (2003) found that principal support of technology training was vital for teachers to incorporate the training into their practice. This study found that the lack of administrative support caused teachers to either not be trained. In the future, schools will be given additional training tools, such as a video of the researcher conducting a training session to use in training. A contact person, an administrator within the school, will distribute information and materials regarding StART must be designated in each school. This person will be the contact for all data information for the school and will therefore be in position to determine how StART fits within the overall school data plan. Also, each school will need to create a training plan for their faculty. This plan will include the schedule for training, materials that will be used and a list of those individuals who will be conducting the training.

Recommendations

Recommendations for further research are as follows:

1. Significant differences were found between the school levels in this study and further investigation should be done to determine if that pattern would again occur with enhancements to the training program.
2. Additional research should be conducted to determine if there is a statistical difference in use and effectiveness of training between schools that made AYP and those schools that did not make AYP.
3. Research should be conducted to determine the impact of the information gained from use of StART on the content of local school plans for improvement and individual teacher goals.
4. If a connection is found between the use of StART and goals set, research should be conducted on the impact on student achievement that takes place due to the use of assessment data within StART.

Recommendations for the StART tool:

1. Make the website interface more user-friendly including the replacement of hyperlinks with buttons, replacement of report descriptions with those used in the data guide, and an addition of pictures to indicate which historical reports are for schools and which are for teachers.
2. The designers should create a new folder structure on the main menu that clearly indicates which reports are created from the most recent extraction of student schedules.

3. Schedules should be extracted on a schedule that is logical at each school level and those dates will not be all identical for the 2004-2005 school year.
4. Focus groups should be created that are able to meet regularly to provide feedback that can be used to further enhance StART.
5. Additional tools that go beyond the capabilities of StART should be explored for those people identified as power data users to further increase the usage of data within the schools.
6. An analysis of report usage should be done to determine if certain reports are not useful to schools and other reports should be created to take the place of those reports that are eliminated.

Recommendations for StART training:

1. Schools will be required to designate a StART coordinator that will be responsible for setting a training schedule, informing the staff of updates to the program and scheduling follow-up survey times.
2. New principals will be given the expectation that StART should be used by their leadership team and their staff.
3. Additional centrally located training sessions will be offered for administrators and local school technology coordinators.

REFERENCES

A Plus Education Reform Act (2000). Georgia Stat. §§ LC 27 1066-EC.

Ausbrooks, C. (2000). *Technology and the changing nature of school administration.* (Report No. EA030845). East Lansing, MI: National Center for Research on Teacher Learning (ERIC Document Reproduction Service No. ED449557)

Blum, R. (2000). Standards-based reform: Can it make a difference for students? *Peabody Journal of Education, 74(4),* 90-113.

Bradshaw, L. K. (2002a). Maximizing technology's role in school reform. *The Delta Kappa Gamma Bulletin, 68(3),* 5-9.

Bradshaw, L. K. (2002b). Technology for teaching and learning: Strategies for staff development and follow-up support. *Journal of Technology and Teacher Education, 10(1),* 131-150.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston, MA: Houghton Mifflin Company.

Champion, R. (2002). Map out evaluation goals. *The Journal of the National Staff Development Council, 23(4),* 78-79.

Champion, R. (2003). The real measure of a professional development program's effectiveness lies in what participants learn. *The Journal of the National Staff Development Council, 24(1),* 75-76.

Cooley, V. E. (2001). Implementing technology using the teachers as trainers staff development model. *Journal of Technology and Teacher Education, 9(2),* 269-284.

Cooper, C., & Cromey, A. (2000). Teachers and students as action researchers: Using data daily. *NCREL's Learning Point, 2(2),* 8-11.

Davenport, P., & Anderson, G. (2002) *Closing the achievement gap: No excuses*. Houston, TX: American Productivity and Quality Center.

Dawson, C., & Rakes, G. C. (2003). The influence of principal's technology training on the integration of technology into schools. *Journal of Research on Technology in Education, 36(1),* 29-49.

Decker, G. (2003). Using data to drive student achievement in the classroom and on high-stakes tests. *T.H.E. Journal, 30(6),* 44-45.

Doyle, D. P. (2003) Data-drive decision-making. *T.H.E. Journal, 30(10),* 19-21.

Education Commission of the States. (2002). *Data-driven decisionmaking: No Child Left Behind issue brief* (GP-02-10). Denver, CO: Author.

Educational Research Service. (2001). *Using assessment data to improve instruction.* Arlington, VA: Author.

The Elementary and Secondary School Act of 1965, Pub, L. No. 89-10, (1965).

Elmore, R. F., & Fuhrman, S. H. (2001). Holding schools accountable: Is it working? *Phi Delta Kappan, 83(1),* 67-72.

Finn, C. E. (2002). Making school reform work. *Public Interest, 148,* 85-95.

Fisher, D. (2001). Trust the process: Increasing student achievement via professional development and process accountability. *NAASP Bulletin, 85,* 67-71.

Geer, C. (2002). Technology training for school administrators: A real world approach. *Tech Trends, 46(6),* 56-59.

Georgia Department of Education (2003). *Georgia Consolidated State Application Accountability Workbook.* Accepted by USED May 19, 2003.

Goertz, M. E., & Duffy, M. C. (2001). *Assessments and accountability across the 50 states* (CPRE Policy Briefs No. 33). Consortium for Policy Research in Education.

Goertz, M. E., & Duffy, M. C. (2003). Mapping the landscapes of high-stakes testing and accountability programs. *Theory into Practice, 42(1),* 4-11.

Guskey, T. R. (1995). *Results- oriented professional development: In search of an optimal mix of effective practices.* University of Kentucky: North Central Regional Educational Laboratory.

Guskey, T. (2003). Scooping up meaningful evidence. *The Journal of the National Staff Development Council, 24(4),* 27-30.

Hardy, L. (2003, July). Information, please: Managing all the data required under NCLB is a task made for technology. *Technology Focus.* Retrieved January 19, 2004, from http://www.asbj.com/2003/07/0703/technologyfocus.html.

Heller, J. I., Daehler, K. R., & Shinohara, M. (2003). Connecting all the pieces. *The Journal of the National Staff Development Council, 24(4),* 36-41.

Holland, P. E. (2001). Professional development in technology: Catalyst for school reform. *Journal of Technology and Teacher Education, 9(2),* 245-267.

Jenson, J., Lewis, B., & Smith, R. (2002). No one way: Working models for teachers' professional development. *Journal of Technology and Teacher Education, 10(4),* 481-496.

Johnson, J. H. (1997). *Data-driven school improvement* (Report No. EDO-EA-97-1). Eugene, OR:ERIC Clearinghouse on Educational Management. (ERIC Document Reproduction Service No. EA027548)

Kalay, P., & Chen, D. (2002). Integrating a decision support system into a school: The effects on student functioning. *Journal of Research on Technology in Education, 34(4),* 435-452.

Kellaher, A., & Maher, J. (2003). My mentor, myself. *The Journal of the National Staff Development Council, 24(4),* 31-35.

Killion, J. (2003). Solid footwork makes evaluation of staff development programs a song. *The Journal of the National Staff Development Council, 24(4),* 15-21.

Kinder, A. (2000). D^3 M: Helping schools distill data. *NCREL's Learning Point, 2(2),* 4-7.

Levine, E. (2002, November). Building a data warehouse. *Technology Focus.* Retrieved January 19, 2004, from http://www.asbj.com/2002/11/1102/technologyfocus2.html.

Levine, E. (2003, September). The data trek: New tools make sifting through mounds of information easier. *American School Board Journal: September 2003 Special Report.* Retrieved January 19, 2004 from http://www.asbj.com/specialreports/0903SpecialReports/S4.html

Liddle, K. (2000). Data-driven success: How one elementary school mined assessment data to improve instruction. *American School Board Journal: Electronic School, 187 (3),* 30-32.

Linn, R. L., Baker, E.L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the no child left behind act of 2001. *Educational Researcher, 31(6),* 3-16.

Lohr, N. J. (2003). Implementing the complexities of NCLB. *T.H.E. Journal, 30(10),* supp. 25-28.

Mertler, C.A. (2002). *Using standardized test data to guide instruction and intervention* (Report No. EDO-TM-02-07). Washington, DC: ERIC Clearinghouse on Assessment and Evaluation. (ERIC Document Reproduction Service No. ED470589)

Mintrop, H., & MacLellan A. M. (2002). School improvement plans in elementary and middle schools on probation. *The Elementary School Journal, 102(4),* 275-300.

Mizell, H. (2003). Facilitator:10 refreshments:8 evaluation:0. *The Journal of the National Staff Development Council, 24(4),* 10-13.

National Staff Development Council. *NSDC standards for staff development.* Retrieved January 3, 2004, from http://www.nsdc.org/standards/index.cfm

No Child Left Behind Act of 2001, Pub, L. No. 107-110, 115 Stat 1425 (2002).

Pearson, P. D., Vyas, S., Sensale, L. M., & Kim, Y. (2001). Making out way through the assessment and accountability maze: Where do we go now? *The Clearing House, 74,* 175-182.

Poplin, C. J. (2003). Models of professional development. *T.H.E. Journal, 30(11),* 38-40.

Sharkey, N. S., & Murnane, R. J. (2003). Learning from student assessment results. *Educational Leadership, 61(3),* 77-81.

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

Telem, M. (1999). A case study of the impact of school administration computerization on the department head's role. *Journal of Research on Computing in Education, 31(4),* 385-401.

United States Department of Education (n.d.) *Goals 2000, Building bridges: The mission & principles of professional development.* Retrieved January 3, 2004 from http://www.ed.gov/G2K/bridge.html.

Vitali, G. J. (1994). *Factors influencing teachers' practices in an assessment driven reform.* (Report No. SP035399). East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED373053)

Wade, H. H. (2001). *Data inquiry and analysis for educational reform.* (Report No. EDO-EA-01-10). Eugene, OR: ERIC Clearinghouse on Educational Management. (ERIC Document Reproduction Service No. ED461911)

Whittaker, A., & Young, V. M. (2002). Tensions in assessment design: Professional development under high-stakes accountability. *Teacher Education Quarterly, 29(3),* 43-60.

Zimmerman, J. A., & May, J. J. (2003). Providing effective professional development: What's holding us back? *American Secondary Education, 31(2),* 37-48.

APPENDIX A

START PILOT SCHOOL SURVEY AND SCORING RUBRIC

Appendix A

School Name _____

2003-2004 StART Pilot School Survey

1. What is your school level?
   A. Elementary school **B.** Middle School **C.** High School

2. How many years of experience do you have (not including this year)?
   A. 0-5 **B.** 6-16 **C.** 17 or more

3. What is your position in the school?
   A. Teacher **B.** Administrator **C.** LSTC/TST or other

4. What is your education level?
   A. Bachelors degree **B.** Masters degree **C.** Specialist or higher

5. What is your previous experience (prior to this year) using assessment data to make instructional decisions?
   A. Never tried **B.** Tried once **C.** Tried more than once

6. What program do you use to access StART?
   A. Lotus Notes **B.** Internet Explorer **C.** SASI

7. What report can be used to quickly see the strand level strengths and weaknesses based on the scores for the students currently in your school?
   A. Strand Distribution Ave. **B.** Strand Ranking **C.** Strand Distribution Student

Please log-in to StART to answer the following 3 questions:

8. What are the two weakest strands for the Gateway (4, 7 or 10) for the students in your school who took the assessment for the first time in 2002-2003?
   1. _____
   2. _____

9. What were the two weakest strands for the Gateway (4, 7 or 10) for the students in your school in 2001-2002?
   1. _____

2. _____

10. What are two changes that you might make to instruction based on the information you found in StART regarding student performance?

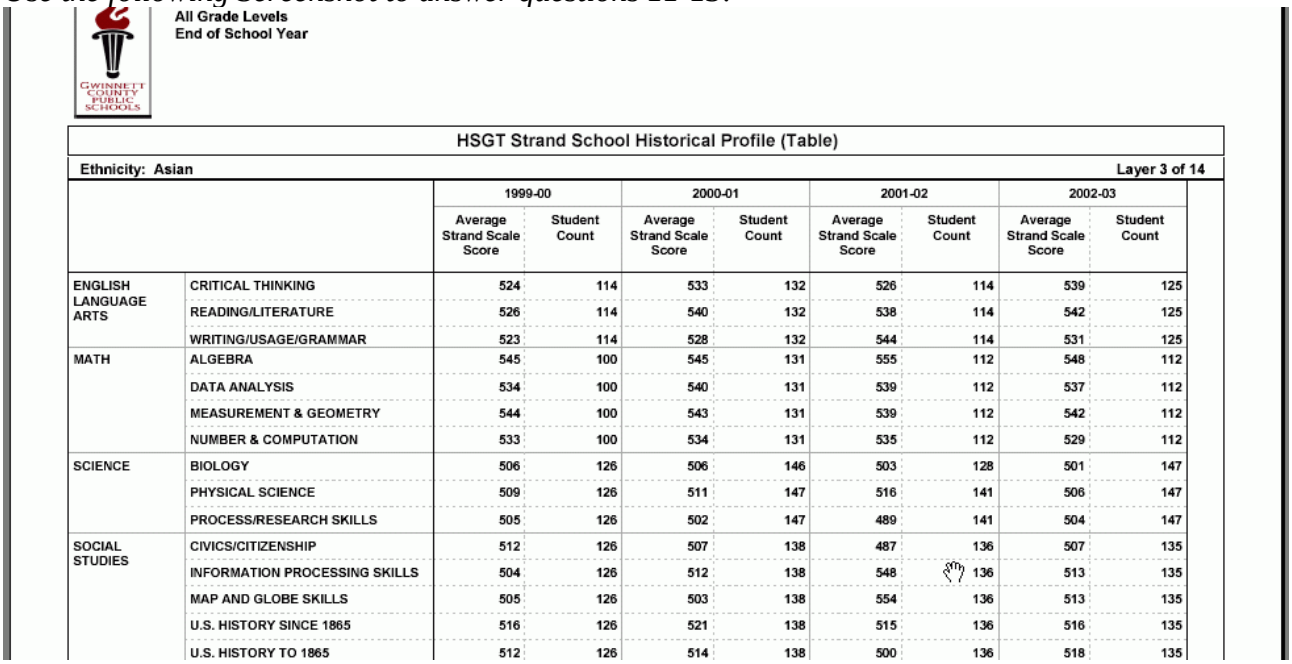
1._____

_____

2._____

_____

Use the following Screenshot to answer questions 11-13:



All Grade Levels
End of School Year

HSGT Strand School Historical Profile (Table)

Ethnicity: Asian — Layer 3 of 14

		1999-00		2000-01		2001-02		2002-03	
		Average Strand Scale Score	Student Count	Average Strand Scale Score	Student Count	Average Strand Scale Score	Student Count	Average Strand Scale Score	Student Count
ENGLISH LANGUAGE ARTS	CRITICAL THINKING	524	114	533	132	526	114	539	125
	READING/LITERATURE	526	114	540	132	538	114	542	125
	WRITING/USAGE/GRAMMAR	523	114	528	132	544	114	531	125
MATH	ALGEBRA	545	100	545	131	555	112	548	112
	DATA ANALYSIS	534	100	540	131	539	112	537	112
	MEASUREMENT & GEOMETRY	544	100	543	131	539	112	542	112
	NUMBER & COMPUTATION	533	100	534	131	535	112	529	112
SCIENCE	BIOLOGY	506	126	506	146	503	128	501	147
	PHYSICAL SCIENCE	509	126	511	147	516	141	506	147
	PROCESS/RESEARCH SKILLS	505	126	502	147	489	141	504	147
SOCIAL STUDIES	CIVICS/CITIZENSHIP	512	126	507	138	487	136	507	135
	INFORMATION PROCESSING SKILLS	504	126	512	138	548	136	513	135
	MAP AND GLOBE SKILLS	505	126	503	138	554	136	513	135
	U.S. HISTORY SINCE 1865	516	126	521	138	515	136	516	135
	U.S. HISTORY TO 1865	512	126	514	138	500	136	518	135

11. Which strand had the largest scale score increase between 2000-01 and 2001-2002?

_____

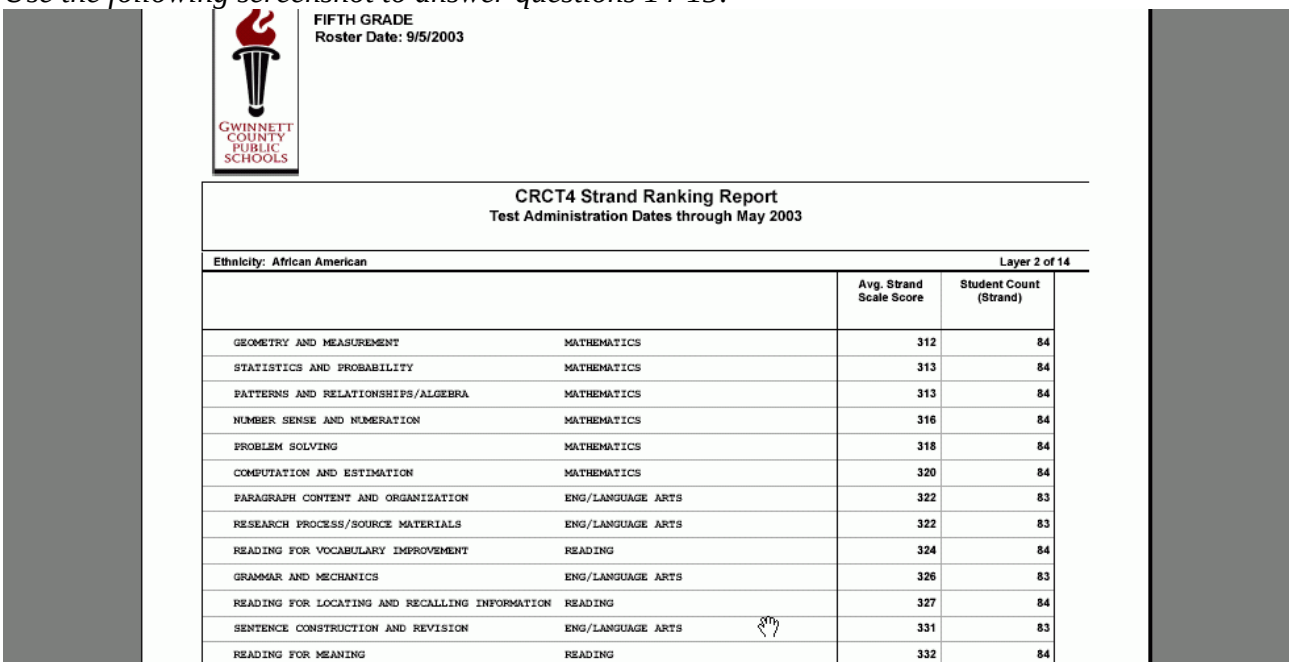
12. What was the weakest strand for all students in 2002-2003?

_____

13. What has the trend been for civics/citizenship over the four years shown?

_____

Use the following screenshot to answer questions 14-15:



FIFTH GRADE
Roster Date: 9/5/2003

GWINNETT COUNTY PUBLIC SCHOOLS

CRCT4 Strand Ranking Report
Test Administration Dates through May 2003

Ethnicity: African American — Layer 2 of 14

		Avg. Strand Scale Score	Student Count (Strand)
GEOMETRY AND MEASUREMENT	MATHEMATICS	312	84
STATISTICS AND PROBABILITY	MATHEMATICS	313	84
PATTERNS AND RELATIONSHIPS/ALGEBRA	MATHEMATICS	313	84
NUMBER SENSE AND NUMERATION	MATHEMATICS	316	84
PROBLEM SOLVING	MATHEMATICS	318	84
COMPUTATION AND ESTIMATION	MATHEMATICS	320	84
PARAGRAPH CONTENT AND ORGANIZATION	ENG/LANGUAGE ARTS	322	83
RESEARCH PROCESS/SOURCE MATERIALS	ENG/LANGUAGE ARTS	322	83
READING FOR VOCABULARY IMPROVEMENT	READING	324	84
GRAMMAR AND MECHANICS	ENG/LANGUAGE ARTS	326	83
READING FOR LOCATING AND RECALLING INFORMATION	READING	327	84
SENTENCE CONSTRUCTION AND REVISION	ENG/LANGUAGE ARTS	331	83
READING FOR MEANING	READING	332	84

14. Which was the weakest subtest for the students in this strand ranking report?

_____

15. Based on this information, which strand area could the students use the most instructional focus?

_____

16. For which group of students does this report pertain?

_____

17. What information would you like to see included in StART that is not there?

_____

_____

18. What types of additional assistance, training or support would make StART more useful for you?

_____

_____

SAMPLE SCORING RUBRIC

StART Scoring Rubric

1. School Level	**A**	**B**	**C**	
2. Years of Experience	**A**	**B**	**C**	
3. Position in School	**A**	**B**	**C**	
4. Education Level	**A**	**B**	**C**	
5. Previous Data Experience	**A**	**B**	**C**	
6. Program Access	**A**	***B***	**C**	**4 pts**
7. Strand Information	**A**	***B***	**C**	**4 pts**

8. Two Weakest Gateway Strands
   _____Use of Scientific Process_____ __/**2 pts**
   _____Knowledge & Use of Science_____ __/**2 pts**

9. 2001-2002 Weakest Gateway Strands
   _____SS: Analysis & Interpretation_____ __/**2 pts**
   _____ Use of Scientific Process/Knowledge & Use of Science __/**2 pts**

Access Total Points _____/ 16

10. Two Instruction Based Change
    1. Practice Gateway in SS, Writing in SS ___/**4pts**
    2. Practice Gateway in SC, Writing in SC ___/**4pts**

15. Instructional Focus
    ___Geometry & Measurement_____ ___/**8pts**

Instruct. Total Points _____/ 16

11. Strand with largest increase __Map & Globe Skills___ ___/**4pts**
12. Weakest Strand for 2002-03 ___Unknown (Asian Students)_____ ___/**4pts**
13. Trend for Civics __decreasing for 3 years, then increasing_____ ___/**4pts**
14. Weakest Subtest __Math_____ ___/**4pts**
16. Which subgroup ____African Americans_____ ___/**4pts**

Reports Total Points _____/ 20

17. What information would you like to see in StART?

_____No Score Given_____

18. What additional support, training or assistance would make StART more useful?

_____No Score Given_____

TOTAL SCORE _____/ 52

APPENDIX B

SCHOOL SYSTEM AND HUMAN SUBJECT APPROVALS

The University of Georgia
Office of The Vice President for Research
DHHS Assurance ID No. : FWA00003901

Institutional Review Board
Human Subjects Office
606A Graduate Studies Research Center
Athens, Georgia 30602-7411
(706) 542-6514; 542-3199
Fax No. (706) 542-5638

APPROVAL FORM

Date Proposal Received: 2003-12-17 **Project Number:** H2004-10432-0

Name	Title	Dept/Phone	Address	Email
Ms. Deborah A. Durrence	PI	Educational Leadership River's Crossing+4808	3520 Duncan Bridge Drive Buford GA 30519 (678)482-9372	debbie_durrence@gwinnett.k12.ga.us
Dr. William W. Swan	CO	Educational Leadership 313 Rivers Crossing +4808 542-4058		

Title of Study: The Effectiveness of StART Training on Building Level Personnel

45 CFR 46 Category: Administrative 1

Modifications Required for Approval and Date Completed: 2004-02-03
Application revision

Approved : 2004-02-03 Begin date : 2004-02-03 Expiration date : 2005-02-02

NOTE: Any research conducted before the approval date or after the end date collection date shown above is not covered by IRB approval, and cannot be retroactively approved.

Number Assigned by Sponsored Programs:

Funding Agency:

Form 310 Provided: No

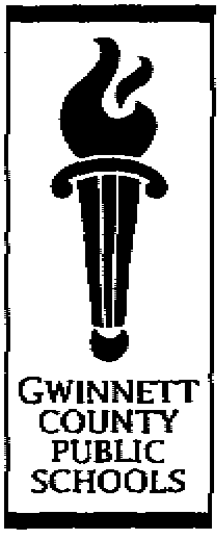
Your human subjects study has been approved as indicated under IRB action above.

Please be aware that it is your responsibility to inform the IRB . . .
. . . of any adverse events or unanticipated risks to the subjects or others within 24 to 72 hours; . .
. . . of any significant changes or additions to your study and obtain approval of them before they are put into effect; . . .
. . . that you need to extend the approval period beyond the expiration date shown above; . . .
. . . that you have completed your data collection as approved, within the approval period shown above, so that your file may be closed.

For additional information regarding your responsibilities as an investigtor refer to the IRB Guidelines.
For your convenience in obtaining approval of changes, extending the approval period, or closing your file, we are providing you with a blue Researcher Request form. Detach this blue form, complete it as appropriate, sign and date it, then return it to the IRB office. Keep this original approval form for your records.

Copy:
Dr. William G. Wraga
Dr. Sally J. Zepeda

Christina A. Joseph, Ph.D.
Chairperson, Institutional Review Board



December 3, 2003

Ms. Debbie Durrence
Asst. Principal on Assignment for Data Reporting
Lawrenceville East

Dear Ms. Durrence:

This is to advise you that your revised research proposal, "The Effectiveness of StART Training on Building Level Personnel" (our file number 2004-08) has been approved.

Please forward a copy of your results to me when they are completed. Also, would you please provide us with some feedback on the research approval process by completing the enclosed survey and returning it in the enclosed postage-paid envelope.

When contacting local schools regarding your research, please provide a copy of this approval letter to the principal. Please note that schools and teachers may elect not to participate in your research study, even though the district has granted permission.

Best wishes for a successful research project. Please call me at (770) 513-6628 if I may be of further assistance.

Sincerely,

Colin Martin

Colin Martin, Ph.D., Director
Research and Accountability

Enclosure

cc: Dr. William Swan, UGA, 350 River's Crossing, 850 College Station Rd. Athens, GA 30602
Dr. Cindy Loe, Associate Superintendent

P.O. Box 343
Lawrenceville, GA
30046-0343
770.963.8651
www.gwinnett.k12.ga.us

APPENDIX C

TRAINING PROGRAM DESCRIPTION

The training program was conducted at the school site for each pilot school involved in the study. The program was a train-the-trainer redelivery method. All schools had a leadership team trained by the primary researcher in a one and one-half hour session in a computer lab in the school. Attendance at the initial training session at each school varied based on the school level (elementary, middle, or high school). The elementary schools had all administrators (average of three people), the local School technology coordinator and a representative teacher from grades one through five. The middle schools had just the administrators (average of 6 people) and the local school technology coordinator. The high schools had the administrators (average of 8 people), the department heads (average of 8 people), and the local school technology coordinator.

The training involved review of a PowerPoint presentation, review of a data guide, and guided practice in the use of StART. User ids, passwords and computer settings at the school were verified and modified as necessary during each training session. Additional materials were provided to each school regarding using StART data for writing school goals, teacher goals, and grade level or department goals. All teachers in each school were given the same initial password and then trained on how to change the password. Each teacher was given access to information regarding their students and school aggregated information. However, teachers were not able to see class or student specific information for any other teachers. Administrators, counselors and local school technology coordinators were given access to all classes and students within their school building and aggregated school level information about all other schools in

the district. All materials used during the initial training session were given to each school on a CD-Rom so that the same presentation would then be used to train the teachers and other personnel in the school. Instructional videos that could be played on a computer were also included on the CD-Rom. Only certificated personnel are given access to StART.

Once the initial training session was completed, each leadership team trained the teachers in their school in small group sessions in the computer labs. The researcher was a resource for the leadership teams as questions arose during the training sessions, but was not present during the training of the rest of the school.

Included is a copy of the elementary school training PowerPoint and the data guide used with elementary schools. Additional PowerPoint presentations were created for the middle and high schools with examples specific to those school levels, but with the same material as the elementary school presentation.

GCPS 2003-2004 Elementary School StART Training

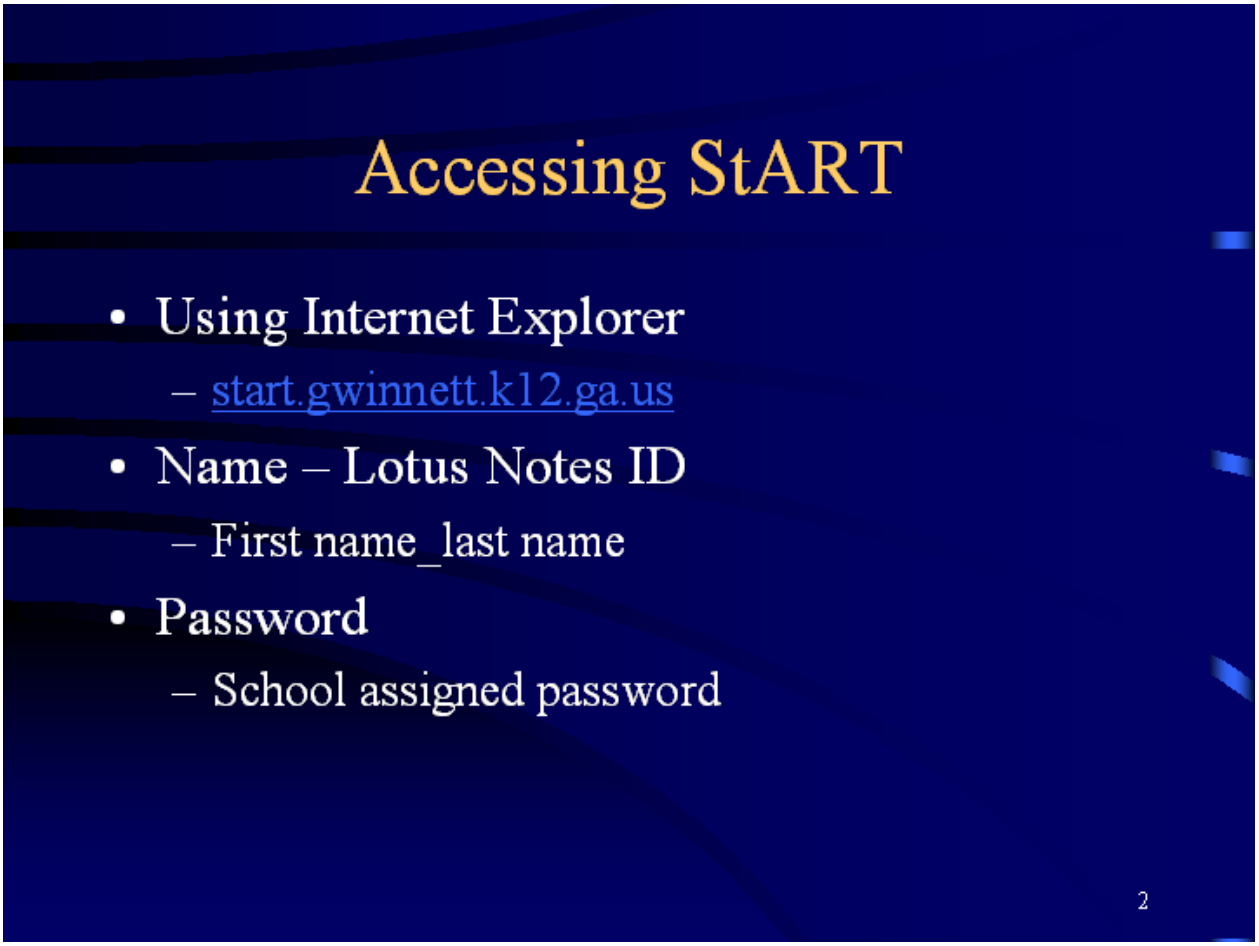
STudent Assessment Reporting Tool

The StART application has been designed to provide test data for the district cluster, school and classroom. Data can be viewed on an aggregated level, as well as by different disaggregated subgroups.

Debbie Durrence, APOSA for Data Reporting

Linda Mitchell, Executive Director, Student Assessment, Accountability and Advisement

1


Accessing StART
• Using Internet Explorer
– start.gwinnett.k12.ga.us
• Name – Lotus Notes ID
– First name_last name
• Password
– School assigned password
2

Guides for Using Data

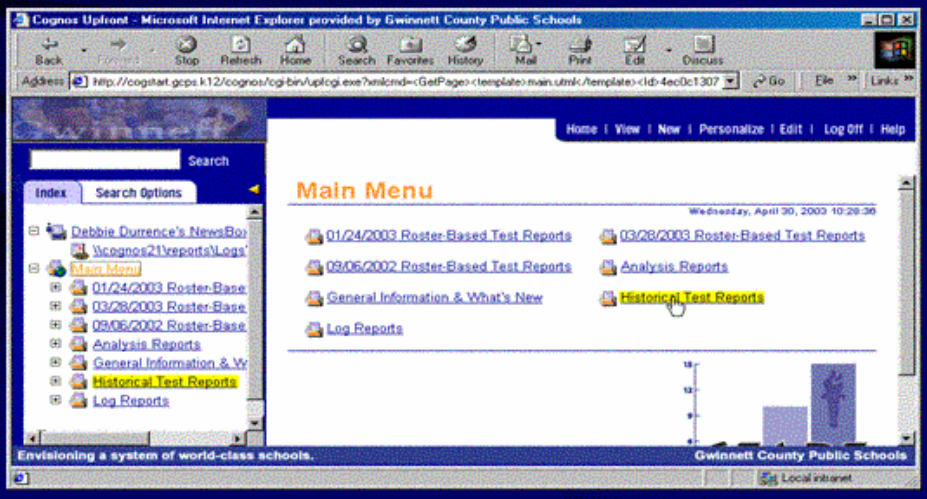
- Data Questions
  - What do you want to know?
- Test Reports Available
  - Which reports provide answers to your data questions?
- Tests Available & Grade Level Administered

Roster-based & Historical Reports

- Roster-based test reports include information for students that were on a teacher's roster on the date the roster was pulled
- Historical based reports show trend data over time using the final spring rosters for each teacher

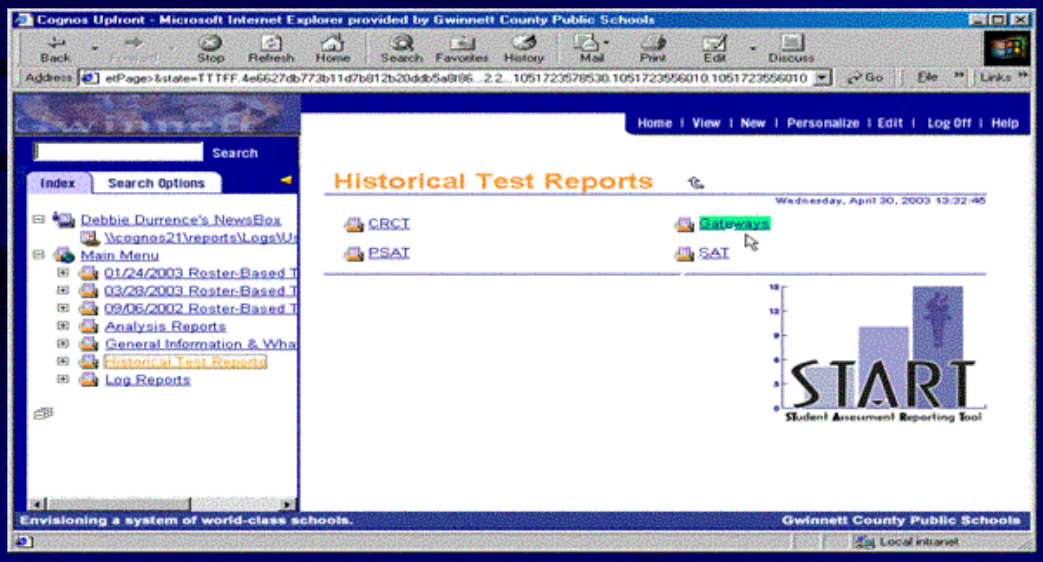
Accessing Historical Reports

- From the Main Menu Choose Historical Reports



Accessing Historical Reports

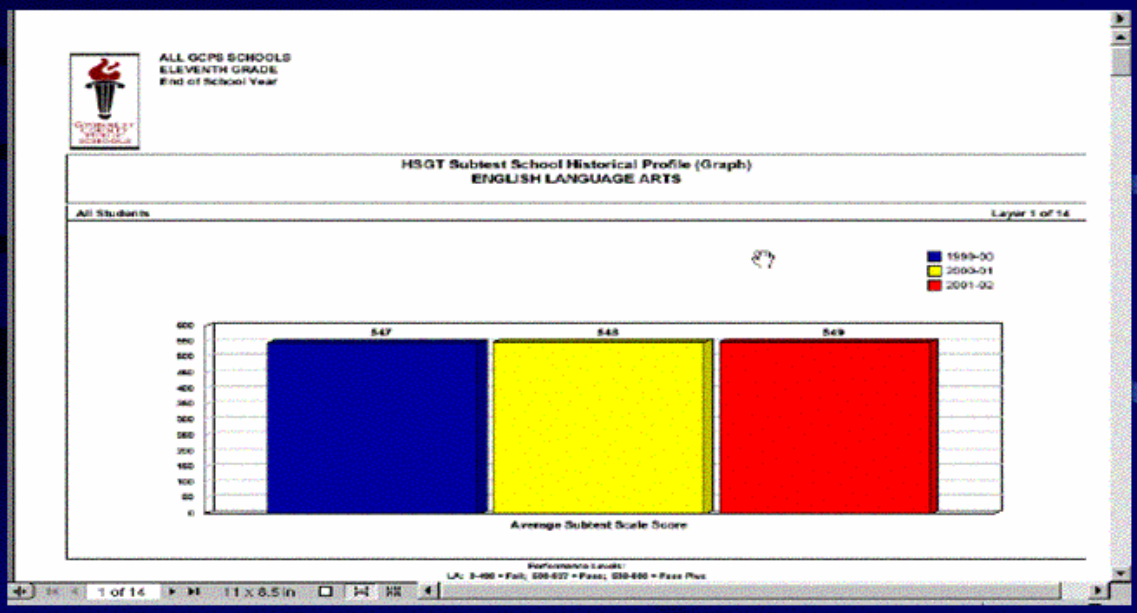
- Choose Gateway



Accessing Historical Reports

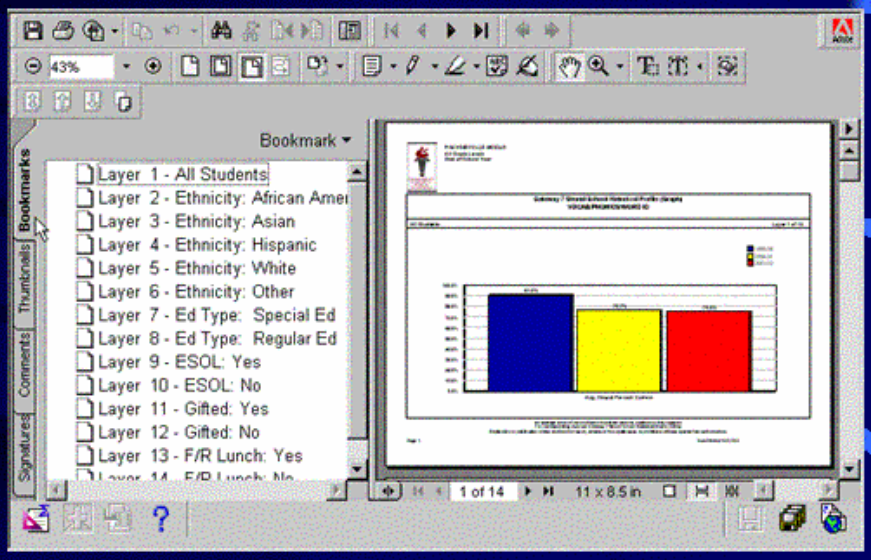
- Choose Gateway 4
- Choose report and then select your school or teacher as appropriate
- Reports Available include:
  - Teacher or School Subtest Historical Profile (Graph & Table)
  - Teacher or School Strand Historical Profile (Graph & Table)

Sample Subtest School Historical Profile Graph

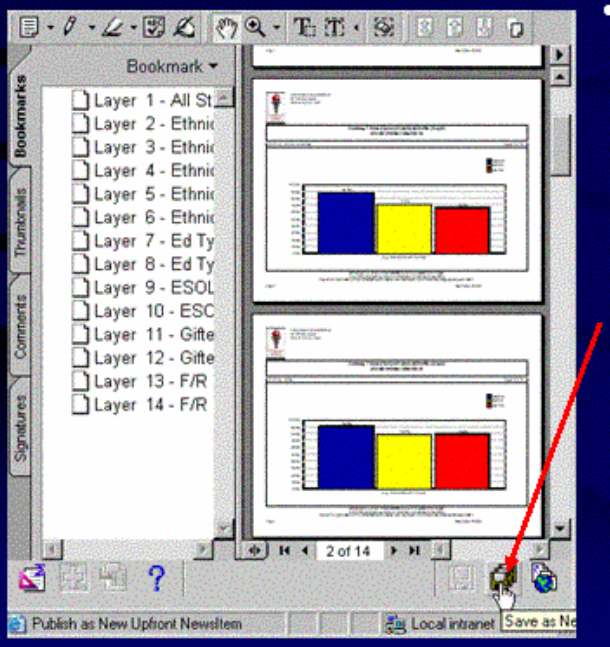


Disaggregated Views

Open the bookmarks for a disaggregated list for your report.



Saving Reports to the News-Box

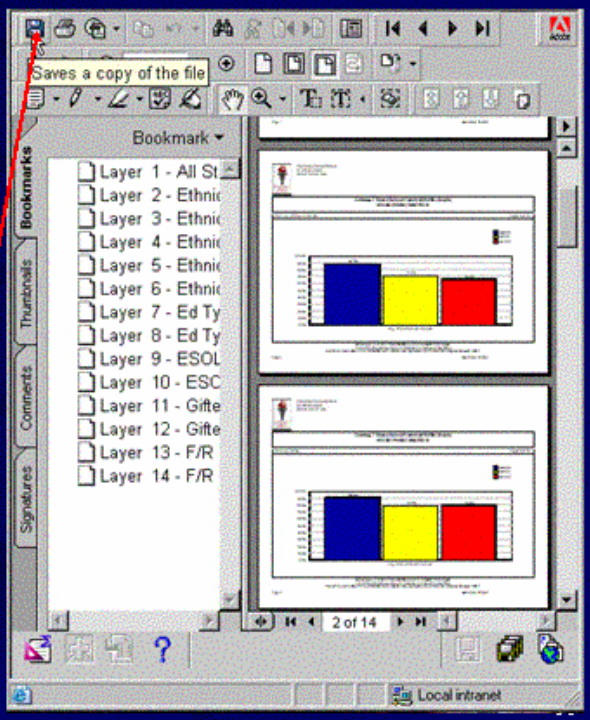


- You can save your report to you personal news-box found on the Main Menu screen by clicking on the Save as a News-Item icon located at the bottom of your screen. For complete directions on saving to your News-Box see the FAQ sheet in the General Information and What's New folder in the Main Menu.

10

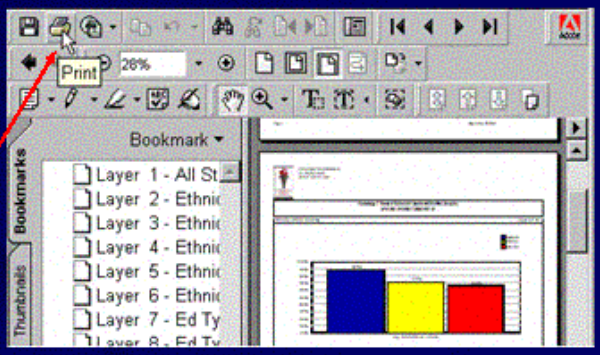
Saving Reports to a Disc or Into a File

- You can save a report to a disc or into a file on your computer by clicking on the single disc icon at the top of your report.
- Complete instructions for saving reports can be viewed from the FAQ sheet in the General Information and What's New Folder in the Main Menu.

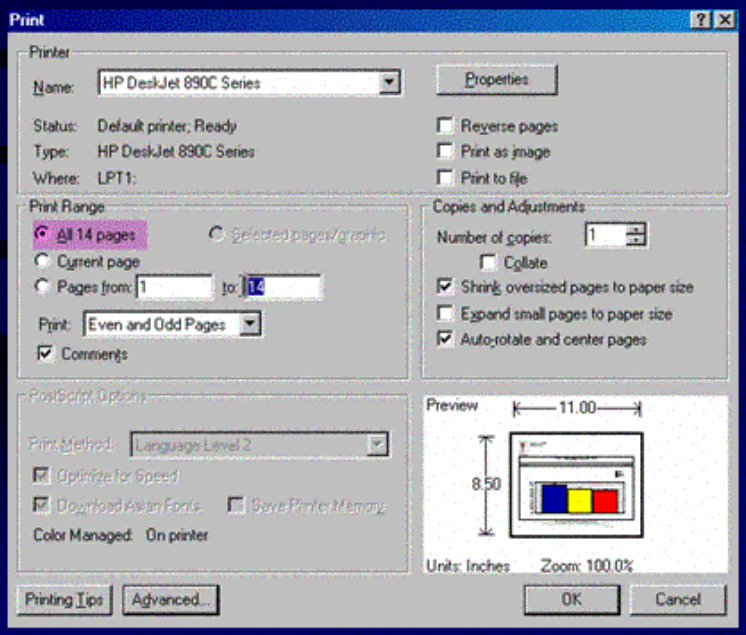


Printing Reports

- Printing Reports IS NOT RECOMMENDED! The START application is intended for on-line analysis. If you must print a report it is recommended that you select particular pages with-in the report. Reports can range from 1 page to 28 pages in length.
- To print select the
- Printer icon at top of
- report.



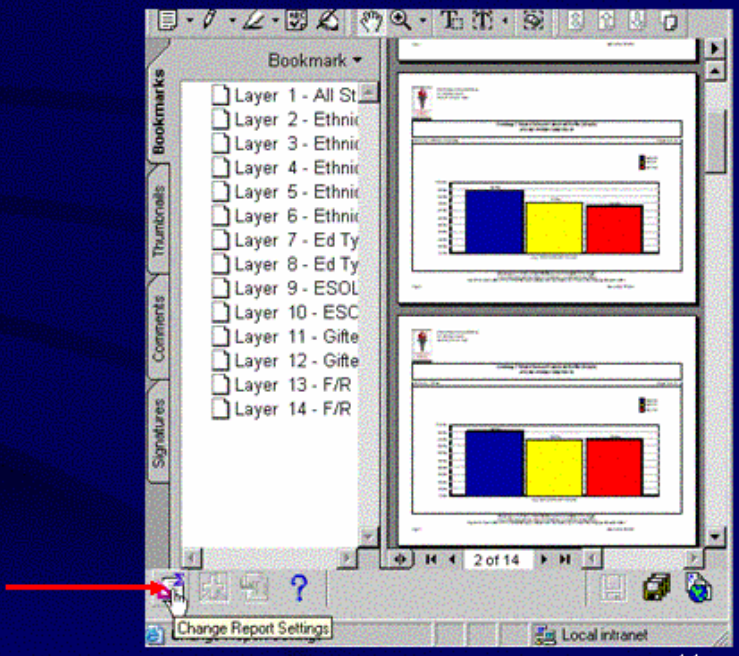
Print Dialog Box



- Before printing be sure to select the page or pages to be printed. All pages will print if no selection is made.

Changing the Report

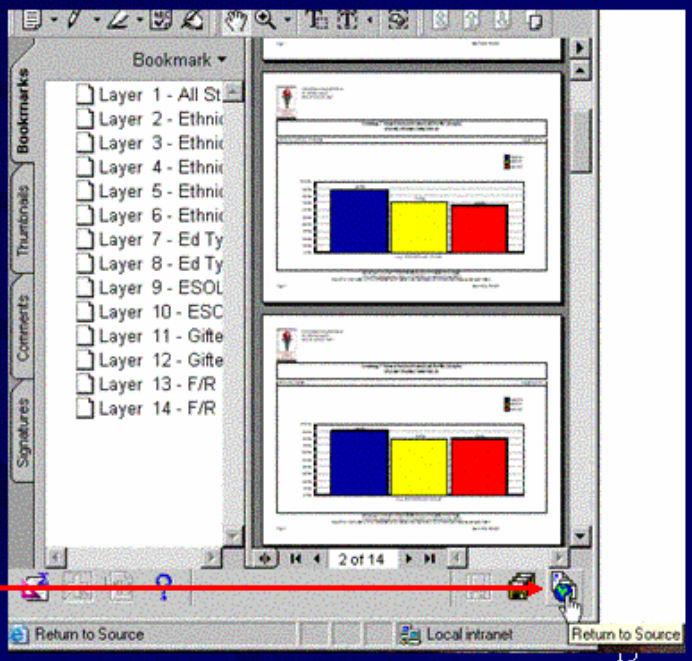
- You can return to the dimension line to make changes in your view of the report or to build alternate views of the test report selected. Click on the triangle at the bottom left of the screen to return to the dimension line.



14

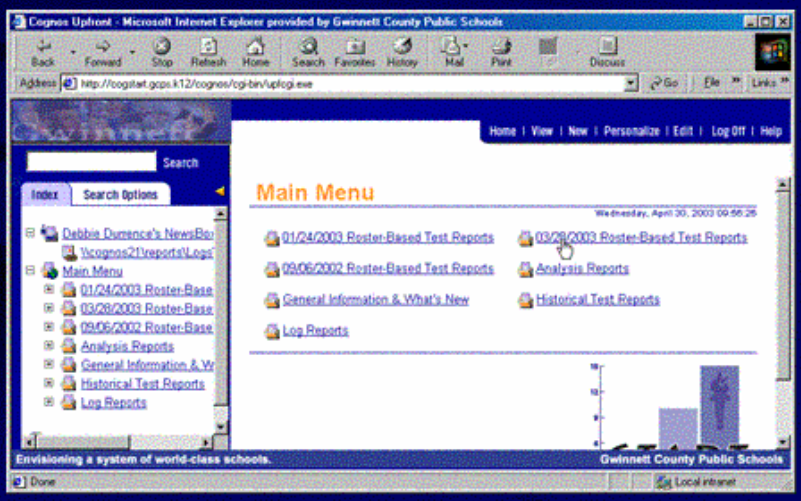
Return to the Source Menu

- To return to the test menu of the report you have built you may click on the green globe at the bottom right corner of your report. From this menu you may select alternate report types.



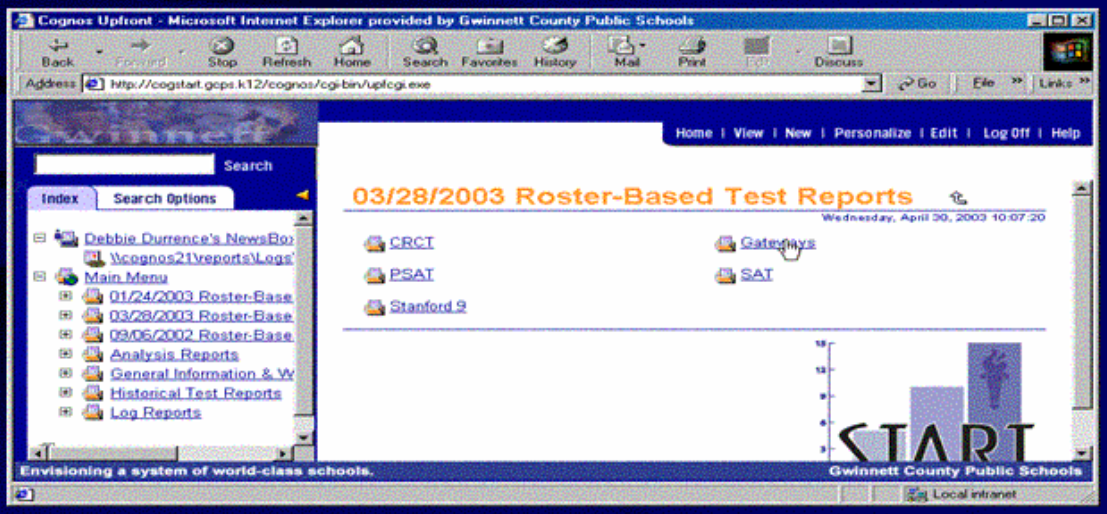
Viewing Individual Student Data

- Step 1:Select a Roster-based Report



Viewing Individual Student Data

- Step 2:Select a test

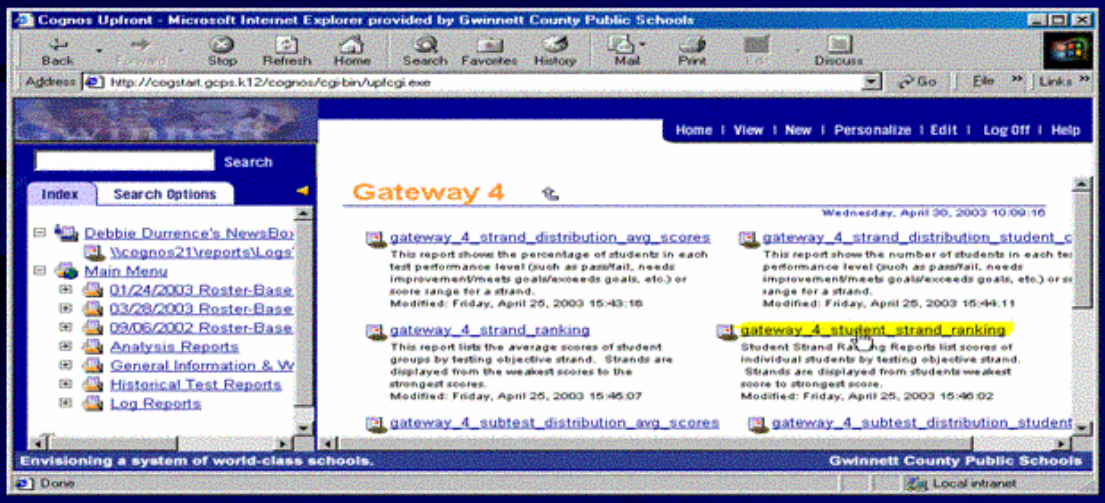


Viewing Individual Student Data

- Step 3:Select Gateway 4

Viewing Individual Student Data

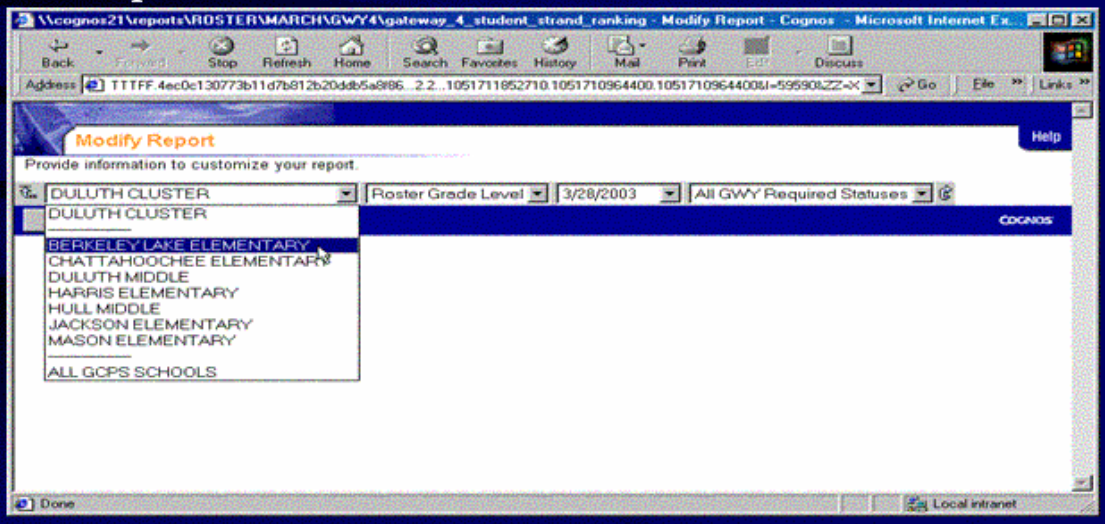
- Step 4:Select Student Strand Ranking Report



19

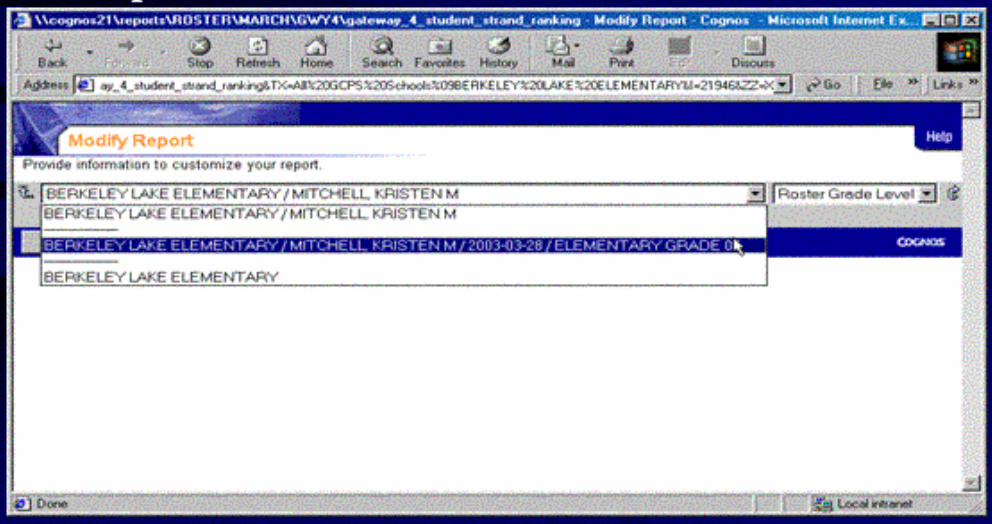
Viewing Individual Student Data

- Step 5: Select Cluster and then School



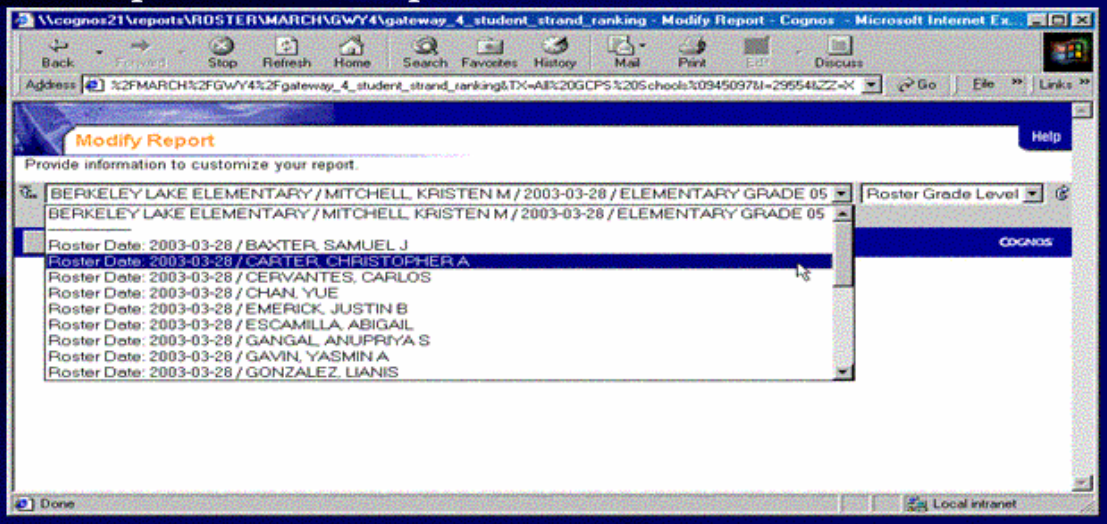
Viewing Individual Student Data

- Step 6: Select Teacher and then Class



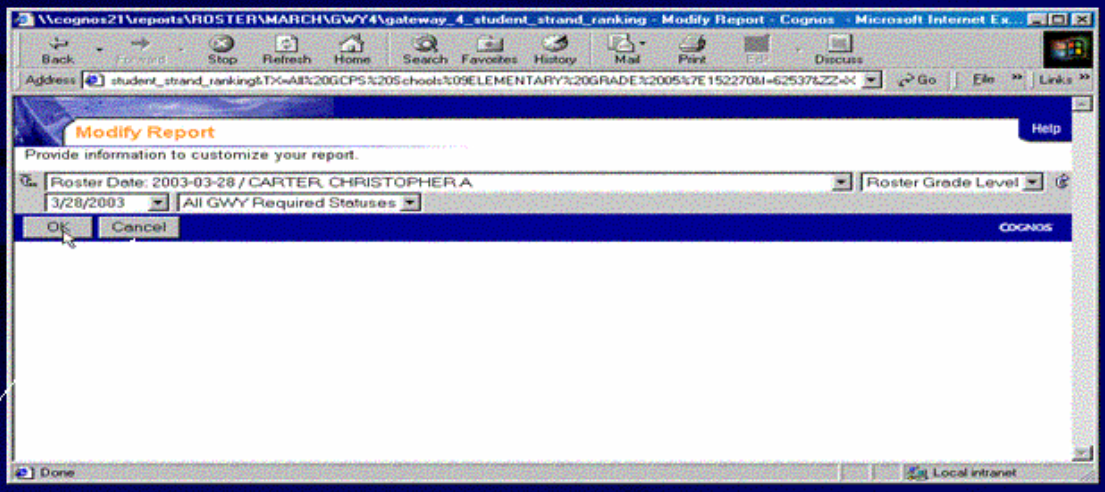
Viewing Individual Student Data

- Step 7: Select a specific student



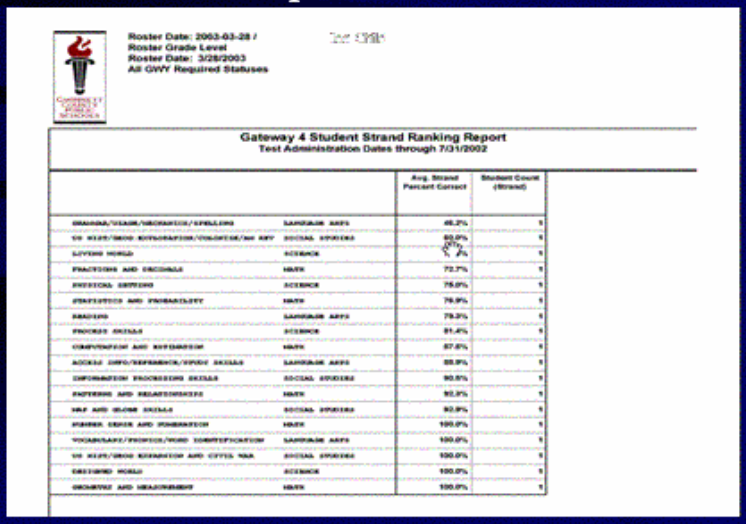
Viewing Individual Student Data

- Step 8: Click on OK to run the report



Viewing Individual Student Data

- Student Report



Roster Date: 2003-03-28 /
Roster Grade Level
Roster Date: 3/28/2003
All GWY Required Statuses

Gateway 4 Student Strand Ranking Report
Test Administration Dates through 7/31/2002

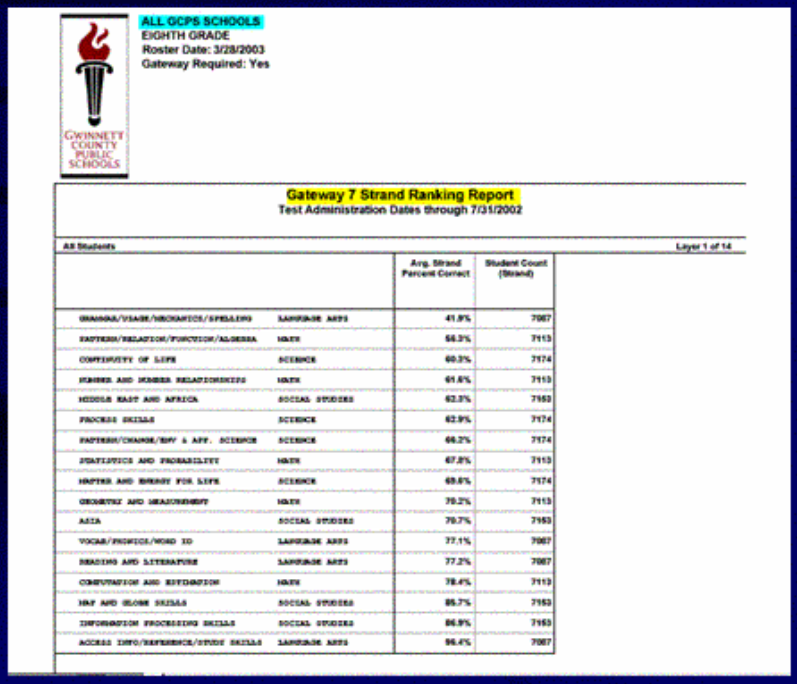
		Avg. Strand Percent Correct	Student Count (Strand)
[illegible]	LANGUAGE ARTS	46.2%	1
[illegible]	SOCIAL STUDIES	60.0%	1
[illegible]	SCIENCE	[illegible]	1
[illegible]	MATH	72.7%	1
[illegible]	SCIENCE	75.0%	1
[illegible]	MATH	76.9%	1
[illegible]	LANGUAGE ARTS	79.2%	1
[illegible]	SCIENCE	81.4%	1
[illegible]	MATH	87.5%	1
[illegible]	LANGUAGE ARTS	88.9%	1
[illegible]	SOCIAL STUDIES	90.5%	1
[illegible]	MATH	92.9%	1
[illegible]	SOCIAL STUDIES	92.9%	1
[illegible]	MATH	100.0%	1
[illegible]	LANGUAGE ARTS	100.0%	1
[illegible]	SOCIAL STUDIES	100.0%	1
[illegible]	SCIENCE	100.0%	1
[illegible]	MATH	100.0%	1

Viewing Strand Distribution Scores

- Follow steps 1-3 from the Individual Student Instructions
- Select Strand Distribution Average Scores Report
- Select
  - Cluster, School or Teacher
  - Strand
  - Grade level
- Click OK to run the report

25

Sample Strand Ranking Report



ALL GCPS SCHOOLS
EIGHTH GRADE
Roster Date: 3/28/2003
Gateway Required: Yes

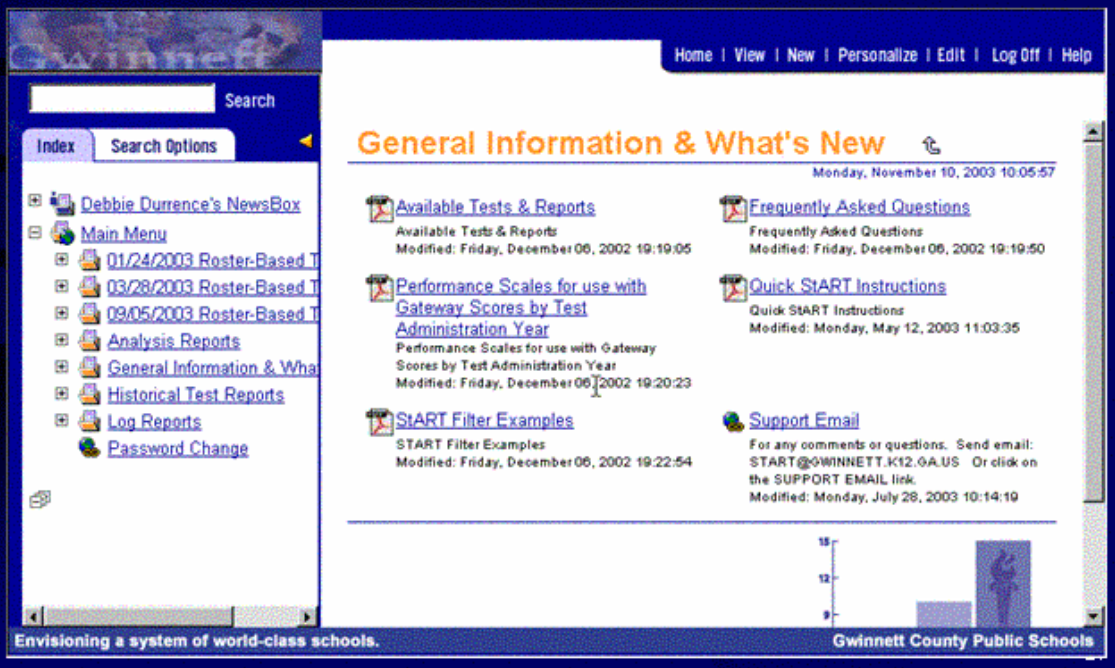
Gateway 7 Strand Ranking Report
Test Administration Dates through 7/31/2002

All Students — Layer 1 of 14

		Avg. Strand Percent Correct	Student Count (Strand)
GRAMMAR/USAGE/MECHANICS/SPELLING	LANGUAGE ARTS	41.9%	7087
PATTERN/RELATION/FUNCTION/ALGEBRA	MATH	58.3%	7113
CONTINUITY OF LIFE	SCIENCE	60.3%	7174
NUMBER AND NUMBER RELATIONSHIPS	MATH	61.6%	7113
MIDDLE EAST AND AFRICA	SOCIAL STUDIES	62.3%	7153
PROCESS SKILLS	SCIENCE	62.9%	7174
PATTERN/CHANGE/ENV & APP. SCIENCE	SCIENCE	66.2%	7174
STATISTICS AND PROBABILITY	MATH	67.8%	7113
MATTER AND ENERGY FOR LIFE	SCIENCE	69.6%	7174
GEOMETRY AND MEASUREMENT	MATH	70.2%	7113
ASIA	SOCIAL STUDIES	70.7%	7153
VOCAB/PHONICS/WORD ID	LANGUAGE ARTS	77.1%	7087
READING AND LITERATURE	LANGUAGE ARTS	77.2%	7087
COMPUTATION AND ESTIMATION	MATH	78.4%	7113
MAP AND GLOBE SKILLS	SOCIAL STUDIES	85.7%	7153
INFORMATION PROCESSING SKILLS	SOCIAL STUDIES	86.9%	7153
ACCESS INFO/REFERENCE/STUDY SKILLS	LANGUAGE ARTS	90.4%	7087

26

More information within StART



Guide to Elementary School Data in StART

Data Questions (What do we want to know?)	Test Report Available	Tests Available & Grades Administered
How ***many students*** were in each performance level for each ***subtest***?	Subtest Distribution Student Counts (TABLE)	GATEWAY – 4 CRCT** –1,2,3,4
What is the ***percentage*** of students in each performance level for each ***subtest***?	Subtest Distribution Average Scores (GRAPH)	GATEWAY – 4 CRCT** –1,2,3,4
How ***many students*** were in each performance level for each ***strand***?	Strand Distribution Student Counts (TABLE)	GATEWAY – 4 CRCT** –1,2,3,4
What is the ***percentage*** of students in each performance level for each ***strand***?	Strand Distribution Average Scores (GRAPH)	GATEWAY – 4 CRCT** –1,2,3,4
What is the list of ***strands*** from weakest to strongest for a ***group*** of students (class, grade, school, etc.) or an individual student?	Strand Ranking Report (LIST)	GATEWAY – 4 CRCT** –1,2,3,4
What is the ***average*** score for ***several years*** in each subtest at the school or teacher level?	Subtest Historical Profile (GRAPH) *available for system, school & teacher	GATEWAY – 4
What is the ***average*** number of students scoring in each performance level for ***several years*** in each subtest at the school or teacher level?	Subtest Historical Profile (TABLE) *available for system, school & teacher	GATEWAY – 4
What is the ***average*** scale score for ***several years*** in each strand at the school or teacher level?	Strand Historical Profile (GRAPH) *available for system, school & teacher	GATEWAY – 4
What is the ***average*** distribution of scores for ***several years*** in each subtest at the school or teacher level?	Strand Historical Profile (TABLE) *available for system, school & teacher	GATEWAY – 4

***Teacher historical reports include all students through the school year, the year the test was administered**

****CRCT Domains are identified as CRCT Strands in START**

Strand Ranking reports **Historical Reports (TREND ANALYSIS)**

APPENDIX D

ANOVA AND t-TEST TABLES FOR NON-SIGNIFICANT RESULTS

Table 41

Analysis of Variance for Interpreting Reports and Education Level

Report Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	44.755	2	22.378	1.578	.209
Within Groups	2438.525	172	14.177		
Total	2483.280	174			

Table 42

Analysis of Variance for Instruction Score and Education Level

Instruction Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	58.546	2	29.273	1.794	.169
Within Groups	2806.003	172	16.314		
Total	2864.549	174			

Table 43

Analysis of Variance for Accessing StART Score and Years of Experience

Access Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	90.168	2	45.084	2.301	.103
Within Groups	3370.689	172	19.597		
Total	3460.857	174			

Table 44

Analysis of Variance for Interpreting Reports Score and Years of Experience

Report Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	24.987	2	12.493	.874	.419
Within Groups	2458.293	172	14.292		
Total	2483.280	174			

Table 45

Analysis of Variance for Instruction and Years of Experience

Instruction Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15.484	2	7.742	.467	.627
Within Groups	2849.065	172	16.564		
Total	2864.549	174			

Table 46

Analysis of Variance for Total Score and Years of Experience

Total Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	39.235	2	19.617	.257	.774
Within Groups	13137.943	172	76.383		
Total	13177.177	174			

Table 47

Analysis of Variance for Interpreting Reports Score and School Level

Report Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	14.280	2	7.140	.497	.609
Within Groups	2469.000	172	14.355		
Total	2483.280	174			

Table 48

Analysis of Variance for Instruction Score and School Level

Instruction Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	14.836	2	7.418	.448	.640
Within Groups	2849.712	172	16.568		
Total	2864.549	174			

Table 49

Analysis of Variance for Total Score and School Level

Total Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	298.424	2	149.212	1.993	.139
Within Groups	12878.753	172	74.876		
Total	13177.177	174			

Table 50

t-test for Accessing StART and Title I Status

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Access Score	Equal variances not assumed	.248	28.177	.806	.23	.917	-1.652	2.106

Table 51

t-test for Interpreting Reports and Title I Status

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Report Score	Equal variances not assumed	.960	26.272	.346	.82	.853	-.933	2.570

Table 52

t-test for Instruction Score and Title I Status

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Instruction Score	Equal variances not assumed	-1.659	25.252	.110	-1.61	.971	-3.609	.388

Table 53

t-test for Total Score and Title I Status

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Total Score	Equal variances not assumed	-.114	30.577	.910	-.19	1.637	-3.527	3.155

Table 54

t-test for Accessing StART and Leadership Tenure

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Access Score	Equal variances not assumed	-1.795	90.983	.076	-1.37	.762	-2.880	.145

Table 55

t-test for Interpreting Reports and Leadership Tenure

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Report Score	Equal variances not assumed	1.074	79.107	.286	.75	.695	-.637	2.131

Table 56

t-test for Instruction Score and Leadership Tenure

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Instruction Scor	Equal variances not assumed	.880	86.693	.381	.63	.714	-.791	2.048

Table 57

t-test for Total Score and Leadership Tenure

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Total Score	Equal variances not assumed	.296	80.727	.768	.47	1.591	-2.695	3.637