

Rowan University

Rowan Digital Works

Theses and Dissertations

5-9-2005

Investigation into the effects of an after school GEPA math program in raising student achievement

Denise J. Costigan
Rowan University

Follow this and additional works at: <https://rdw.rowan.edu/etd>



Part of the [Disability and Equity in Education Commons](#)

Let us know how access to this document benefits you - share your thoughts on our feedback form.

Recommended Citation

Costigan, Denise J., "Investigation into the effects of an after school GEPA math program in raising student achievement" (2005). *Theses and Dissertations*. 983.
<https://rdw.rowan.edu/etd/983>

This Thesis is brought to you for free and open access by Rowan Digital Works. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Rowan Digital Works. For more information, please contact LibraryTheses@rowan.edu.

**INVESTIGATION INTO THE EFFECTS OF AN AFTER SCHOOL GEPA
MATH PROGRAM IN RAISING STUDENT ACHIEVEMENT**

by
Denise J. Costigan

A Thesis

Submitted in partial fulfillment of the requirements of the Master of Arts Degree
of
The Graduate School
at
Rowan University
April 2005

Approved by:

Date Approved May 9, 2005

© 2005 Denise J. Costigan

ABSTRACT

Denise J. Costigan

INVESTIGATION INTO THE EFFECTS OF AN AFTER SCHOOL GEPA MATH PROGRAM IN RAISING STUDENT ACHIEVEMENT 2004/05

Dr. Stanley Urban
Master of Arts in Learning Disabilities

The intent of this paper is to determine the effectiveness of an after school math program on raising student achievement on the Grade Eight Proficiency Assessment (GEPA). Regular education students were selected based on a review of their fourth grade standardized test, and the trend in their yearly Terra Nova scores from 5th, 6th and 7th grade, if available. Based on a proficient score of 200, students who scored in the range of 185 to 210 were invited to attend the eighteen-week after school program. In addition, all special education students were invited to attend. Forty-five students elected to participate in the voluntary program.

Students took a pretest at the onset of the program and a posttest at the conclusion and results were tabulated and analyzed. The results supported the continuation of the after school program for both populations of students. The percentage change calculation in student scores on the pretest and posttest provided evidence of the program's effectiveness. One major variable, student participation rate was analyzed to help

determine the relationship between performance and attendance. The data suggests that those students who attended greater than fifty percent of the time showed a marked improvement over those students who attended less than fifty percent.

Table of Contents

CHAPTER 1	1
INTRODUCTION.....	1
BACKGROUND.....	1
PURPOSE OF THE STUDY	3
ASSUMPTIONS	4
NEED OF THE STUDY	4
VALUE OF THE STUDY.....	4
RESEARCH QUESTIONS	5
Research Question 1:	5
Research Question 2:	5
DEFINITION OF TERMS.....	5
Standardized Tests.....	5
Norm referenced tests.....	6
National Assessment of Educational Progress.....	6
Core Curriculum Content Standards.....	6
Learning Disability (LD)	7
LIMITATIONS OF THE STUDY	7
CHAPTER 2	8
REVIEW OF THE LITERATURE.....	8
SAFEGUARDS/ACCOUNTABILITY.....	10
CONTROVERSIAL ISSUES.....	11
NEW JERSEY TESTING PROCEDURES	11
TEST PREPAREDNESS.....	12
LATEST UPDATE.....	13
STATE ASSESSMENT	14
TRENDS/ CHANGES IN TEST FORMAT	15
FOCUS OF MATHEMATICS ASSESSMENT	16
UNDERSTANDING STATE ASSESSMENT/STANDARDS	16
EXAMPLES OF STATE TESTING	18
POTENTIAL HAZARDS OF EXCESSIVE TESTING.....	18
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS	19
WHY ASSESSMENT IS IMPORTANT.....	20

Table of Contents (Contd.)

CHAPTER 3	22
DESIGN OF THE STUDY	22
POPULATION	22
METHOD OF SAMPLE SELECTION	22
INSTRUMENTATION	23
COLLECTION OF DATA.....	24
RESEARCH DESIGN.....	24
CHAPTER 4	26
ANALYSIS OF RESULTS	26
VARIABLES	35
CHAPTER 5	36
SUMMARY	36
REFERENCES.....	41

List of Tables

TABLE 1 - ALL PARTICIPATING STUDENTS PRETEST AND POSTTEST RESULTS	27
TABLE 2 - SPECIAL EDUCATION STUDENT PRETEST AND POST-TEST RESULTS	30
TABLE 3 - REGULAR EDUCATION STUDENT PRETEST AND POSTTEST RESULTS	31
TABLE 4 – ALL STUDENTS WITH 50% OR > PARTICIPATION RATE	33
TABLE 5 – ALL STUDENTS WITH 50% OR < PARTICIPATION RATE	34
TABLE 6 – STUDENT PROGRESS BY PARTICIPATION RATE	39

Chapter 1

Introduction

Background

In 1975, the New Jersey Legislature enacted the Public School Education Act which guaranteed all children in New Jersey the educational opportunity and knowledge of how to function in a democratic society. An amendment was added in 1976 that established uniform standards of minimum achievement in communications and computation skills. This amendment legalized the use of a test as a graduation requirement in New Jersey.

The “standards based” reform movement of the 1990’s brought major changes to most state public education systems. In 1996, the New Jersey State Board of Education adopted the Core Curriculum Content Standards (CCCS) which delineated what New Jersey students should know or be able to perform at the end of 4th, 8th and 12th grades. “All New Jersey school districts are required to organize instruction and design curricula so that virtually all students achieve the new content standards” (NJ DOE GEPA Manual, 2004). Subsequently, New Jersey altered its statewide student assessment to align tests with the new standards. The hope was that the test results would track the local districts success in achieving benchmarks and would also enable educators to provide remedial help to students in need.

The testing programs used to measure benchmark achievement towards these requirements are the Elementary School Proficiency Assessment (ESPA), now known as

New Jersey Assessment of Skills and Knowledge (NJASK) in fourth grade, the Grade Eight Proficiency Assessment (GEPA), which replaced the Grade Eight Early Warning Test (EWT), and the High School Proficiency Assessment (HSPA), which will replace the Grade Eleven High School Proficiency Test (HSPT 11). All of these assessment instruments are aligned with New Jersey CCCS.

The GEPA, first administered in 1999, provides valuable information about individual student progress toward mastery of the CCCS. It is used for placement purposes and program planning and is geared toward enabling students to pass the state required graduation test. All eighth grade public school students must take the GEPA. The only exception is students who participate in the Alternate Proficiency Assessment (APA). The APA is for students whose disabilities are so severe that they are not receiving instruction in skills measured by the statewide assessment, and cannot complete any of the questions in the content areas. Categories of students tested include:

- General Education student,
- Limited English Proficient (LEP) students,
- Special Education (SE) students,
- Students with Disabilities (Section 504 of Rehabilitation Act of 1973).

The three content areas tested are Science, Mathematics and Language Arts Literacy. The Math content area has 40 multiple choice and 8 open ended questions, and lasts 2 hours and 27 minutes. Test items are highly confidential because some items may appear on future versions of the exam. This confidentiality maintains the stability of the test item pool over time and allows for comparisons from year to year.

The new federal education law, *No Child Left Behind (NCLB) Act* mandates its own accountability system for any New Jersey school that receives Title I funds. Under this system:

- DOE identifies schools in need of improvement, based on student performance on the states 4th and 8th grade assessments;
- Schools identified must develop plans to show what actions they will take to improve teaching and learning;
- Districts identified as in need of improvement must develop plans for “intra-district choice”;
- DOE identifies additional schools in need of improvement based on test results for subgroups of students such as Limited English Proficiency (LEP), low income, disabilities or Special Education (SE);
- Schools who fail to meet their “annual yearly progress” targets for three years must provide supplemental services to students including tutoring at public expense;
- Schools that fail to meet targets for four years are subject to severe “corrective action” and can be closed.

Purpose of the Study

The purpose of this study is to empirically demonstrate that selected students who attend an after school GEPA support program will improve their scores to the proficient range (200+) due to additional instructional time and review of test taking strategies. Also, their experience with sample GEPA tests should reduce test anxiety and increase familiarity with test format.

Assumptions

The major premise of this study is that, an after school support program in math, offered to select borderline students, (defined as -15 points below and +10 above the average score of 200 on the ESPA), will help boost their scores on standardized state testing. Common folk wisdom states that one can not study for this type of test, however, "Having said that, though, there's no question that test preparation can make a significant difference in your child's performance" (Cookson & Halberstam,1998).

Need of the study

This study will help justify the additional cost of teacher time allotted to the program. It will also help district's meet the continuing requirement of NCLB that of ensuring that all children are making progress.

Value of the study

The major value of this study is that individual student progress is tracked so that deficiencies can be remediated and consequently more students will meet the criterion of "proficient" on their GEPA test. Overall, the district should experience an improvement in state testing scores. In addition, areas of weakness can be communicated to math instructors so that classroom curriculum is aligned with the test requirements. Lastly, the district will continue to meet the National Assessment of Educational Progress (NAEP) goals for continuing success. This study will provide data to support continuation of an after school support program. If successful, the district may wish to consider the possible benefit, i.e. improved state assessment scores, and offer such a program to 6th and 7th grade students, as well as, 8th graders.

Research Questions

In order to accomplish the purpose of this study the following research questions will be answered.

Research Question 1:

Can selected eighth grade students who scored in the borderline range of between 185 to 210 on the math section of the ESPA (200 being considered proficient or the mean), score in the proficient range on the GEPA (eighth grade test) after attending a specialized after school program?

Research Question 2:

What is the relationship between participation in an after school support program and success on the GEPA, when compared to the improvement on the GEPA for invited students who do not attend the after school specialized support program on a consistent basis (>50% participation rate)?

Definition of Terms

The following terms have a specific definition within the framework of this study.

Standardized Tests

Standardized tests such as the GEPA, are a form of assessment that has been normed against a specific population, so that subsequent student performance on the test can be meaningfully scored and interpreted (Cookson & Halberstam, 1998). They are used by the U.S. and other countries to assess children's abilities and achievement.

Norm referenced tests

Norm referenced tests such as the Terra Nova, are based on a large national sample of students in grades K-12, stratified by geographic region, community type, district size and socioeconomic status. Test scores are reported as percentile rankings. The percentile figure reflects a comparison of students performance with that of the “norm” or average student at that particular grade level (Cookson & Halberstam, 1998). The purpose is to help teachers help students to achieve academic goals.

National Assessment of Educational Progress

NAEP is a nationally administered assessment program designed to measure educational accomplishment and progress of selected U.S. students in 4th, 8th, and 12th grade students over time, in selected content areas (Cookson & Halberstam, 1998). Their goal is to gain statistical information on the level of American student achievement. This data is published annually by the U.S. Department of Education’s Public Affairs Office.

Core Curriculum Content Standards

CCCS were adopted in 1996 by New Jersey and delineate the expectations of what all students need to know in the various content areas at various grade levels. Content standards are used increasingly to direct curriculum and to define what schools need to teach and when they need to teach it.

Learning Disability (LD)

LD is a deficiency in some step of the learning process from intake of information to storing and retrieval of information. “These children show a measurable gap between potential and performance” (Cookson & Halberstam, 1998).

Limitations of the study

A limitation of this study is that official GEPA test results will not be available before completion of this thesis; therefore, the research design includes a pretest and posttest simulated GEPA test administered and scored within the district. Other limitations center around inability to control other variables which might effect the results, such as outside tutoring, teachers whose curriculum is more closely tied to the test, and attendance/participation rate during the program.

Chapter 2

Review of the Literature

This chapter will review the legal requirements related to participation in various assessments including: Section 504 of the Rehabilitation Act (1973), Americans with Disabilities Act (ADA), Individuals With Disabilities Education Act (IDEA), Goals 2000 and No Child Left Behind (NCLB). In addition, it will include trends in state assessments, a review of how various states are meeting the new challenges, the debates surrounding testing, some of the recommendations for preparedness for testing and lastly, some of the hazards associated with testing.

History/ Legislation

In 1983, *A Nation At Risk*, a national report on the state of American education concluded that our educational system was failing significantly. In response, President George H. Bush's "Goals 2000/Educate America", a national educational objective was created in 1991 and called for the adoption of a national standard. Goals 2000 provided the impetus for states to adopt curriculum standards that described, year to year what students would study in each subject area (Cookson & Halberstam, 1998). It was further felt that the adoption of national goals would ensure that all schools throughout the U.S. would maintain equivalent minimum levels of instructional quality with all students being taught to the same standards and assessed by standardized tests. Subsequently, President Clinton's "Voluntary National Test Program" reintroduced the concept of establishing a national standard that these tests are designed to assess. "These exams are based on the National Assessment of Educational

Progress, or NAEP” (Cookson & Halberstam, 1998). Some legislators view it as an encroachment on states and local districts rights, while others see it as a “unifying and vital step toward improved education for all children” (Cookson & Halberstam, 1998).

The New Jersey Statewide Systematic Initiative (NJ SSI) has created software that imports files from a vendor’s data disks and can be used to generate charts and tables that present the data based on several variables. Using these charts, administrators can see what a targeted group of students don’t know in what particular area, taught by what teacher, and can analyze the causes and take corrective action. The collection and dissemination of quality information will aid in accountability. Accountability requires continued tracking of individual student performance data throughout the students school career. New Jersey spends approximately \$17 million per year on testing programs. However, if accountability systems are not in place, and if test results are not used in education planning, it’s money wasted. Test results often arrive too late for remediation, scheduling of fall classes, or appropriate placement of students, or for the scheduling of staff training opportunities during the summer.

It was clear that simply increasing funding would not improve student performance, rather it was how states targeted their resources that was important. No Child Left Behind expands upon Goals 2000, which officially began standards-based educational reform in the U.S. and required testing of students at three points during their school years (Allen, 2004).

According to Allen, testing high school students in the 1970’s was followed by an eighth grade test. When the Core Curriculum Content Standards were introduced in 1996 in response to Goals 2000, a fourth grade state assessment was implemented.

The No Child Left Behind Act was signed into law on January 8, 2002 with the mission to improve schools nationwide. Provisions of the act include mandated annual testing of all students in grades 3 through 8 in reading and math. This provision was to ensure that children do not fall behind in school. The Act also ties federal funding to the test results. NCLB sets escalating targets that schools must reach when they administer state tests to students. By 2014, all students will be expected to perform at or above the proficient level on state tests. It is, however, up to the states to define what “proficient” means, and those definitions vary widely (Gewertz, 2004).

Safeguards/Accountability

The first time that there was a large scale commitment to accountability for results in return for government financial assistance was in the 1960’s, with the Title I program of federal aid to schools with low income students. The need for educational accountability was highlighted by Goals 2000 and Improving America’s School Act. Up until 1995, the IEP team was the primary decision-maker about whether or not students with disabilities participated in achievement assessment.

The 1997 amendment to IDEA makes Part B funding contingent upon participation of students with disabilities in state assessments, with accommodations or alternate assessment. Accommodations include changes to standardized testing conditions that prevent irrelevant factors from affecting test scores. According to Yen and Henderson (2002), some of the more common accommodations include extended time, small group or individual settings, and the response format using an alternate method of communication or presentation, which includes Braille or large print versions.

States are required to report on participation and performance. Test results must be disaggregated demographically by race, poverty, English proficiency, gender and disability (Education USA 2002). Varying state policies on participation and accommodations will have a major impact on the reporting process. While most states offer accommodations, there was little agreement or consistency in the acceptability of various accommodations. This makes for added difficulty when attempting to compare performance across states.

Controversial Issues

The most important issue revolves around whether it is more psychometrically sound to base decision making on a smaller number of students, e.g. the general education students who participate fully in a non-accommodated tests, or to base decisions on all students, some of whom have experienced changes to the test. The debate continues very similarly to the debate on instructional models of inclusion.

New Jersey Testing Procedures

The New Jersey Department of Education (NJ DOE) releases annual reports on how public schools in the state fared on meeting the federally mandated standards under the No Child Left Behind Act. All public schools, including charters, are evaluated based on their standardized test scores. The federal law requires that schools show “adequate yearly progress” (AYP) toward 100% proficiency by 2014. AYP results are based on year to year comparisons of school scores on High School Proficiency Assessment (HSPA) administered in the 11th grade, Grade Eight Proficiency Assessment, and New Jersey Assessment of Skills and Knowledge, given in 3rd and 4th grades. To achieve AYP, the school’s students must

meet both proficiency targets and 95% participation rates in both Language Arts Literacy and Mathematics for each of ten sub-groups.

GEPA is part of the NJ. DOE's strategic plan to meet the federal testing mandates of NCLB. GEPA replaced the Early Warning Test which had been administered from 1991-1998. GEPA's primary purpose was for identifying eighth grade students who may need instructional intervention in three content areas, Math, Language Arts Literacy and Science. In March of 1999, the first GEPA was administered to all eighth grade students except those taking the APA. In 1999, only Language Arts Literacy and Math were tested. Science was added in the 2000 administration. The 1999 GEPA was the first assessment measuring progress toward CCCS. The results will differ from past results on commercially available norm-referenced tests or state's EWT, because New Jersey raised the bar with a more demanding curriculum and assessment. In 2002, charts were recast to include the scores of all students including, Special Education and English as a Second Language, as required by the provision of the federal NCLB legislation.

Test Preparedness

Test preparedness includes an understanding of the purpose of the test, becoming familiar with the various kinds of test items (open-ended versus multiple choice), and understanding the direction and scoring procedures used. States should ensure that parents and students are aware of test requirements. In this regard, Voorhees Middle School (VMS) has two GEPA nights, open to parents and their eighth grade student(s), that present sample questions from GEPA Math and Language Arts Literacy. In addition, our GEPA after school program,

which runs for an hour per week, for eighteen weeks, is offered to increase familiarity with test format and content, and to provide practice on test taking strategies.

Latest Update

South Jersey's Report Card on schools in Burlington, Camden and Gloucester counties that did not meet all the standards, was published in the Philadelphia Inquirer on September 30, 2004. It listed the various sanctions if a school misses any of the 40 indicators and the consequences:

- ✓ Early Warning- no sanctions;
- ✓ Needs Improvement-school has not met requirement of AYP for one year-no sanction;
- ✓ Choice-no school improvement for two consecutive years-parents notified and given transfer option within district to a school where AYP is being made;
- ✓ Supplemental Services-school fails to meet AYP for three consecutive years-must continue to offer transfers and begin offering tutoring and supplemental educational services;
- ✓ Corrective Action-schools failing to make progress for four consecutive years-school must implement one of the following: replace staff relevant to the failure to make AYP; institute a new curriculum, including professional development; decrease management authority at school level; appoint an outside expert to advise the school; extend the school year or school day; restructure the school's internal structure.

An article which appeared in the Camden Courier Post (2004), stated that 109 South Jersey schools failed to meet the minimum education standards set by the federal government.

However, there were 48 schools in South Jersey that improved enough to be removed from the “early warning” list of lagging schools. Schools are removed from the list when they meet the standards for two consecutive years.

This is the second year that New Jersey has identified schools that need improvement under the federal guidelines. State Education Commissioner William L. Librera is proud of NJ’s progress and critical of some portions of the federal act. Librera attributed statewide gains to changes implemented by schools to boost performance. Modified guidelines also contributed to the progress, as the size of subgroups doubled from 20 to 40, meaning schools with smaller classes were not evaluated using the 40 benchmarks, and were deemed as making adequate progress. Another revision is that the performance of Special Education students counted in 2003, only if there were at least 35 students in a class. Interestingly enough, middle schools led the list of those not meeting AYP.

State Assessment

The states and school districts select which standardized test are administered. This process is often guided as much by “economic and political factors as by educational need” (Cookson & Halberstam). Whether or not standardized tests measure what they claim to measure is an ongoing debate. Some states favor a lot of testing and some states use them restrictively, including only a few age groups.

Large-scale state assessments are increasing in use in education and are being used in helping to make decisions about individual students. Appropriate interpretation and use is

necessary and critical. Most high stakes applications (when test results contribute to student promotion or graduation decisions), require that students reach a performance standard or predetermined cutoff score on a test. Test scores are also used to identify student's strengths and weaknesses so that appropriate instructional planning is implemented.

The validity of a test or its ability to measure adequately what it purports to measure, is a central standard of all tests. In addition, the test design should incorporate a description of the constructs, i.e. knowledge and skills that it intends to measure. Reliability and validity are often sacrificed due to cost factors. Content validity of state assessments needs to be evaluated in light of their content standards. Gaps between test results and real learning need to be continually evolving. Ideally, there should be no "teaching to the test", only teaching to the content standards represented by the test. Accountability systems are being installed enthusiastically, however, there is not yet any proof that they improve education, according to Sheppard (2002-03). Sheppard also believes that high stake testing harms the climate for teaching and learning.

Trends/ Changes in Test Format

State testing programs hold a lot of power. They can determine whether students are promoted or retained. They also affect curriculum and instruction. Since the 1920's the dominant form of standardized tests have been norm and criterion referenced multiple choice tests. Norm referenced tests compare scores with a national sample while criterion referenced tests compare students against a predetermined level of achievement.

In the 1990's based on new discoveries in cognitive psychology, education reformers began to change assessments so that performance was evaluated based on authentic tasks in real

life context (Neill, 2000). The switch occurred because it was believed that prior tests did not adequately measure critical thinking and problem solving skills. Performance tasks would involve issues that were meaningful to students. They would study fewer things but in greater depth. However, such change has not occurred in most state testing programs. “Today, all state programs include one or more of the following components: criterion and or norm-referenced multiple-choice items; some open-ended short answer questions; and a writing sample. While open ended questions are intended to measure critical thinking, in reality many open ended questions are similar to multiple choice with the answer options removed” (Neill, 2000). In the classroom, Neill suggests that teachers use comprehensive assessment strategies, i.e. observations of students, real investigations and reports by students in different subject areas, rubrics and portfolio or learning records.

Focus of Mathematics Assessment

The math focus has shifted from routine computation to conceptual insights and analytical skills. It is believed that the major purpose of math instruction is to help students solve math problems of everyday life; prepare them for jobs and help them to participate fully in civic affairs. Therefore, math assessment needs to change to a more authentic assessment. In authentic assessment the teacher’s role is that of a facilitator as students work to construct knowledge. The California Math Program has used authentic assessment for several years and it involves engaging students in thinking and reasoning.

Understanding State Assessment/Standards

The state Board of Education sets the standards used in state assessments and the process is usually a “criterion referenced process that relates judgements about the desired level of

student knowledge and abilities to a test cut score” (Yen and Henderson, 2002). In New Jersey, the *New Jersey Core Curriculum Content Standards* delineates what content needs to be taught and learned by students in order to be successful on the standardized test. Most state assessments are designed to measure proficiency in regards to challenging content standards. The standards define what students need to know but they don’t specify what strategy or pedagogy the districts must use. Districts, however, are under extreme pressure to perform well on state tests due to Annual Report Cards and increasingly severe sanctions under NCLB.

Tienken and Wilson (2001) recommend a three-part process of delineation, alignment and calibration for the development and understanding of state assessment and using that understanding to improve instruction. “Delineation is the process of identifying all aspects or dimensions of a particular subject domain” (Tienken and Wilson, 2001). The standards include knowledge specifications or content standards which describe the specific processes and content that all students must know by the end of a particular year, and problem solving specifications or process standards, which describe what students should be able to do with the content knowledge. Each subject tested has distinct dimensions to be mastered. Once the dimensions are known, instructional techniques that best teach them can be employed. Then, during the alignment phase, educators examine textbooks, materials and instructional strategies to ensure consistency with the district’s delineation of state assessment. The last phase, calibration, involves communication and interaction with teaching staff to ensure that the content, cognitive process, complexity etc. is understood for each subject area.

“Calibration is any action that helps teachers design activities and construct assessment based on the dimensions of state assessment and standards” (Tienken and Wilson, 2001).

Examples of State Testing

In Texas, teachers break down and teach skills and knowledge required to pass the Texas Assessment of Academic Skills or TAAS. In the months prior to TAAS, students who need extra help stay one hour longer in school. In Freeport, Texas teachers obtain student scores on the TAAS and break them into five groups reflecting their level of mastery. Based on the data teachers identify those particular areas that each group needs to work on to improve. A timeline for what they will teach and how much time will be spent on each objective is set. Each teacher receives an "instructional focus" sheet (Bradley, p.5) stating the objective to be taught. Students are then assessed and those that need extra help get it after school or during an extended day program.

Potential Hazards of Excessive Testing

Changes in what is taught are not necessarily good. The effect of high stakes accountability pressure has some teachers eliminating or reducing time spent on Social Studies, etc. to spend additional time on tested subjects. Test preparation often dominates instruction particularly in poorer districts. The best teachers always taught conceptual understanding in addition to teaching what they needed to know for the test.

In addition, states have the freedom to choose and design tests that will produce results that they desire of school progress. A new testing plan shows an understanding of the need for independent verification of reported test score gains in state accountability tests. For example, by setting performance standards that are low such as Texas did (25% of national standard), it is possible to show a dramatic improvement above the standard. Colorado, on

the other hand, set very high standards (90% of national performance). It is therefore difficult to interpret test results accurately due to the politics involved (Sheppard).

In New York, where schools received scores based on the performance of their students, there was evidence that some schools achieved high scores by either retaining in grade those who were not likely to perform well, or by referring students to special education, where students were exempted from testing (Elliott, Erieken, Thurlow and Shriner).

National Assessment of Educational Progress

Since 1969, the National Assessment of Educational Progress has been assessing what public and private school students can do in several curriculum areas including math, reading, science, writing, US history and geography. "NAEP is a congressionally-mandated project of the National Center for Education Statistics (NCES), United States, Department of Education" (Hanson, 2001). In 1988, Congress authorized voluntary participation in state level assessment during 1990 and 1992 and Educational Testing Services (ETS) was awarded the contract for conducting these assessments. ETS was responsible for developing the assessment instrument, analyzing the results and preparing reports on student achievement. The National Assessment Governing Board (NAGB) formulates the policy guidelines for NAEP and is composed of twenty-four members from a variety of occupations. Their job is to select the subject areas to be assessed, develop assessment objectives and specifications, identify appropriate achievement goals for students and ensure that all items are unbiased.

The NAEP is given to a sample of children across the country. Its sole purpose is to provide an accurate comparative picture of what students are learning throughout the country.

“NAEP is the one assessment mechanism that can be used to gauge real progress.” (Sheppard, 2002-03). “Comprehensiveness of NAEP content is critical to its role as an independent monitor of achievement trends.” (Sheppard, 2002-03). In order to track progress toward the eight national education goals, the National Education Goals Panel is compiling a survey of what each state is doing to measure academic achievement of its students, according to Lawton (1996). Since states use different batteries, state to state comparisons can’t be made, but the report will present a profile of each state’s assessment system, describing how often and what kinds of tests are used, which grade levels are tested and how the results are used.

NAEP is the only ongoing assessment of academic achievement and its findings are known as The Nation’s Report Card, which summarizes achievement across items and describes the relationship between achievement and a variety of background characteristics. The annual School Report Card allows parents and interested citizens the opportunity to review the progress of their local schools and evaluate whether or not they are providing students with high quality education in a cost-effective way (Bowman).

Why assessment is important

The NJ United for Higher School Standards web page (www.newjerseyunited.org) lists several reasons why assessment is important. They include:

- 1) raise expectations for all students and ensure that no child falls through the cracks;
- 2) provide individual feedback on student progress in achieving standards;

- 3) provides uniform system of determining whether all students are making satisfactory progress in learning skills in CCCS;
- 4) helps schools determine strengths and weaknesses of their programs;
- 5) furnished the DOE with an annual picture of states progress toward the goal of achieving standards;
- 6) helps to measure whether schools have shown Adequate Yearly Progress, the minimum level of improvement that states and schools must achieve each year according to NCLB.

Chapter 3

Design of the Study

Introduction:

In this Chapter, the population for the study is defined; the method used in selecting a sample for the population is described; the instrumentation used and the procedures for the collecting and analyzing the data are explained.

Population

The subject pool included the population of eighth grade students who attend Voorhees Middle School, Voorhees, New Jersey in both regular and Special Education during the 2004-05 school year.

Method of Sample Selection

All students whose standard scores on the ESPA (now NJ ASK) were between 185 and 210 were invited to participate in this after school program. A score of 200 indicates proficiency. A score below 200 indicates partial proficiency. In addition, all Special Education students who scored below 210 were invited to attend. In conjunction with ESPA scores, these same students' Terra Nova stanine scores from 6th and 7th grade, if available, were reviewed. The intent was to determine if they were maintaining a comparable level of proficiency in comparison with other students in regard to both local level and national level stanines

Instrumentation

The measure for the pretest and posttest was *Measuring Up® to the New Jersey Core Curriculum Content Standards* (Peoples Publishing Group, 2003) test 2, grade 8 which is a diagnostic practice test. This test consisted of 40 questions, including 34 multiple choice and 6 open-ended questions to be administered in 90 minutes. The Scoring guide (Peoples Publishing Group, 2003) includes a zero to three-point rubric for the open-ended questions. A three-point response would show complete understanding. A two-point response would show nearly complete understanding, a one-point response shows limited understanding and a 0-point response shows insufficient understanding. The answer guide also lists which New Jersey standard is being tested and the Cumulative Progress Indicators in those questions.

Individual performance on these Diagnostic Practice Tests can be interpreted based on ranges. Since only half of the test was administered the ranges were adjusted accordingly. Based on one point for each multiple choice question (17 points) and three open ended with a maximum of three points each (9 points) a maximum of 26 points can be earned. A score in the 0-10.75 range indicates that there is a good chance that the student would not score at the proficient level. A score between 11.0-15.75 indicates that the student would score either just above or just below the proficient level cut-score. A score in the range of 16.0-26.0 would be a good indicator that the student is at or above the proficient level.

These practice tests cannot totally simulate the “real” GEPA test and cannot therefore be used to predict student scores on the GEPA. Several factors influence student performance that are outside the realm of control of this study. The first is that student performance on

any test can vary from day to day. Secondly, these tests are administered under less standardized conditions than those used during the real test. Thirdly, the tests are scored locally by two math educators, without the extensive training and accuracy controls used to score the live test. Finally, continued instruction will occur between the initial pretest and the posttest.

Collection of Data

In October 2004, a GEPA prep test prepared by Measuring Up® was administered to all students participating in the after school program. Students who attended the program were given 45 minutes to complete the test and asked to complete the odd numbered questions only, which included 17, multiple choice and three open ended questions. At the conclusion of the program in March 2005, all students who took the original test will be requested to complete the Measuring Up test's even numbered items. These tests will be scored and areas of weakness noted.

Research Design

This pretest/posttest design will be employed in order to measure gains attributable to this remedial program. In this matched pair t-test, the null hypothesis would state that there is no significant difference between pretest and posttest scores.

The variables in the above design include the participation rate of the chosen students, various teachers' expertise during the school year and student motivation. However, since standardized tests are independent and objective, they provide a good indication of whether or not remedial or enrichment programs are needed. If a student scores worse than expected on a standardized test based on his or her past performance it could have just been a "bad"

day, or it could signal a problem or difficulty with the new form of standardized test which could be remediated by test taking strategies and or familiarity with the test design and format, such as practice tests.

One of the major attributes of standardized tests are their ability to pinpoint specific deficits such as math computation or problem solving ability which can then become the focus of improved instructional time. Review and practice taking sample tests will familiarize students with the test format and enable them to become better test takers because of reduced anxiety and experience with the type of questions asked.

I will be comparing the results of the posttest administered in March, 2005 to the results obtained from the pretest given in early October, 2004. The objective is to show improvement in individual scores associated with the after school GEPA program. This improvement is important to the school district since NCLB stipulates that all students will be proficient by 2014. Schools are also being monitored for their continuing improvement and face severe consequences for failure to meet these objectives.

Chapter 4

Analysis of Results

A tabulation of the results of all participants in the after school GEPA program is available in Table 1. Results include both the pretest and posttest GEPA simulated diagnostic practice test 2 by Measuring Up ®. This diagnostic practice test was correlated to the *New Jersey Core Curriculum Content Standards* for Mathematics and the associated Cumulative Progress Indicators.

Table 1 lists all 45 students, identified by student number, who participated to some degree in the after school GEPA program with their respective, pretest and posttest scores. The scores are broken down into the total number correct for both multiple choice (MC) and open-ended (OE), followed by the percentage correct. Both the pretest and posttest included a total of 17 MC questions, each counting as one point, and three OE questions worth a maximum of nine points based on a scoring rubric of zero to three. The pretest total is a combination of the MC and OE pretest scores. The posttest totals likewise consist of both the MC and OE posttest scores combined. The last column in table 1 represents the total overall percentage change in students scores between the initial testing, which took place in October of 2004, and the final testing which took place in March 2005.

Table 1 includes all students who participated for at least two days in the program. If students were present for only test taking (both pretest and posttest), it is assumed that they did not benefit from any instructional time offered in the after school GEPA program. They did however, continue to receive daily math instruction in their respective classrooms which would account for any improvement noted.

Table 1 - All Participating Students Pretest and Posttest Results

Student	Pretest Scores					Posttest Scores				Test Totals		
	PR	MC	%	OE	%	MC	%	OE	%	Pretest Total	Posttest Total	% Change
1	39	6	35	3	33	6	35	1.5	17	9	7.5	-17
2	56	4	24	0	0	3	18	0	0	4	3	-25
3	5	1	6	0	0	4	24	2	22	1	6	+500
4	100	6	35	2	22	10	59	3	33	8	13 *	+63
5	67	8	47	2	22	12	71	7	77	10	19**	+90
6	61	8	47	1	11	9	53	4	44	9	13 *	+44
7	61					4	24	4.5	5		8.5	
8	67	3	18	2	22	6	35	3	33	5	9	+80
9	64	2	12	0	0	5	29	3	33	2	8	+300
10	22	7	41	0	0					7		
11	17	6	35	4	44	9	53	4	44	10	13*	+30
12	44	7	41	4	44	6	35	6	67	11*	12*	+9
13	89	5	29	0	0	9	53	3.5	39	5	12.5*	+150
14	100	9	53	3	33	10	59	0	0	12 *	10	-17
15	44	3	18	0	0	6	35	0	0	3	6	+100
16	55	7	41	3	33	10	59	5.5	61	10	15.5*	+55
17	11	4	24	3	33					7		
18	5	7	41	2	22	8	47	7	78	9	15*	+67
19	67					9	53	3	33		12*	
20	89	1	6	2	22	6	35	2	22	3	8	+167
21	5	6	35	1	11	5	29	2	22	7	7	NC
22	67	2	12	1	11	6	35	2	22	3	8	+167
23	61	6	35	2	22	11	65	8.5	94	8	19.5**	
24	28	5	29	0	0	4	24	1	11	5	5	NC
25	61	6	35	2	22	8	47	4	44	8	12*	+63
26	5	3	18	1	11	2	12	1	11	4	3	-25
27	78	7	41	2	22	9	53	6	66	9	15*	+67
28	83	6	35	1	11	11	65	5.5	61	7	16.5**	+136
29	72	6	35	2	22	6	35	4	44	8	10	+25
30	89	2	12	2	22	9	53	6	66	4	15*	+275
31	6	3	18	2	22	9	53	4	44	5	13*	+160
32	11	5	29	0	0					5		
33	83	2	12	0	0	7	41	1	11	2	8	+300
34	89	6	35	2	22	12	71	4	44	8	16**	+100
35	89	5	29	2	22	6	35	2	22	7	8	+14

Table 1 - All Participating Students Pretest and Posttest Results (Contd.)

Student	Pretest Scores					Posttest Scores				Test Totals		
	PR	MC	%	OE	%	MC	%	OE	%	Pretest Total	Posttest Total	% Change
36	94	10	59	3	33	6	35	4	44	13*	10	-23
37	75	2	12	0	0	8	47	4	44	2	12*	+500
35	89	5	29	2	22	6	35	2	22	7	8	+14
36	94	10	59	3	33	6	35	4	44	13*	10	-23
37	75	2	12	0	0	8	47	4	44	2	12*	+500
38	72					6	35	2	22		8	
39	33					2	12	2	22		4	
40	72					9	53	6	66		15*	
41	5	8	47	2	22					10		
42	100	4	24	1	11	3	18	0	0	5	3	-40
43	67	2	12	1	11	5	29	3	33	3	8	+166
44	6	3	18	0	0	4	24	1.5	17	3	5.5	+83
45	44	7	41	3	33	7	41	7	77	10	14*	+40
	54.6									6.525	10.434	+60
N=45	Avg.									N=40	N=38	

PR = Participation Rate (number of days /hours out of a total of 18)
 MC = Multiple Choice Questions (Number correct out of a total of 17)
 OE = Open Ended Questions (Number points based on rubric total 9)
 * Borderline ** Proficient No asterisk Not Proficient

The chart below lists the ranges established by Measuring Up® for proficiency.

Ranges of level of proficiency as determined by Measuring Up® GEPA test	
00.0 – 10.75	Not Proficient (NP)
11.0 – 15.75	Just above or just below Proficient/Borderline (B)
16.0 – 26.00	At or above Proficient (P)

Students were selected to participate in the after school GEPA program based on the results achieved on the standardized test administered to them in fourth grade which was the Elementary School Proficiency Assessment or ESPA now the New Jersey Assessment of Skills and Knowledge. The range varied from 15 points below proficient or 185 points, to

10 points above proficient or 210 points. In addition, student scores on *Terra Nova* administered in 5th, 6th and 7th grades, if available, were considered as a means of determining a trend in individual performance on both a national and local level.

A significant proportion of students attending the after school GEPA program moved from a category of not proficient (NP) or borderline (B) before the program, to a category of borderline or proficient (P) after the program. Of forty students pretested, only three tested in the borderline range, with thirty seven scoring in the non proficient range. In the post-test survey, four students moved up to the proficient range, fourteen were borderline and 20 remained non-proficient, although several just missed the cut off to the borderline range. Twenty five students increased their performance score, anywhere from +9 % to +500%. Positive results were attributable to increased familiarity with the test format, ongoing review of math nomenclature, practice on sample questions and suggested test taking strategies.

Results were further differentiated into Special Education students (Table 2) and regular education students (Table 3), to determine if classification had any impact on the results. In addition, it would serve as documentation as to whether or not the program was more beneficial to one group. Both groups experienced an increase in performance. Special Education students increased in the aggregate 41%, and regular education students improved their overall performance by 66%, regardless of rate of participation.

Table 2 - Special Education Student Pretest and Posttest Results

Student	Pretest Scores					Posttest Scores				Test Totals		
	PR	MC	%	OE	%	MC	%	OE	%	Pretest Total	Posttest Total	% Change
1	39	6	35	3	33	6	35	1.5	17	9	7.5	-17
2	56	4	24	0	0	3	18	0	0	4	3	-25
3	5	1	6	0	0	4	24	2	22	1	6	+500
4	100	6	35	2	22	10	59	3	33	8	13	+63
8	67	3	18	2	22	6	35	3	33	5	9	+80
15	44	3	18	0	0	6	35	0	0	3	6	+100
19	67					9	53	3	33		12	
20	89	1	6	2	22	6	35	2	22	3	8	+167
24	28	5	29	0	0	4	24	1	11	5	5	NC
26	5	3	18	1	11	2	12	1	11	4	3	-25
27	78	7	41	2	22	9	53	6	66	9	15	+67
29	72	6	35	2	22	6	35	4	44	8	10	+25
32	11	5	29	0	0					5		
33	83	2	12	0	0	7	41	1	11	2	8	+300
41	5	8	47	2	22					10		
42	100	4	24	1	11	3	18	0	0	5	3	-40
43	67	2	12	1	11	5	29	3	33	3	8	+166
										N=16	N=15	
	56.9									5.25	7.77	+48

In the Special Education population (Table 2) improvement ranged from no change (NC) to 500% improvement. Four students showed negative results ranging from 0% to -40% decline in performance. These results can be attributable to low participation rate (PR) and the fact that no extended time considerations were given during administration of either pretest or post-test. In fact, student number 42, who showed the greatest decline, failed to complete a substantial portion of the post-test which skewed the results.

Table 3 - Regular Education Student Pretest and Posttest Results

Student	Pretest Scores					Posttest Scores				Test Totals		
	PR	MC	%	OE	%	MC	%	OE	%	Pretest Total	Posttest Total	% Change
5	67	8	47	2	22	12	71	7	77	10	19	+90
6	61	8	47	1	11	9	53	4	44	9	13	+44
7	61					4	24	4.5	5		8.5	
9	64	2	12	0	0	5	29	3	33	2	8	+300
10	22	7	41	0	0					7		
11	17	6	35	4	44	9	53	4	44	10	13	+30
12	44	7	41	4	44	6	35	6	67	11	12	+9
13	89	5	29	0	0	9	53	3.5	39	5	12.5	+150
14	100	9	53	3	33	10	59	0	0	12	10	-17
16	55	7	41	3	33	10	59	5.5	61	10	15.5	+55
17	11	4	24	3	33					7		
18	5	7	41	2	22	8	47	7	78	9	15	+67
21	5	6	35	1	11	5	29	2	22	7	7	NC
22	67	2	12	1	11	6	35	2	22	3	8	+167
23	61	6	35	2	22	11	65	8.5	94	8	19.5	+144
25	61	6	35	2	22	8	47	4	44	8	12	+63
28	83	6	35	1	11	11	65	5.5	61	7	16.5	+136
30	89	2	12	2	22	9	53	6	66	4	15	+275
31	6	3	18	2	22	9	53	4	44	5	13	+160
34	89	6	35	2	22	12	71	4	44	8	16	+100
35	89	5	29	2	22	6	35	2	22	7	8	+14
36	94	10	59	3	33	6	35	4	44	13	10	-23
37	75	2	12	0	0	8	47	4	44	2	12	+500
38	72					6	35	2	22		8	
39	33					2	12	2	22		4	
40	72					9	53	6	66		15	
44	6	3	18	0	0	4	24	1.5	17	3	5.5	+83
45	44	7	41	3	33	7	41	7	77	10	14	+40
										7.38	12.38	+68
	N=29									N=24	N=26	

To understand more completely the negative results posted by the four special education students and two regular education students, one of the variables analyzed was participation rate as seen in tables 4 and 5. Participation rate was one of the critical variables in this study. A cutoff of 50% participation was used as a benchmark for determining whether or

not the program met with success. Unfortunately, as the weeks wore on, student interest waned, with many students failing to show up for the program.

Other students had commitments such as after school sports, theater, etc. that interfered with their steady participation. Since participation was voluntary, no pressure was brought upon students to attend, at least not by the Voorhees Middle School staff. It was reported that one student's mother was forcing him to attend and he maintained a surly and negative attitude throughout the program. Tables 4 and 5 below break down student participation results.

Table 4 – All Students With 50% Or > Participation Rate

Student	Pretest Scores					Posttest Scores				Test Totals		
	PR	MC	%	OE	%	MC	%	OE	%	Pretest Total	Posttest Total	% Change
2	56	4	24	0	0	3	18	0	0	4	3	-25
4	100	6	35	2	22	10	59	3	33	8	13	+63
5	67	8	47	2	22	12	71	7	77	10	19	+90
6	61	8	47	1	11	9	53	4	44	9	13	+44
7	61					4	24	4.5	5		8.5	
8	67	3	18	2	22	6	35	3	33	5	9	+80
9	64	2	12	0	0	5	29	3	33	2	8	+300
13	89	5	29	0	0	9	53	3.5	39	5	12.5	+150
14	100	9	53	3	33	10	59	0	0	12	10	-17
16	55	7	41	3	33	10	59	5.5	61	10	15.5	+55
19	67					9	53	3	33		12	
20	89	1	6	2	22	6	35	2	22	3	8	+167
22	67	2	12	1	11	6	35	2	22	3	8	+167
23	61	6	35	2	22	11	65	8.5	94	8	19.5	+144
25	61	6	35	2	22	8	47	4	44	8	12	+63
27	78	7	41	2	22	9	53	6	66	9	15	+67
28	83	6	35	1	11	11	65	5.5	61	7	16.5	+136
29	72	6	35	2	22	6	35	4	44	8	10	+25
30	89	2	12	2	22	9	53	6	66	4	15	+160
33	83	2	12	0	0	7	41	1	11	2	8	+300
34	89	6	35	2	22	12	71	4	44	8	16	+100
35	89	5	29	2	22	6	35	2	22	7	8	+14
36	94	10	59	3	33	6	35	4	44	13	10	-23
37	75	2	12	0	0	8	47	4	44	2	12	+500
38	72					6	35	2	22		8	
40	72					9	53	6	66		15	
42	100	4	24	1	11	3	18	0	0	5	3	-40
43	67	2	12	1	11	5	29	3	33	3	8	+166
		5.0		1.5		7.8		3.6		6.5	11.3	+74
N=28		N=24		N=24		N=26		N=26		N=24	N=28	

Table 5 – All Students With 50% Or < Participation Rate

Student	Pretest Scores					Posttest Scores				Test Totals		
	PR	MC	%	OE	%	MC	%	OE	%	Pretest Total	Posttest Total	% Change
1	39	6	35	3	33	6	35	1.5	17	9	7.5	-17
3	5	1	6	0	0	4	24	2	22	1	6	+500
10	22	7	41	0	0					7		
11	17	6	35	4	44	9	53	4	44	10	13	+30
12	44	7	41	4	44	6	35	6	67	11	12	+9
15	44	3	18	0	0	6	35	0	0	3	6	+100
17	11	4	24	3	33					7		
18	5	7	41	2	22	8	47	7	78	9	15	+67
21	5	6	35	1	11	5	29	2	22	7	7	NC
24	28	5	29	0	0	4	24	1	11	5	5	NC
26	5	3	18	1	11	2	12	1	11	4	3	-25
31	6	3	18	2	22	9	53	4	44	5	13	+160
32	11	5	29	0	0					5		
39	33					2	12	2	22		4	
41	5	8	47	2	22					10		
44	6	3	18	0	0	4	24	1.5	17	3	5.5	+83
45	44	7	41	3	33	7	41	7	77	10	14	+40
		6.75	40	1.56	17	5.5	32	3.0	33	6.625	8.54	+29
N=17	N=17	N=16		N=16		N=12		N=12		N=16	N=13	

Both groups aggregate pretest scores were similar, 6.625 and 6.5, for the < 50% participation and > 50% participation groups, respectively. At a glance, only three students in the <50% group moved from NP to B. While in the >50% group, ten students moved from NP to B, and four student moved from NP to P. Student performance was improved 76% when attendance was >50%, versus 23% improvement in the <50% participation group. The fact that students improved even at lower attendance could be attributable to beneficial effects of the program and ongoing teacher input in the classroom.

Variables

The Posttest GEPA was given on the last two weeks of the program because of scheduling conflicts with students. During the last day of the program some students took the test, while the other students who had taken it the week before, had pizza and donuts. They were also chatting, so the atmosphere was not as conducive to high performance as I would have preferred.

Several of the students who came once and never returned, needed to be tracked down and took the test under varying conditions, such as being in a classroom while other instruction was going on. Again, this was not ideal testing conditions for optimal performance.

One could conclude definitively that the results obtained strongly support continuation of the GEPA Math after school program as a vehicle to improve student performance.

Chapter 5

Summary

This paper's purpose is to determine the effectiveness of an after school math program on raising student achievement on the Grade Eight Proficiency Assessment or GEPA. This norm referenced, standardized test is a requirement of the federal government and student progress on it is a factor in the school's rating based on the new No Child Left Behind legislation. This law requires that all students be proficient defined for eighth graders as a score greater than 200 by the year 2014.

Students in the study were selected based on a review of their fourth grade standardized test, the Elementary School Proficiency Assessment or ESPA now the New Jersey Assessment of Skills and Knowledge or NJASK, and the trend in their Terra Nova scores from 5th, 6th and 7th grades, if available. Based on a proficient score of 200, students who scored in the range of 185 to 210 were invited to attend the eighteen-week after school program. In addition, all special education students were invited to attend. Forty-five eighth grade students elected to participate in the program.

The program consisted of sample GEPA tests, review of math nomenclature and various test-taking strategies. The goal was to reduce student anxiety due to increased familiarity with the test format, review basic math concepts and vocabulary, and to suggest appropriate test taking strategies in an effort to boost their overall performance.

Students were given a pretest at the onset of the program and a posttest at the conclusion and results were tabulated and analyzed. Some students missed the beginning of the program and or the end when the tests were administered. The data collected is included but did not figure into the overall results. The results supported the continuation of the after school program for both regular and special education students. A significant increase in student scores provided evidence of the program's effectiveness. The one major variable, student participation rate was analyzed to help ascertain performance as a function of attendance. The data suggests that those students who attended the program greater than fifty percent of the time showed a marked improvement (+78%) over those students who attended less than fifty percent (+23%), although positive improvement was noted in both groups. A more detailed breakdown of participation rates is included in table 6 as a point of interest.

There are many school districts offering programs similar to the one in this study because all schools are under pressure to demonstrate continuous improvement. The hardest hurdle will be the special education population who by the very nature of their classification may be unable, despite exhaustive efforts by educators, to reach the goals established by No Child Left Behind. While this study showed an improvement trend in this select group, it should be noted that none of the special education students tested achieved proficiency as measured by the posttest. The intent of NCLB is admirable; however, the feasibility of achieving its lofty goals seems implausible. When the costs begin to outweigh the benefits one needs to readdress the priorities.

Within the Voorhees school district, we have altered our after school GEPA program which began four years ago. Initially, the program ran for the six weeks prior to the GEPA test. It

included two, one and half-hour sessions per week. The program now runs for eighteen weeks beginning in early October and running into the first week in March and includes once a week one-hour sessions. The actual number of hours instruction (18 hours total) did not change. The hope was that the material presented over a longer time would remain in students long-term memory rather than a cram session. The down side was the time commitment and the competition of other more engaging after school activities such as sports and theater. The program remains voluntary. The hope here was that only those students who were truly committed and interested in improving their scores would attend and behavior issues would be eliminated.

The program consisted of student participation in reviewing GEPA simulated test questions and familiarity with the GEPA format of multiple choice and open-ended questions.

Students could therefore miss a class here or there, and not lose a lot of ground. Some content, such as the Pythagorean Theorem, that had not as yet been introduced to the majority of eighth graders in their regular math classes prior to the GEPA testing, was reviewed as needed to solve questions.

Test taking strategies were emphasized throughout the program. Since GEPA testing does not penalize students for guessing students were instructed to bubble in multiple choice questions if time was running out, rather than leave blanks. A big improvement was noted in the posttest measurements for open ended questions. The strategy stressed here was to answer each bullet. The open ended questions were more heavily weighted and the score is based on a three point rubric. Students were instructed to draw a table or chart, and make some attempt at an answer, to at least get partial credit. Students did a nice job in following

this suggestion in the posttest and hopefully on the actual GEPA, which they just completed the week of March 7, 2005.

Some school districts have full day Saturday classes for several weeks prior to the GEPA. Others have various permutations on the programs mentioned above. It would be interesting to compare the results achieved under the varying conditions to see which is optimal. My concern as with any measurement of student knowledge is whether “teaching to the test” is an effective methodology for long term retention and applicability of mathematical concepts in real life situations. Unfortunately, the actual GEPA test results for eighth graders are received too late to be helpful in guiding instruction. The Terra Novas or other standardized tests given at other grade levels can and should be utilized to help guide both instruction and remediation efforts.

At the conclusion of the program, students were queried on their thoughts, feelings and suggestions for the GEPA program in the future. Student feedback was positive overall; however, a lot of students felt that 18 weeks were too long. A further breakdown of participation rate, point spread and percentage change (Table 6) tends to support the notion that cutting back the number of sessions to 10-12 would not dramatically impact results.

Table 6 – Student Progress by Participation Rate

Participation Rate	# of Students	Avg. Point Increase	% Increase
0-33%	8	+2.9	+53
34-50%	4	+1.6	+19
51-67%	10	+5.3	+85
68-100%	14	+3.7	+56

The above breakdown of both regular education and special education students by participation rate showed the greatest gains in point increase occurred in the 51-67% or 10-

12 hour category. The real test of the beneficial and predictive value of the program will come when the actual GEPA numbers come in and an analysis is performed on their performance under actual testing conditions. This will be completed over the summer and results submitted to our administration to justify the cost of the program and substantiate its effectiveness.

References

- Allen, D. (2004). Developing A New State Testing System. Retrieved September 27, 2004, from www.educating.com.
- Boser, U. (2000). Teaching To The Test? *Education Week*, 19, Issue 39.
- Bradley, A (2000). Put To The Test. *Education Week*, 20, Issue 5.
- Cookson, Peter W. Jr. & Halberstam, Joshua. (1998). A Parent's Guide to Standardized Tests on School-How To Improve Your Child's Chances of Success. *N.Y. N.Y. Learning Express*.
- Elliott, J., Erieken, R, Thurlow, M & Shriner, J. (2000). State- Level Accountability For The Performance Of Students With Disabilities: Five Years of Change? *Journal of Special Education*, 34, Issue 1.
- Gewertz, C. (2004). Urban Students Show Reading, Math Gains On State Assessments *Education Week*, 23, Issue 29.
- Goertz, M. & Duffy, M. (2003). Mapping the Landscape of High-Stakes Testing and Accountability Programs. *Theory Into Practice*, 42, Issue 1.
- Hanson ,C. (2001). The National Assessment of Educational Progress. *Eric/TM Digest*.
- Hi, A. (2002). The Generalizability of CBM Survey-Level Mathematics Assessments: Just How Many Samples do We Need? *School Psychology Review*. 31, Issue 4.
- Kane, M., Khattri, N. (1995). Assessment Reform. *Phi Delta Kappan*, 77, Issue 1.
- Lawton, M. (1996). Goals Panel To Survey State Assessments, Ask For More NAEP Data. *Education Week*, 15, Issue 21.
- Lewis, A. (2002). Variation Among State Standards. *Gifted Child Today Magazine*, 56, Issue 8.
- Miller, D. (2003) Measuring Up to the New Jersey Core curriculum Content Standards Teacher Edition
- Neill, M. (2000). Old Tests In New Clothes. *Instructor Vol. 109*, Issue 4
- Pandey, T. (1990). Authentic Mathematics Assessment. ERIC/TM Digest.
- Pascopella, A. (2004). Inside The Law. New England's Bold Test: Multi-State Assessments. Retrieved August, 2004, from www.DistrictAdministration.com

References (Contd.)

- Sheppard, L. (2002-03). The Hazards of High-Stakes Testing, *Issues in Science and Technology*, 19, Issue 2.
- Thurlow, M, Allison, L., Scott, D.& Ysseldyke, J. (2000). Students With Disabilities In Large-Scale Assessments: State Participation And Accommodation Policies. *Journal of Special Education*, 34, Issue 3.
- Tienken, C. & Wilson M. (2001). State Standards, Assessments and Instruction. *ERIC/TM Digest*
- Valenza, J.& Berkowitz, R. (2003). Improving Test Scores. *School Library Journal*, 49, issue 6.
- Yen, Wendy M. & Henderson, Diane L. (2002). Professional Standards Related To Using Large-Scale State Assessments in Decisions for Individual Students Measurement and Evaluation in Counseling and Development. *Measurement and Evaluation in Counseling and Development* 35, pp. 132-142.
- No child Left Behind Act Now One For The Books. (2002). *Education USA*, 44, No. 2.
- Interpreting Test Results. (2004). *New Jersey Professional Education Port*. Retrieved October 10, 2004, from www.njpep.org/tutorials/interpreting_results.
- New Jersey United for Higher School Standards Assessment Fact Sheet. (2004). Retrieved February 27, 2004, from www.newjerseyunited.org/educationa/assessments.asp.
- The Peoples Publishing Group, Inc. (2003) *Measuring Up to the New Jersey Core Curriculum Content Standards*.