EFFECTS OF MATH BENCHMARKING ON STUDENTS AT THE ELEMENTARY LEVEL

By

JOSEPHINE K MCCLERNON GOEKE

Submitted to

Educational Leadership Faculty

Northwest Missouri State University Missouri

Department of Educational Leadership

College of Education and Human Services

Maryville, MO 64468

Submitted in Fulfillment for the Requirements for
61-683 Research Paper

Summer 2012

June 21, 2013
ABSTRACT

The purpose of this study was to determine whether or not there is a significant difference in elementary students’ math performance, as measured by the Scantron Performance Series Diagnostic Assessment, if they receive instruction in math benchmarking. It is considered best practice to provide students with clear, meaningful learning objectives, assess those objectives with formative assessments conducted throughout the learning process, and provide re-teaching or adjusted instruction, as appropriate, depending upon results of the ongoing formative assessments. It is also considered best practice to include students in the process of setting learning goals and monitoring progress. These are the components of Dennis Chaconas’ math benchmarking system. The findings of this study indicate that, at the 3rd and 4th grade levels, there is no difference in math performance between groups of students who receive a semester of traditional instruction in math (without benchmarking), and those who received a semester of math instruction with the benchmarking system. At both grade levels, students who received benchmarking instruction had a slightly higher mean score on the Scantron Performance Series Diagnostic Assessment, however there was not a significant difference in scores between those who received traditional math instruction and those who received math benchmarking instruction. Furthermore, a variety of other variables could have contributed to the slight differences in mean scores between the two groups. Based upon these findings and the review of research and literature available on the topic, it would seem that the findings are inconclusive, and further analysis should be done in another four to five years to more accurately determine whether the benchmarking system is an effective method for improving student performance in math.
INTRODUCTION

Background, issues and concerns.

The infamous No Child Left Behind Act (NCLB), signed into law in 2002, brought a new level of expectation for student performance in the United States. The law required 100% of American students at the elementary and secondary levels to perform at a minimum of a proficient level in core academic areas by the year 2014. While President Barack Obama has waived the requirement to meet this deadline for at least 35 states (McNeil, 2013), there is still the expectation that districts are achieving continued academic (House, 2013). However, the reality is that many students continue to lag behind, especially in districts with high populations of students living in poverty (N.A., 2011). Many of these students are faced with multiple obstacles outside of the school setting, making academic success much more difficult. Hunger, violence, homelessness, and lack of adequate adult interaction outside of school are just a few of the issues that can interfere with student learning (Carter, 2013). Although these factors are often beyond the control of the school system, teachers are still held responsible for getting all students to reach high levels of academic success.

These requirements lead districts to be in constant search for a solution to the pervasive academic inequities existing among various groups of students. Some districts are hopeful that the answer to poor student performance in mathematics lies in the implementation of math benchmarking.

Practice under investigation.

The practice under investigation will be the effectiveness of Dennis Chaconas’ math benchmarking system in achieving improved academic performance in math for students at the elementary level. There will be an investigation to determine whether or not math benchmarking
systems have a positive effect on elementary student performance in math. Student test scores for mathematics from the Scantron Performance Series Diagnostic Assessment will be analyzed to determine whether or not groups of students receiving math instruction with the benchmarking system earn significantly higher Performance Series scores than students who receive traditional math instruction without benchmarking.

*School policy to be informed by study.*

Elementary schools are charged with the task of continually increasing the percentage of students who are performing at proficient and advanced levels each year on the state assessment. These increases must be significant enough to meet criteria known as Adequate Yearly Progress (AYP), a goal that is determined by the state’s Department of Elementary and Secondary Education. Results from the 2012 MAP test indicate that, at a Midwestern Elementary School, 57.7% of third graders and 41.3% of fourth graders performed at proficient or advanced levels in Mathematics. The NCLB Act requires 100% of students to be performing at a minimum of a proficient level by 2014. While Missouri has been granted an NCLB waiver, Missouri schools are still required to make continuous improvement in students’ academic performance. If math benchmarking can lead to a greatly accelerated rate of improvement, “Midwestern Elementary” would have a better chance at meeting this goal.

*Conceptual underpinning.*

Master teachers recognize the importance of identifying essential skills and knowledge students must master, providing students with clear objectives, administering frequent formative assessments that reliably assess the desired student outcomes, responding with immediate feedback, and adjusting instruction or providing “re-teaching” when the content has not been mastered. Moreover, they understand the power associated with intrinsic motivation, in which
students set their own goals and monitor their own progress. Finally, effective teachers know how crucial it is to celebrate student successes when goals are reached.

The reality, however, is that teachers often find themselves “covering” material in a textbook in order to keep up with the pacing guide, rather than teaching a concept until they are certain that actual learning has occurred. Tests tend to be given for the purpose of a grade, rather than for the purpose of informing instruction. In the subject of math, in which most skills and knowledge that are taught are prerequisites for upcoming skills, students who didn’t master the material are left behind, with little hope of catching up.

The concept of benchmarking in math is designed to be closely aligned with the former scenario, in which both teachers and students are keenly aware of student progress, and remediation is provided immediately in an effort to prevent students from falling behind. It is also designed to be motivating for students, and provide students and parents alike with straightforward information about which skills a student has mastered, and those they have not. The expectation is that all students will ultimately master all objectives within the math curriculum. One can easily conclude, then, that the implementation of math benchmarking will ultimately result in improved student performance in math.

Statement of the problem.

In order to determine whether or not to continue implementing math benchmarking, teachers and administrators need to know whether or not it is making a difference in elementary student performance in math.
Purpose of the study.

To determine whether there is a significant difference in math performance between groups of elementary students who receive traditional instruction in math (without benchmarking) and groups of students who receive instruction in math using the benchmarking system.

Research questions.

RQ#1: Is there a difference in Scantron Performance Series Diagnostic Assessment scores in Mathematics between groups of elementary students who receive traditional instruction in math and those who receive math instruction using benchmarking?

Null hypothesis.

There is no difference in Performance Series scores in Mathematics between groups of elementary students who receive traditional instruction in math and those who receive math instruction using benchmarking.

Anticipated benefits of the study.

If elementary students who receive math instruction using the benchmarking system earn significantly higher Performance Series scores in Mathematics than those who receive traditional math instruction, teachers should continue to implement math benchmarking at the elementary level.

Definition of terms.

Scantron Performance Series Diagnostic Assessment (Performance Series)- diagnostic assessment used to determine current levels of student performance in math and reading, typically administered to students in grades two through five in the Fall, Winter, and Spring.
No Child Left Behind Act (NCLB) – Legislation signed into law by former President George W. Bush in 2001. The law requires 100% of students to achieve at proficient or advanced levels on state assessments.

NCLB Waiver – waivers granting flexibility regarding certain requirements of the NCLB Act

Adequate Yearly Progress (AYP) – a goal for yearly improvement in state assessment scores, determined by individual states, designed to result in ultimate proficiency for all students

Missouri Assessment Program (MAP) – Missouri’s state assessment, administered in the spring semester of each school year at grades 3 – 8.

Formative Assessments – assessments used during the learning process to inform students and teachers of students’ current understanding of target skills and concepts in order to better inform instruction

Summary.

A study was conducted to determine whether there is a significant difference in math scores on the Performance Series Diagnostic Assessment between groups of students who received traditional instruction in math (without benchmarking) and those who received math instruction with the benchmarking system. A t-test was administered. If the results of the test indicate that students receiving instruction in math with benchmarking perform at significantly higher levels than students who receive traditional math instruction without benchmarking, teachers should continue to implement the benchmarking system. Ongoing use of formative assessments and subsequent adjustments to instruction based on analysis of results are widely-accepted best practices, which should result in improved student achievement. Math benchmarking is strategically designed to facilitate this process. This study will contribute to an
ongoing conversation about the effectiveness of the benchmarking system and will allow districts to better determine whether or not the system would benefit their students.
REVIEW OF LITERATURE

The U.S. Department of Education (ED) recommends that schools utilize research-based instructional “best practices”. To communicate those practices to educators, ED houses a collection of information about best practices on its “Doing What Works” (DWW) website. The department also provides a link on its website to the Institute of Education Sciences’ site: “What Works Clearinghouse” (WWC). Both DWW and WWC provide reports that promote the extensive and ongoing collection and analysis of student achievement data at student, classroom, school, and district levels in an effort to pinpoint areas in which student performance can be improved. Both include recommendations that teachers work collaboratively to determine students’ academic needs and adjust instruction accordingly. Finally, both sites recommend that students participate in the process of setting goals and monitoring their progress. Clearly, the use of ongoing assessment is valued and encouraged as best practice in the field of education.

The concept of frequently assessing student performance and responding appropriately with adjusted or improved instruction is hardly a new concept. In a 1947 article from Educational Leadership, author Paul R. Grim (1947), who describes evaluation as “the process of gathering and interpreting evidence regarding the progress and problems of the learner in achieving desirable educational objectives,” says that “evaluation must operate as an integral part of the learning process, not in isolation from it, nor as an end in itself. It is, rather, a continuous, cumulative process, an aspect of functional learning, serving to diagnose learning difficulties and to aid in guiding desirable child growth.” (p.438-439). A 1996 publication by Dylan Williams reiterates the importance of this practice, referring to an assessment cycle in which “evidence of performance or attainment is elicited, interpreted and acted upon, in some way” (William & Black, 1996). Williams further illustrates the differences between summative
assessments, “which are designed to judge the extent of students’ learning of the material in a course, for the purpose of grading, certification, evaluation of progress or even for researching the effectiveness of a curriculum,” and formative assessments, which are characterized by an element of feedback given to students in an effort to “alter the gap” that exists between their level of understanding and the desired level of understanding of a skill or concept.

During the Cold War, student assessment data began to transcend the field of education when the federal government started to take more of an interest in the status of academic achievement among American students (Alderson & Martin, 2006). With this shift came an effort to better quantify the learning that occurs among students. G.E. Dart described such efforts in his 1971 article for Educational Leadership, in which he discusses his school district’s efforts to establish “verifiable performance objectives”. He references a document compiled by a district that had been piloting the strategy, describing the document as “an inch-and-a-half thick” and containing “literally hundreds of particularized objectives – all the reading skills, every item of knowledge and performance in the primary, intermediate, and upper mathematics program, etc.” (Dart, 1971, p. 728). The process described by Dart was likely influenced by Benjamin Bloom, who in the 1960s and 1970s was conducting his own research on how learners acquire new knowledge and skills, and what role teachers play in that process. Bloom’s research ultimately led him to conclude that teachers had been missing the mark by organizing their teaching in such a way that resulted in testing students at the end of a unit with summative evaluations. Conversely, he resolved that “these checks on learning progress…would be much more valuable if they were used as part of the teaching and learning process to provide feedback on students’ individual learning difficulties and then to prescribe specific remediation activities” (Guskey, 2010, p. 53). Bloom was referring to the importance of utilizing formative assessments. He
subsequently developed a strategy, which he called *mastery learning*, in which “teachers organize the important concepts and skills they want students to acquire into learning units, each requiring about a week or two of instructional time. Following high-quality initial instruction, teachers administer a *formative assessment* that identifies precisely what students have learned well and where they still need additional work. The formative assessment includes explicit, targeted suggestions – termed *correctives* – about what students must do to correct their learning difficulties and to master the desired learning outcomes.” (Guskey, 2010, p. 53).

Although the concept of such formative assessments has been recognized among educators for years as an example of best practice, the notion of frequent testing is often resisted, perhaps as a consequence of the federal government’s interest in the results of such assessments. As Dart writes in his 1971 article, “we all have the feeling that verifiable performance objectives, district and state educational aims and goals, and all of planning, programming, budgeting are really so that someone outside the children and the teachers can use them to judge the children and the teachers.” (p. 728) He contends that the measure of a child’s success cannot be fully recognized through a collection of test scores, and he expresses his hope that schools would not be judged by test scores alone (Dart, 1971). Of course, this very fear was realized in 2001 with the passage of the No Child Left Behind Act, which charged educators with the task of getting 100% of American students to perform at a minimum of a proficient level by the year 2014, thereby using results of student performance on state tests “to judge schools” on their ability to educate students. (Mika, 2005). In 2006, Ken Petress wrote that “the amount of time used to prepare for…testing and the testing itself occupies an alarming proportion of class time for teachers and students” and “is counterproductive due to its frequency, content and question style, and stress it places on students and teachers.” (Petress, 2006, p, 80). Yet, Petress also admits that
“testing is valuable and needs to be supported and used.” (p. 81). Since the law requires 100% proficiency, teachers must continue to find a method that can help them to achieve this goal.

Through the 1970s, mastery learning was a widely accepted method for organizing the cycle of teaching and assessing students in the classroom setting. Then, in 1980, a man named William Spady, together with a cohort of about 40 other individuals, developed the concept of Outcome-Based Education, or OBE. In a 1998 interview, Spady recalled that the group believed that, through this system, they “could get higher learning- success for far more students than we had previously thought possible” (Tucker, 1998). In the 1998 interview, Spady explained that “the primary motivation, initially, was simply generating higher test scores. That has always been the principal measure of success to schools.” In 1984, Spady had co-authored an article in an issue of Educational Leadership, in which he outlined the successes that had been achieved by a school in Connecticut that had adopted OBE practices. One such indicator of OBE’s success in the school was that “no more than one or two students per year in the entire school failed to reach grade level on standardized mathematics tests (including mildly handicapped students)” (Rubin & Spady, 1984, p. 38). However, in 1998, when reflecting upon the beginning years of OBE, Spady notes that “it didn’t matter to most educators whether or not the knowledge reflected by the tests was useful…We called what we were doing ‘outcome-based,’ but the outcomes were anything that anybody happened to be doing that resulted in a higher score” (Tucker, 1998). Consequently, Spady and others involved in the OBE movement worked to make adjustments in fine-tuning the final objectives students would work toward to be sure that they reflected meaningful, useful learning. Instruction and assessments leading up to the ultimate goal were designed specifically to aid students in acquiring desired skills and knowledge.
Although OBE has gained momentum over the years, its existence has not gone without criticism. Even those who have been in support of the system recognize its flaws, including the level of difficulty in implementing something so complex. In 1991, Jean A. King and Karen M. Evans wrote about the challenges associated with OBE, including the fact that “a massive curriculum development effort is required to, in Spady’s words, ‘design down’ from exit outcomes to specific lesson outcomes for every student” (King & Evans, 1991, p. 74). This portion of the work is what Dennis Chaconas is doing with his math benchmarking systems. Inspired by the work of Benjamin Bloom, William Spady, and Dylan Williams, Chaconas designed his own instruction, based in part on a combination of their ideas when he was a teacher in the late 1970s and early 1980s. Chaconas’ benchmarking system was most greatly influenced, however, by OBE. After spending some time in 2003 and 2004 as a consultant for UCLA, Chaconas branched off to become an independent consultant, and now works with school districts in seven states to implement math benchmarking (Chaconas, 2013). Chaconas describes benchmarking as “an effective way to operationalize standards-based math instruction and make real the promise of higher expectations and college readiness for all of our students” (Chaconas, 2012, p. 1). He says that “the work must be done through a collaborative effort of teachers, instructional coaches, administrators, and external support” to “craft and implement benchmarks” (Chaconas, 2012, p. 1). Essentially, Chaconas leads these individuals through a collaborative effort to identify the minutest of math learning objectives (reminiscent of the verifiable learning objectives of the early 1970s) from a district’s standards. In most districts, these objectives are derived from the Common Core State Standards. Chaconas next guides the group in developing a series of formative assessments to be used in collecting ongoing information about which objectives students have met, and which need to be retaught until they
are mastered. Chaconas outlines several “key components” of benchmarking in the document he recently distributed to a group of educators who will be embarking upon this process during the summer of 2013. The components include “clear benchmarks”, written as “I Can” statements which communicate the desired learning outcome. Students’ progress toward meeting benchmarks is to be “public”, meaning that a chart is posted in the classroom, displaying which benchmarks each student has met, thus putting the “focus of the class on learning” and providing “public accountability for students and for teachers”. Assessments and materials are to be “aligned”, “effective teaching strategies” are to be used to help students attain benchmark goals, and “additional learning time” is to be provided for students when they do not pass a benchmark. The grading system used in math benchmarking is standards-based. “An A, B, C, I system of grading is used to ensure that students are expected to attain a minimum level of proficiency.” (p.1-2).

Another component of the system is the “student individual accountability” in which students maintain their own log of benchmarks to keep track of which benchmarks they have passed and set personal goals for how they will progress until all benchmarks are met. “Recognition of success” is another component of Chaconas’ benchmarking system. Under the system, students should be recognized or even rewarded for positive academic progress. Finally, Chaconas places a strong emphasis on the requirement that all those involved in the benchmarking system have an adequate knowledge base about how benchmarking works. His system requires teachers, instructional coaches, principals, and district leaders to have “a level of fluency in the benchmarking process” (Chaconas, 2012, p. 3). According to a study conducted in Philadelphia, this wide knowledge base could make or break the effectiveness of the system. The study of a separate Benchmark system being used in Philadelphia found that “the most
robust predictors of learning growth were instructional leadership, collective responsibility, and use of the Core Curriculum” (Blanc, Christman, Liu, Mitchell, & Travers, 2010, p. 206).

Chaconas’ system has been recognized for the impact it has had on students at the high school level. A February, 2013, article in the Kansas City Star features two Kansas City-area school districts that have been using math benchmarking. “Kansas City Area School District #1” began implementing the system during the 2009-2010 school year, and “Kansas City Area School District #2” adopted the system soon after. According to the article in the Kansas City Star, in School District #1, “the percentage of…students scoring proficient or advanced in Algebra I jumped from 27.8 percent to 61.2 percent” from 2009 to 2012, and in School District #2, this percentage rose “from 28.5 percent to 51.6 percent” within the same time frame (Robertson, 2013).

The reality for current educators in the public education system is that they must work to prepare their students for high stakes summative evaluations, such as the MAP test, in order to provide American citizens with information about the state of academic performance among American students. Certainly, most teachers recognize the need for administering ongoing formative assessments throughout the learning process in order to guide students toward ultimate learning outcomes. Dennis Chaconas’ relatively new math benchmarking system is gaining attention for early successes. Continued data and research over the coming years will reveal whether or not this system can yield consistent positive results.
RESEARCH METHODS

Research design.

The research design consisted of a quantitative study to determine whether there was a significant difference in math scores on the Performance Series Diagnostic Assessment between groups of students who receive traditional instruction in math (instruction without the use of benchmarking) and those who received math instruction with the use of benchmarking. The independent variable was instruction in math benchmarking. The dependent variable was student test scores in math on the Performance Series Diagnostic Assessment. If students receiving instruction in math benchmarking score significantly higher on the Performance Series assessment than those who received traditional instruction, the benchmarking system should continue to be implemented.

Study group description.

Scores were obtained from sixty-one third graders who had received a semester of traditional math instruction (Without Benchmarking), sixty-four third graders who had received a semester of instruction in math benchmarking (With Benchmarking), sixty-eight fourth graders who had received a semester of traditional math instruction (Without Benchmarking), and seventy-six fourth graders who had received a semester of instruction in math benchmarking (With Benchmarking). The cohort groups were reasonably comparable in terms of previous math instruction and socioeconomic status.

Data collection and instrumentation.

Data was obtained from the instructional coach at “Midwestern Elementary School,” which included individual Standard Scores in math for third and fourth grade students who took the Performance Series Diagnostic Assessment in the winter of the 2011-2012 school year, prior
to implementation of math benchmarking, and third and fourth grade students who took the same test in the winter of the 2012-2013 school year, after having received one semester of math benchmarking.

Statistical analysis methods.

A t-test was conducted for each grade level to determine whether or not there is a significant difference in math performance on the Scantron Performance Series Diagnostic Assessment between groups of students who received traditional instruction in math (without benchmarking), and those who received instruction in math using the benchmarking system. The source of data for third graders was split into two categories: those who had received one semester of traditional math instruction (without benchmarking) and those who had received one semester of instruction in math benchmarking. The source of data for fourth graders was split into the same two categories. The test revealed the mean, mean difference, t-test score, degrees of freedom, and p-value. The null hypothesis states that “there is no difference in Performance Series scores in Mathematics between groups of elementary students who receive math benchmarking instruction and those who do not.” The alpha level was set at 0.10 for the purposes of testing the null hypothesis with a 90% confidence interval.
FINDINGS

A t-test was conducted to determine whether there was a difference in student performance on the Scantron Performance Series Diagnostic Assessment in mathematics between groups of students who received traditional instruction in math (without benchmarking), and those who received instruction in math benchmarking. Below are the results, depicted in tables and followed by explanations of the results. The data used was provided by the Instructional Coach at “Midwestern Elementary School.” It represents Standard Scores of third and fourth graders at “Midwestern Elementary” during Winter 2011, when students were receiving traditional math instruction (without benchmarking); and Winter 2012, after students had received one semester of instruction in math benchmarking.

Figure 1

**t-Test Analysis Results for 3rd Graders: Without Benchmarking versus With Benchmarking**

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean</th>
<th>Mean D</th>
<th>t-test</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without Benchmarking (n=61)</td>
<td>2272.00</td>
<td>-17.53</td>
<td>-0.62</td>
<td>123</td>
<td>0.53</td>
</tr>
<tr>
<td>With Benchmarking (n=64)</td>
<td>2289.53</td>
<td>-17.53</td>
<td>-0.62</td>
<td>123</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Note: Significant when p<=0.10

Scores were obtained reflecting the performance of sixty-one third graders in the “Without Benchmarking” group, and sixty-four third graders in the “With Benchmarking” group. Performance Series nationally normed data indicates the interquartile range (range of student scores between the 25th and 75th percentiles) for 3rd graders taking this test to be 2153 – 2369. The mean score of the “Without Benchmarking” group was 2272.00. Scores for this group ranged from 1869 to 2516. The mean score of the “With Benchmarking” group was 2289.53.
Scores from this group ranged from 1978 to 2820. The mean difference was -17.53. The t-test score was -0.62. The degrees of freedom was 123. The p-value was 0.53. The alpha level was set at 0.10.

The purpose of the analysis was to determine whether there was a difference in scores between groups of third graders who had received traditional instruction in math (without benchmarking) and those who had received math instruction with benchmarking. The analysis shows that, although students who received benchmarking instruction had a higher mean score, there was not a significant difference, based on the comparison of the p-Value, 0.53, which was greater than the alpha level, 0.10. Therefore, for third graders, the researcher failed to reject the null hypothesis. There is no difference in Performance Series scores in Mathematics between groups of third grade students who receive math benchmarking instruction and those who receive traditional instruction in math. The findings suggest that one semester of math instruction using the benchmarking system does not make a difference in math performance for third graders.

Figure 2

**t-Test Analysis Results for 4th Graders: Without Benchmarking versus With Benchmarking**

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean</th>
<th>Mean D</th>
<th>t-test</th>
<th>Df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without Benchmarking (n=68)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=68)</td>
<td>2347.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With Benchmarking (n=76)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2370.93</td>
<td>-23.01</td>
<td>-1.27</td>
<td>142</td>
<td>0.21</td>
<td></td>
</tr>
</tbody>
</table>

Note: Significant when p<=0.10

Scores were obtained reflecting the performance of sixty-eight fourth graders in the “Without Benchmarking” group, and seventy-six fourth graders in the “With Benchmarking” group. Performance Series nationally normed data indicates the interquartile range (range of
student scores between the 25th and 75th percentiles) for 4th graders taking this test to be 2263 - 2476. The mean score of the Without Benchmarking group was 2347.93. Scores for this group ranged from 2061 to 2572. The mean score of the With Benchmarking group was 2370.93. Scores for this group ranged from 2112 to 2585. The mean difference was -23.01. The t-test score was -1.27. The degrees of freedom was 142. The p-value was 0.21. The alpha level was set at 0.10.

The purpose of the analysis was to determine whether there was a difference in scores between groups of fourth graders who received traditional instruction in math (without benchmarking) and those who received math instruction with the benchmarking system. The analysis shows that, although fourth graders who had received instruction in math benchmarking obtained a slightly higher mean score than fourth graders who had not received traditional math instruction, there was not a significant difference. The p-Value, 0.21, was greater than the alpha level, 0.10. Therefore, for fourth graders, the researcher failed to reject the hypothesis. There is no difference in Performance Series scores in Mathematics between groups of fourth grade students who receive math benchmarking instruction and those who receive traditional instruction in math. The findings suggest that one semester of math benchmarking instruction does not result in a significant difference in math performance for fourth graders.
CONCLUSIONS AND RECOMMENDATIONS

Educators understand the importance of administering ongoing formative assessments, collaboratively analyzing results, and adjusting instruction accordingly. They further understand the power associated with communicating clear, meaningful objectives, and involving students in setting personal goals and monitoring progress toward those goals. The problem is that a system like this can be difficult to administer. The math benchmarking system, developed by teachers with the guidance of consultant Dennis Chaconas, if designed to accommodate this cycle of instruction and assessment. One would assume that a strategic system such as this one would be likely to yield improved student performance.

The results of this study would suggest that there is no significant difference between scores of 3rd and 4th grade students who receive one semester of traditional math instruction and those who receive one semester of math instruction using the benchmarking system. At the 3rd grade level, although the “With Benchmarking” group had a slightly higher mean score, there was not a significant difference between their scores and the scores of students who received traditional instruction in math, based on the comparison of the p-Value, 0.53, which was greater than the alpha level, 0.10. At the 4th grade level, although the “With Benchmarking” group had a slightly higher mean score, there was not a significant difference between their scores and the scores of students who received traditional instruction in math, based on the comparison of the p-Value, 0.21, which was greater than the alpha level, 0.10. When considering these results, it is important to recognize, however, that the data available for this study only reflected the impact of one semester of math benchmarking instruction. Also, all of the teachers who taught benchmarking to the “With Benchmarking” groups were using the benchmarking system for the first time, so a lack of experience with the system could have impacted its effectiveness.
Another factor to consider is that, at both grade levels, only two of the three teachers providing math instruction remained constant over the two years. It is, therefore, also possible that results could have been affected by the fact that each grade level had one new teacher during the first semester of benchmarking instruction. Finally, different classes of students comprised the “Without Benchmarking” students and the “With Benchmarking” students at each grade level. It is a possibility that there are simply differences in achievement levels between the different groups of students, independent of math benchmarking.

Initial results indicate that math benchmarking does not significantly improve student performance in math. However, due to the limited data available for this study, the researcher considers the results to be inconclusive. It is recommended that student performance data is gathered over the course of four to five more years of math benchmarking instruction, and analyzed again to determine, with more certainty, whether or not math benchmarking yields a higher level of effectiveness in improving student performance in math after a longer implementation period. It could also be useful to use assessment data only for teachers that remain constant, so as to minimize the chance that differences in teacher effectiveness would have an impact on results.

Another consideration for educators would be whether or not a similar system would be effective in other subject areas at the elementary level, such as Reading or Writing. It would be desirable to develop a system of benchmarks for one or both of these subject areas, administer the instruction to students, and conduct a similar study to this one. If future analyses indicate that a benchmarking system can improve student performance in math, perhaps a system such as this one is the answer for low-performing students across content areas.
As teachers continue to implement math benchmarking, it is essential that they continue to receive learning opportunities on a variety of related topics in order to ensure the most successful possible results. Teachers should be very knowledgeable about the Common Core State Standards, and the “I Can” statements which have been generated for each math benchmark. They should have a clear understanding of what each standard requires students to be able to do, so that they can guide students toward that ultimate learning outcome. Teachers should also be equipped with a wide variety of effective instructional strategies for teaching math concepts to students in a way that appeals to a variety of learners. When students do not pass a benchmark, teachers must be ready to reintroduce the material in yet another way until they find the method of learning that will make sense to those students. In order to communicate with parents about students’ math instruction and grades, teachers must have a clear understanding of what benchmarking is and why it is being used. They must be able to explain to parents how the benchmarking grading system works and why it is being used. Finally, because student involvement is such an important component of math benchmarking, teachers would benefit from professional development to learn more about how to conference effectively with individual students.
REFERENCES


x1aWQmc2l0ZT1laG9zdC1saXZl#db=ehh&AN=18647465


http://ezproxy.nwmissouri.edu:2065/ehost/detail?vid=10&sid=f8a710d9-210e-4823-97cc-b52af1929151%40sessionmgr114&hid=122&bdata=JkF1dGhUeXIbPNvb2tpZSxpcC
x1aWQmc2l0ZT1laG9zdC1saXZl#db=ehh&AN=18620228

http://ezproxy.nwmissouri.edu:2065/ehost/detail?vid=16&sid=f8a710d9-210e-4823-97cc-b52af1929151%40sessionmgr114&hid=122&bdata=JkF1dGhUeXIbPNvb2tpZSxpcC
x1aWQmc2l0ZT1laG9zdC1saXZl#db=eric&AN=EJ913779


http://ezproxy.nwmissouri.edu:2065/ehost/detail?vid=8&sid=f8a710d9-210e-4823-97cc-b52af1929151%40sessionmgr114&hid=122&bdata=JkF1dGhUeXBlPWNvb2tpZSxpcCxlAQmc2l0ZT1laG9zdC1saXZl#db=eric&AN=EJ432790

Retrieved June 5, 2013 from
http://www.edweek.org/ew/section/infographics/nclbwaivers.html

http://www.gallup.com/poll/19552/testing-testing-schools-too-much.aspx

http://www.edweek.org/ew/issues/achievement-gap/

http://ezproxy.nwmissouri.edu:2065/ehost/detail?vid=12&sid=f8a710d9-210e-4823-97cc-b52af1929151%40sessionmgr114&hid=122&bdata=JkF1dGhUeXBlPWNvb2tpZSxpcCxlAQmc2l0ZT1laG9zdC1saXZl#db=eric&AN=EJ732737

