INTERVENTION THAT ADDS UP: THE IMPACT OF MERIT SOFTWARE ON STANDARDIZED ACHIEVEMENT TEST SCORES OF MIDDLE SCHOOL STUDENTS

By

Dr. SAMUEL SECURRO, JR.,
Professor of Education
Marshall University Graduate College.
West Virginia

Dr. JERRY D. JONES,
Professor of Education
Marshall University
Graduate College
West Virginia

Mr. DANNY CANTRELL,
Project Evaluation Coordinator
West Virginia State University Institute,
West Virginia

Mr. JAMES "MICKEY"
Principal
Horace Mann Middle School
West Virginia

ABSTRACT

The "No Child Left Behind" Act mandated the need for research-based interventions to increase and to improve learning and achievement for all youngsters. Research in computer-based instruction and intervention for learning basic skills and related achievements in content area subjects has documented the need for controlled investigations of such software and how it may improve the learning and performance of youngsters, and particularly for those who are in the "lower quartile" of school achievement. Although the current study focused on the effects of Merit Mathematics software on the achievement of middle school youngsters, effects of the treatment is also included for social studies, science, and reading/LA as measured by the state-mandated testing program in West Virginia (WESTEST).

A pre to post analysis was performed using a t-test for dependent samples to measure the overall differences in WESTEST mean scores from pre to post conditions for each of the four content areas, and results were statistically significant for all four WESTEST mean score pairs (p .000, SPSS Version 13.0). Effect size measures revealed the following magnitude of change: Mathematics (.844); Reading/LA (.223), Science (.132), and Social Studies (.166). The effect size of .844 for Mathematics is an extremely large value, indicating a very substantial difference (increase) in these scores from pre to post.

Two socioeconomic factors (ethnicity and eligibility for free lunch) were incorporated into the study to determine if these factors affected the outcomes. Inspection of independent t test results were insignificant, indicating that ethnicity and free lunch were not major factors in the overall outcomes.

INTRODUCTION

Merit Software, a publisher of software since the 1980's, has commissioned ongoing studies on the impact of Merit software on student achievement. This study is a follow-up and extension of an earlier study conducted at Calhoun County Middle School in rural West Virginia. The first study demonstrated that Merit Math software improved the achievement of students in grades six and eight on all nine of the SAT-9 variables. The current study investigated whether Merit software has improved the achievement of middle school students in an urban setting in southern West Virginia in science, mathematics, social studies and English/La.

Administrators and teachers at Horace Mann Middle School in Charleston, West Virginia agreed to participate in a treatment versus comparison group study. No Child Left Behind Legislation spotlighted the need to improve the achievement of struggling students and these administrators and teachers wanted to investigate the impact of effective, research-based programs and interventions on the learning of all students. The research team coordinated the project with school administrators, trained the participating teachers in the use of the software, and collected and analyzed the data.

Review of the Literature

A systematic review of published literature was undertaken to assess the effectiveness of educational software, including a few large scale reviews of multiple software programs.

Schacter (1999), writing for the Milken Exchange on education technology, analyzed five large scale studies of education technology that had been done up to 1999 to summarize the impact of educational technology on learning. He reported the findings of Kulik's (1994) meta-analysis study of 500 individual research studies of computer-based instruction which showed that:

1) students who used CAI scored, on average, at the 64th percentile on achievement tests compared to students in control groups without computers, who scored at the 50th percentile; 2) students receiving CAI learn more in less time; and 3) students have more positive attitudes toward their classes and like them more when their teachers include CAI.

Schacter, cites Mann's 1999 study of West Virginia's Basic Skills/Computer Education (BS/CE) program, which analyzed a representative sample of 950 fifth-grade student's achievement from 18 of the state's elementary schools, He found the following: 1) a rise in student test scores on the Stanford 9 corresponded with increased levels of participation in BS/CE; 2) all students scored high on the Stanford 9 because of BS/CE, with the greatest increase in scores among the *lower achieving students*; 3) 50 percent of the teachers reported in the sample, that technology had helped considerably with the state's instructional goals and objectives, and they became more enthusiastic about BS/CE with the passage of time; and 4) boys and girls did not differ in regard to achievement, access, or computer use in the study. CAI has demonstrated a positive impact on mathematics achievement in a national study. Wenglinsky's 1998 National Study of Technology's Impact on Mathematics Achievement (summarized in Schacter), assessed the effects of simulation and higher order thinking technologies on a national sample of 6,227 fourth graders' and 7,146 eighth graders' mathematics achievement on the National Assessment of Educational Progress (NAEP). Wenglinsky found that eighth graders who used CAI displayed gains in math scores of up to 15 weeks above grade level as measured by NAEP. Also, when the teachers of eighth-grade students received professional development on computers, their students showed gains in math scores of up to 13 weeks above grade level. Higher order uses of CAI and professional development positively correlated to student academic achievement in mathematics for both fourth- and eighth-grade students.

The Wenglinsky study is important because it used particularly a national database, (NAEP, 1996) and advanced analysis techniques to isolate the effects of the computer from the myriad other factors involved in student achievement" (Barton, 1998). For eighth graders, the study found that "the frequency of home computer use was positively related to academic achievement and the social environment of the school, [and] the use of computers to teach lower-order thinking skills was negatively related to academic achievement and the social environment of the school" (3).

When the relationship between technology use and achievement is measured in terms of estimated grade levels, the estimates suggest "substantial positive benefits of technology for eighth-graders, but mixed results for fourth-graders" (30). Wenglinsky emphasized three implications of these findings: 1) state and federal policymakers should take every effort to insure that teachers are properly trained to use computers; 2) teachers should focus on using computers to apply higher-order skills learned elsewhere in class, and 3) the primary focus of all technology initiatives should be on middle schools rather than elementary schools because most higher-order concepts are not introduced before middle school.

Brown(2000) found that lower achieving Black students who used the CAI made greater mathematics progress than did those in the control group who did not use the program. His research provides strong evidence for the use of CAI as a supplement to classroom instruction of

mathematics in elementary students and in middle school algebra students. Specifically, he found that Black students who ranked below their White classmates in mathematics achievement levels gained the most from the CAI software. He believed these differences may be attributed to the fact that White students began the program with higher achievement levels than Black students, leaving less potential for growth among White students and greater potential for growth among Black students.

Some researchers have found that CAI has raised achievement scores for lower achieving students. Christmann and Badgett (1999) compared science students who were taught with traditional instructional methods to those who received traditional instruction supplemented with CAI. Their results revealed that students receiving the technology supplemented instruction had higher academic achievement. The authors reported that CAI has been more effective in raising achievement of lower ability students. Christmann further noted that researchers Atkinson (1969) and Watson (1972) pointed out that "the computer can supplement the drill-and-practice of traditional instruction through relevant practice exercises," and that the "tutorial mode of the computer presents the student with an introduction of concepts that is followed by appropriate questioning strategies."

According to the criteria defined by Bindig (2002), Merit Software modules fits to a tutorial category rather than Drill and Practice or Application. The distinction is made primarily due to the involvement of higher levels of cognitive thinking. In addition to developing academic skills, the various modules require the use of meta-cognitive strategies associated with analytical reading. Close inspection of the various software programs and the interaction of the programs by students reveal that students, in order to be successful, are required to monitor their reading, focus upon salient characteristics of the complete text, and to reread to check on their understanding and/or to confirm their selection of target items. Even when unsuccessful, attention to the computer feedback signaled the students to reprocess the information in various ways ranging from rereading to eliminating reexamined information.

A 2002 analysis of 95 reviews of rational number software published over the last 20 years found a lack of implicit rubrics about how and what students will learn mathematically as they utilize particular educational software (Kafai, Franke, & Battey, 2002). The study suggested that review criteria based on principles of mathematical inquiry could help reviewers more accurately evaluate the actual potential and benefits of investigated software and give teachers better information on how to choose and integrate educational software into their classrooms. Merit Software possesses characteristics of exemplary software programs: unit activities support a balanced, integrated mathematics program that includes *reading, writing, thinking and word study*. The software features skill and review that promote automatic recall of core content and concepts that leads to effective, efficient problem-solving with tasks that are more difficult (Bruning, Schraw, and Ronning, 1999). In addition, Merit software supports constructivist learning by providing students with choices, decisions, and multiple completion paths for problem solving, enhancing cognition, and learner motivation. Merit Math software provides positive, formative feedback and scorekeeping and includes record keeping that enables students and teachers to monitor progress (Spitzer, 1996). Merit programs afford students temporary, flexible scaffolding at the point of difficulty. When students encounter a problem with a program unit, they can access several forms of assistance that review concepts, show examples, and provide opportunities to retry a skill in a supported environment. Unlike more structured programs, students and teachers can alter sequence and repetitions within units to maximize learning, a feature Kemp (1997) identified with effective software. A set criteria needs to be attained before a student can satisfactorily complete the exercises. The exercises can be reset as required to master the concept and can immediately re-enter the sequence of unit activities with different examples for additional practice. Students control the degree of assistance.

The review of the literature supports that CAI is an effective classroom tool for raising student scores on standardized achievement tests and appears particularly effective for students of lesser abilities or lower achievement levels. The several studies which report successful applications of CAI must be evaluated separately for the strength of each design.

Studies which follow Wenglinsky's recommendation to follow students over time, and measures academic achievement pre and post to CAI, are currently the exception.

Research Questions

This study, informed by the issues raised in the review of the literature, set out to examine the impact of Merit Math software on student learning and achievement, and specifically on low achieving students through a quasi-experimental study with random groups. Three questions directed the research:

1. Is there an overall change (increase) in pre to post Westest mean scores for 7th and 8th graders at HMS?
2. Is there a difference in the post Westest mean scores among 7th and 8th graders who participated in Merit instruction compared to their peers who were given traditional instruction?
3. Did socioeconomic variables (i.e., ethnicity and eligibility for free lunch) have any influence on Westest mean scores?

Project Setting

Kanawha County The Community

Kanawha County, West Virginia, includes both urban and rural sections with a population of 195,218. The per capita income is $32,789, and the median household income is $35,355. Over one-third of the population report their professions as management, professional, and related occupations. Eighty percent of persons over 25 have graduated high school and 20.6% of those over 25 have a Bachelor's degree or higher. In the city of Charleston, 32.6% of persons over 25 has a Bachelor's degree or higher. Eighty percent of the population of the city of Charleston is White; 15.1% of the city of Charleston's population is African American.

The Kanawha County School District

Horace Mann Middle School is located in Charleston and for the 2004-2005 year it served 411 students in grade 6 through grade 8, with a student teacher ratio of 15 to 1. Fifty percent of Horace Mann Middle School students qualify for free or reduced-price lunch compared with a state-wide average of 54%. Sixty-three percent of students are White and 36% are African American.

Curriculum

The West Virginia Content Standards and Objectives (CSOs) describe the knowledge expected for all students at every grade level, including those with disabilities. The mathematics curriculum for grades six through eight at Horace Mann Middle School is aligned to West Virginia Mathematics Content Standards and Objectives (CSOs). Teachers are expected to use programs and learning experiences that support the CSOs.

The West Virginia Department of Education changed its standardized assessment instrument from the norm-referenced SAT-9 to the WESTEST, a customized, criterion-referenced test aligned to the CSO's noted above. The WESTEST, in addition to scaled scores, identifies students at the Distinguished, Above Mastery, Mastery, Partial Mastery and Novice levels. The cut scores for each range were field tested between 2003 and 2005 and continue to be refined.

Procedures

The Merit Mathematics Program assists students with mathematical problem solving and application exercises employed in four areas of study: Fraction Shape-Up;Pre-Algebra Shape-Up;Basic Algebra Shape-Up; and Word Problem Shape-Up.

Software units consisted of four sets of exercises that promoted skill development and strategic thinking for the following:

- Try-out is a pre-test that can provide the teacher and student with information about relative skill and strength.
- Warm-up isolates a skill and provides several opportunities to perform that skill. Feedback is provided and set criteria need to be reached before a

student can satisfactorily complete the exercises.

- Work-out is a more rigorous exercise that inter-mixes mathematics skills. Feedback is given, and percentage of accurate responses much be reached before completion is achieved.
- Finals serve as a post-test. The program includes record keeping and helps monitor students' progress.

Treatments.

Treatment group received two 45-minute sessions per week of Merit treatment for 9 weeks. Where the Merit programs displaced curricula, the treatment groups did not make up the work that the comparison groups completed. The treatment group did not utilize the Glencoe Math software program that was used by the comparison groups.

The comparison group followed the regular mathematics curriculum, tied to West Virginia CSOs. Teachers selected materials and developed instructional activities that enabled students to master the content standards and taught these during two 45-minute daily math blocks. Grade seven students had lessons based on the *Mathscape- Seeing and Thinking Mathematically* series of books. Grade eight students had lessons based on the *Mathscape-Seeing and Thinking Mathematically* series of books and the *Glencoe Math* software program.

Study Design and Methodology

This study is a quasi-experimental, two group, pretest to posttest design, with random assignment to experimental (Merit) and control conditions (traditional content instruction). Dependent measures included standardized test scores in four content areas: Mathematics, English/LA; Science, and Social Studies on year-end, state-mandated tests. Data was analyzed separately and comparatively using a t-test for dependent samples to measure the differences in WESTEST mean scores for the Merit and control groups. The influence of socioeconomic status (eligibility for free and reduced lunch) and student ethnicity were also analyzed. Finally, a post hoc analysis was obtained for those in the lower quartile.

Findings of the Study

To obtain an overall analysis, the data sets had to be resorted and "cleaned" due to a number of incomplete and missing cases of pre and post test data. For example, some students may have been absent on pretest days but present on posttest days, or vice-versa. Thus, an original data base which numbered 177 was sorted to 109 subjects. The two occasions for data collection were for 7th and 8th graders completing year-end state-mandated assessments in 2004 (Pretest) and in 2005 (Posttest). Descriptive data are shown in Table 1. In addition to the descriptive data, paired samples correlations were obtained to determine equal variances. All four pairs were significant at p .000.

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	WestMath_0304	683.0550	109	31.05307	2.97434
	WestMath0405	699.0092	109	32.52478	3.11531
Pair 2	WestRLA_0304	677.8073	109	33.26200	3.18592
	WestRLA0405	688.4220	109	29.77144	2.85159
Pair 3	WestSci_0304	685.0550	109	30.41955	2.91366
	WestSci0405	695.4404	109	38.04297	3.64386
Pair 4	WestSocSt_0304	679.0092	109	28.99984	2.77768
	WestSocSt0405	687.2477	109	31.50666	3.01779

Table 1 Descriptive Data for Westest Content Area Pairs

A dependent samples t test was obtained to measure the overall differences in Westest mean scores. These measures were obtained independently for each of the content areas (*Mathematics; Reading/LA; Science and Social Studies*).

These results are depicted in Table 2 and are statistically significant for all four Westest mean score pairs (p .000, SPSS Version 13.0).

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	WestMath_0304 - WestMath0405	-15.95413	16.88435	1.61723	-19.15975	-12.74850	-9.865	108	.000
Pair 2	WestRLA_0304 - WestRLA0405	-10.61468	19.90575	1.90663	-14.39394	-6.83542	-5.567	108	.000
Pair 3	WestSci_0304 - WestSci0405	-10.38532	26.87209	2.57388	-15.48720	-5.28344	-4.035	108	.000
Pair 4	WestSocSt_0304 - WestSocSt0405	-8.23853	18.57413	1.77908	-11.76497	-4.71209	-4.631	108	.000

Table 2 Paired Sample t test for Westest Content Pairs

Inspection of the paired mean differences in Table 2 shows a statistically significant increase in mean test scores for the four content areas from pre to post. Additionally a Bonferroni adjustment was made

beforehand for the alpha level with a test of significance at .05 to control, for family wise error, which yielded a p of .0125

How significant or dramatic were these differences? Effect size measures were calculated for the four pairs to determine the magnitude of the change. These results were as follows: Math (.844); Reading/LA (.223); Science (.132) and Social Studies (.166). Following effects size guidelines of Cohen (1998) (small effect = .01; moderate effect = .06; and large effect =.14), the obtained values for the study pairs were large. The effect size of .844 for Mathematics is an extremely large value.

Although significance is obtained for the pre and post scores Westest scores, other variables can potentially account for some of the "good" variance. Socioeconomic factors are always a factor to consider when interpreting gains (or losses) in standardized achievement test scores. These factors (ethnicity and eligibility for free lunch) were incorporated into the current study to determine if such influence affected the outcomes. Separate independent samples t-tests were obtained for the four pairs to compare the groupings. In each case, the equality of means test (Levene's) was not significant, indicating that the various distributions had equal variances. The independent t test results were not significant for the pairs, with values between .441 and .925, indicating that ethnicity was not a major factor in the overall outcomes. A similar, but lesser effect was found for "free lunch" which also resulted in no differences for Mathematics, Science and Social Studies, but a "near effect" for Reading/LA, ($p > .07$).

Lower quartiles are of interest given their importance in school performance compliance. Westest scores for those in the bottom quartile for merit (11) and non merit (13) were compared using an independent t test. These results were not significant (p .104) but yielded means of 654 for Merit compared to 646 for Non-Merit. Although not significant, it is an important 8-point difference if it can be replicated with large sample sizes..

Recommendations

To further increase the validity of the comparison of students in treatment groups (those using computer assisted software) and those in the control groups (those using more traditional methods of instruction), detailed records should be kept regarding the amount of time each group spent on each concept subject or content area. Also, the same pretest and posttest measures should always be administered to both groups.

The amount of time spent utilizing the Merit software may impact the results obtained through the use of the Merit (and other) computer software. The most recent NAEP report (Sandene, 2005) indicates that student degree of familiarity with computers can play a significant role in test results obtained from computer administered tests. To further refine studies of the effectiveness of computer software, a test should be designed which could reliably determine student computer facility for both training and testing purposes. Additionally, treatment periods should be extended for the entire semester, or for a minimum of 18 weeks to ensure "bonding".

Teachers and students using the Merit software modules should be interviewed during and after use of the software to determine how their experiences may contribute to modifications in the design of future editions of the modules. A comparison of the views of high and low achieving students may hold valuable clues as to how educational software might be improved to benefit all students, and not just those who are low achieving.

References

Bindig, Charles W. (2002), "Effect of Tutorial Software on Student Achievement in Writing," Nova Southeastern University, Applied Dissertation Presented for EDD Program for Educational Leaders in Partial Fulfillment of the Requirements for the Degree of Doctor of Education.

Brown, F. (2000). Using Computer Assisted Instruction to Teach Mathematics: A Study. The NABSE Journal, December 2000, *4* (1), 62-72.

Bruning, R. H., Schraw, C. J., & Ronning, R. R. (1999). *Cognitive psychology and instruction*. Upper Saddle River, NJ: Prentice-Hall.

Christmann, E. (2000). Relevant Research on Computer-

Assisted Instruction in Science *New Mexico Middle School Journal*, Spring 2000, 41.

Christmann, E, & Badgett, J. (1999). A comparative analysis of the effects of computer-assisted instruction on student achievement in differing science and demographic areas. *Journal of Computers in Mathematics and Science Teaching*, 18 135-143.

Kafai, Y,, Franke, M..L., & Battey, D. S.(2002). Education, Communication & Information.December 23 (*2*), 163.

Kemp, S. (1997). Software characteristics can affect student achievement. *Curriculum Administrator*, 32 (4), 48.

Kemp, S. (1997). Trained educators well designed software and technology aids student achievement and attitudes. *Curriculum Administrator* 32 (2), 73.

Sandene, B, Horkay, N., Bennett, R. E.; Allen, N.; Braswell, J.; Kaplan, B., and Andreas Oranje Online Assessment in Mathematics and Writing: Reports From the NAEP Technlogy-Based Assessment Project, Research and Development Series. National Center for Education Statistics. Retrieved online September 25, 2005 at Http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005457

Schacter, J. (2000). The Impact of Education Technology on Student Achievement: What the Most Current Research Has to Say. Milken Exchange on education technology.

Spitzer, D. R. (1996). Motivation The neglected factor in instructional design. *Educational Technology* 36(3), 45-49.

Wenglinsky, H. (1998). Does it Compute? The relationship between educational technology and student achievement in mathematics. Educational Testing

ABOUT THE AUTHORS

Dr. Securro is a faculty member in the Elementary and Secondary Program who regularly teaches graduate courses in educational research and writing and advanced, computer-based statistical analysis. He prepared the research design and conducted the statistical analysis, tables and related results for the present study.

Dr. Jones is a faculty member in the Educational Leadership Studies Program who regularly teaches graduate courses in school supervision and the role of the principal. He conducted the study in the field and was responsible for school personnel orientation data collection and subject selection and arrangement in the school.

Mr. Blackwell is the chief school administrator for Horace Mann Middle School who made the arrangements for the study, promulgated the various permission and consent levels needed for participants and managed the technical environment for curriculum software delivery.

Mr. Cantrell was the principle writer for the project and was responsible for the selection and review of related background studies and arranging and narrating the final product. He is currently pursing a doctorate at Marshall University Graduate Center in Educational Leadership.

Dr. Securro

Dr. Jones