



SuccessMaker
Evidence of Effectiveness
Selected Evaluation Studies

Pearson Digital Learning
A business of Pearson Education

SuccessMaker Evidence of Effectiveness

Selected Evaluation Studies

	Page
Pilot Implementation at Ball’s Bluff Elementary School Loudoun County, VA, 1998	1
Duval County Public Schools Evaluation Report for 1999-2000 Duval County, FL, (Pearson Education Technologies 2000)	2
Relationship Study for SuccessMaker Levels and SAT-9 in Hueneme Elementary District, School Year 2000-2001, with Growth Analysis Pt. Hueneme, CA (Pearson Education Technologies 2001)	3
Computer-Integrated Learning System and Elementary Student Achievement in Mathematics: An Evaluation Report Unpublished dissertation, Temple University, 1995	4
Student Achievement in Mathematics and the Use of Computer-Based Instruction in the Hempfield School District, Five-year study, Hempfield School District, Landisville, PA 1998	5
A Three-Year Longitudinal Study Assessing the Impact of the SuccessMaker Program on Student Achievement and Student and Teacher Attitudes in the Methacton School District Elementary Schools Lehigh University, 1998	6
Technology Literacy Challenge Grant Application, Pamlico County Schools, NC	7
An Interim Evaluation of Operation Safety Net, A Five Year Project Miami-Dade County, FL	8
Seminole County Public Schools Relationship Study for 2000-2001 Seminole County, FL (Pearson Education Technologies 2001)	12
Integrated Learning Systems in UK Schools, 1994; Integrated Learning Systems: A report of Phase II of the pilot evaluation of ILS in the UK, University of Leicester, 1996	13
Community School District 6 [New York City Schools] Integrated Technology Reading Support Project: First, Second and Third Year Evaluation Report[s], 1996, 1998, 1999	15
Meta-Analytic Studies of Findings on Computer-Based Instruction James A. Kulik (in <i>Technology Assessment in Education and Training</i> , editors Eva Baker and Harold O’Neil, Lawrence Erlbaum Associates, New Jersey 1994)	17

Title of evaluation: Pilot Implementation at Ball's Bluff Elementary School: Evaluation

Author: Office of Research, Loudoun County Public Schools, Virginia

Relation of author to model: Implementer: District Office of Research

Date evaluation report completed: April 1998

Months/years covered by evaluation: One school year, 1996-97

Number of schools involved: One subject school and two control schools; 254 students at subject school included.

Demographics of sample population: Most students in grades 2-5 in the subject school were included. The school has 22% of students on free/reduced lunch; 31% minority.

Purpose of study: Evaluate the use of SuccessMaker mathematics on student achievement and teacher attitude about its impact on classroom instruction.

Methodology: Standardized achievement test data were used to compare student math achievement from Spring 1996 to Spring 1997. Students were included in the study if they had test scores for both years and had spent sufficient time in SuccessMaker for the program to have an effect. Statistical box plots were used to examine the entire distribution for each grade and median scores were used to quantify the comparison.

Measures/indicators used to assess effectiveness: In Spring 1996, students in grades 2, 3, and 4 took the ITBS. One year and one grade later, students in grades 3, 4, and 5 took the Stanford-9. The distribution of percentile scores was compared for the students with both test scores. Although these are different tests, they are comparable to the extent that they measure similar types of math performance, and are based on similar standardization groups. At Ball's Bluff, distributions of scores for minority students only were also examined. Matched student scores at two other elementary schools were also examined for achievement growth. Schedules at the schools were also compared to be sure results were not because of increased time on math.

Findings: At the subject school (Ball's Bluff) median scores improved at all three grades over the one-year period. The distribution shows that the students improved as a group, especially in grade 5, and not just as a result of a few high scorers. Scores for minority students also improved consistently. At the other two schools (Cool Springs and Leesburg) the scores declined or remained approximately stable, although one was also introducing a strong math initiative.

In summary, math achievement at Ball's Bluff improved broadly during the year of the pilot.

- Comparison schools that did not use SuccessMaker did not show similar improvements
- The improvement at Ball's Bluff does not appear to have been a result of increased time-on-task, nor to any other change in instruction
- The improvement appears to have been specific to math

(See also the story from Multimedia Schools "Reaching Every Student: The Loudoun County Story," January/February 2000, for a teacher's view of the program's impact on learning and teaching.)

Title of evaluation: Duval County Public Schools Evaluation Report for 1999-2000

Author: Barbara Tingey and Anthony Thrall, Ph.D.

Relation of author to model: Research department of developer

Date evaluation report completed: November 27, 2000

Months/years covered by evaluation: 1 school year

Number of schools involved: 12 schools, grade 4 reading and grade 5 mathematics

Demographics of sample population: The 12 schools were low-performing before the program was implemented, earning D and F grades in the Florida Accountability program. In these Duval County schools, 90% of the students have free or reduced lunch, and the schools are 90% minority, primarily Black.

Purpose of study: The study addressed the issue of efficacy of SuccessMaker courseware in improving student achievement. The study was designed to describe reading achievement for grade 4 students and math achievement for grade 5 students as measured by the Florida Comprehensive Assessment Test (FCAT) and to summarize the relationship of FCAT scores and SuccessMaker levels. The study also was designed to test the accuracy of the SuccessMaker target levels set in the previous (1999-2000) evaluation report based on current student data.

Methodology: SuccessMaker data were collected at the time the FCAT was administered. Test scores were matched to student's SuccessMaker data. Students were included who met SuccessMaker usage requirements and had matched FCAT scores. Linear regression analysis was applied to matched scores and SuccessMaker levels. Additionally, logistic regression was applied to determine target SuccessMaker levels set for specific test achievement levels.

Measures/indicators used to assess effectiveness: In the previous year, courseware targets were designed to ensure that about 80% of students reaching the target level would also meet or exceed the statewide average score on the FCAT. The current study examines the updating of the previous targets to meet the district goals of increasing the percentage of students who reach or exceed FCAT Achievement Level 2. Finally, SuccessMaker target levels set in the previous year were evaluated with current data.

Findings: Schools were low-performing before the program was implemented, earning D and F grades in the Florida Accountability program. In 2000, each of these 12 schools maintained or improved their grade; no school remained at an F grade, and 3 schools moved up to C.

In reading, 72 students out of 438 reached the courseware target level in reading for meeting the statewide average. Of these, 86% (62 students) also met or exceeded the year 2000 state average FCAT reading scaled score (293 in 2000 vs. 288 in 1999). By way of comparison, among the remaining 366 students who did not meet the courseware target, only 25% met or exceeded the state average score. In mathematics, 91 students out of 387 reached the courseware target level. Of these 91 students, 79% (72 students) also met or exceeded the year 2000 state average FCAT mathematics scaled score (314 in 2000 vs. 303 in 1999). But among the remaining 296 students who did not meet the courseware target, only 27% met or exceeded the state average score.

Thus the previously set courseware targets appear to be quite reliable. They were designed to assure an 80% success rate, as compared to actual success rates of 86% and 79% in reading and mathematics, respectively. This degree of reliability is especially impressive in light of the changes that occur from one year to the next; in this case, the state average moved up, and it was this more stringent, current state average that was used to evaluate the previous year's target.

SuccessMaker Evidence of Effectiveness

Title of evaluation: Relationship Study for SuccessMaker Levels and SAT-9 in Hueneme Elementary District, School Year 2000-2001, with Growth Analysis 2000-2001

Author: Barbara Tingey and Chuck Simon, Accountability & Implementation, Pearson Education Technologies

Relation of author to model: Developer

Date evaluation report completed: October 2001

Months/years covered by evaluation: 1 school year

Number of schools involved: Nine schools, grades 4 and 5, total of 597 students

Demographics of sample population: 73% free and reduced lunch; 80% minority, primarily Hispanic

Purpose of study: One purpose is to examine the achievement growth of students on the Stanford Achievement Test (SAT-9) from spring 2000 to spring 2001. Another purpose of the study is to summarize the relationship between student scale scores on the SAT-9 in reading and mathematics and the levels attained in SuccessMaker foundation software in order to set target SuccessMaker levels for future implementation.

Methodology: SuccessMaker data were collected at the time the SAT-9 was administered. Student scores on SAT-9 taken spring 2000 and spring 2001 were matched to their SuccessMaker levels at the time of the test in 2001. Students were included who met SuccessMaker usage requirements and had matched SAT-9 scores. For the relationship study, linear regression analysis was applied to matched 2001 scores and SuccessMaker levels. Additionally, logistic regression was applied to determine target SuccessMaker levels set for specific test achievement levels.

Measures/indicators used to assess effectiveness: The difference in each student's NCE scores from 2000 to 2001 was determined, and mean gains derived. Also, changes in the distribution of student scores by quintile was examined from 2000 to 2001.

Findings: Solid growth in mean gain was demonstrated by all groups, with mean gain of at least 5 NCE in grade 4 reading and grade 5 mathematics.

	Reading		Mathematics	
	Number of students	Mean Gain on SAT9	Number of students	Mean Gain on SAT9
Grade 4	307	5.90 NCE	328	1.60 NCE
Grade 5	290	4.14 NCE	321	5.02 NCE

The change in distribution by quintiles revealed a decrease in students in the lowest quintiles and increase in the top quintile. Here are the values for reading for the students in grade 5 in 2001.

	1-19 th percentile	20-39 th percentile	40-59 th percentile	60-79 th percentile	80-99 th percentile
2000	18%	26%	22%	21%	12%
2001	10%	21%	24%	26%	18%

Target SuccessMaker levels set for these grade levels for specific performance are being used in the current implementation. This information is part of a district-wide accountability system for raising student achievement. They monitor student achievement year-round, identify areas of difficulty, and provide focused instruction and interventions to improve learning, and ultimately performance on the Stanford Achievement Test.

Title of evaluation: Computer-Integrated Learning System and Elementary Student Achievement in Mathematics: An Evaluation Report (Hempfield SchDist)

Authors: Cheryl M. Laub (unpublished dissertation to Temple University Graduate Board)

Relation of author to model: Implementer (principal of Centerville Elementary)

Date evaluation report completed: May 1995

Months/years covered by evaluation: 7 months

Number of schools involved : 314 students in grades 4 and 5 at two elementary schools

Demographics of sample population: Suburban, computer-literate students. Students receiving free or reduced lunch were 11% of the population in one school and 18% in the other.

Purpose of study: The study was designed to determine if the program in place was producing desired results. The study was designed to investigate the relationship of student achievement in mathematics and the use of the SuccessMaker courseware Math Concepts and Skills, and to investigate the rate of learning across students who are low-, average-, and high-achieving. The study also gathered teacher perception of implementation, impact on achievement and instructional strategies, and teacher preparation.

Methodology: Students used SuccessMaker 12 minutes a day on computers in their classrooms to supplement their 45-minute math instruction. Math achievement of students in grades 4 and 5 were measured in scaled scores using the Stanford Achievement Test (SAT) and compared to mathematics achievement of two years of previous fourth and fifth grade classes in the same schools. Tests were administered in the fall and the spring. The fall test scores were used to designate students as low (stanines 1,2,3), average (stanines (4,5,6), and high-achieving (stanines 7,8,9). Change scores were calculated and used to measure student achievement. Analysis of covariance was used to adjust initial group differences and determine whether the difference between treatment and control scores of low-, average-, and high-achieving students were statistically significant. The 12 teachers responsible for these students were surveyed to guide the planning of ongoing in-service training and the use of SuccessMaker with specific student populations. Teacher perceptions were measured using two questionnaires designed by the researcher.

Findings: Statistically significant results were found for the following:

- Mean change score of students using the computer (50.15) was greater than the mean change score of earlier classes who did not (28.2). A dependent t-test indicated the gains were affected by the use of SuccessMaker and not by chance.
- Higher-achieving students' mean scores increased more than average-achieving students' scores. Average-achieving students' mean scores increased more than the scores of low-achieving students
- Teacher surveys showed positive response about the adequacy of initial workshops and ongoing support during implementation.
- Teacher perceptions of the advantages of the system centered on student learning and teacher instruction. Disadvantages cited were less focused, and included technical problems with hardware, and coordination of courseware with textbooks.

A majority of teachers indicated the use of the program modified their instructional strategies, especially to examine instructional sequences and modify them based on data.

Title of evaluation: Student Achievement in Mathematics and the Use of Computer-Based Instruction in the Hempfield School District

Authors: C.M. Laub, Ed.D. and R. L. Wildasin, Ed.D., Hempfield School District, Landisville, PA

Relation of author to model: Implementer

Date evaluation report completed: September 1998

Months/years covered by evaluation: Five years (1993-1998)

Number of schools involved : Six elementary schools, more than 500 students

Demographics of sample population: Students in the class of 2004 across the district; average free/reduced lunch by school is 10%-24%; minority students in a school range from 6% to 22%.

Purpose of study: The SuccessMaker program focused on raising achievement in mathematics for students in grades 2 through 6 to significantly reduce the number of students in the lowest quartile and increase the students in the top quartile. The study followed district students who were in grade 2 in 1993 over 5 years to measure growth.

Methodology: Student scores on the Stanford Achievement Test in May of each year from grades 3 through 6 were tracked. (In 1998, the test was the Stanford-9) SuccessMaker was the primary innovation in math instruction. Effectiveness was evaluated through longitudinal analysis of achievement.

Measures/indicators used to assess effectiveness: Analysis included comparison of the national percentile equivalent of the mean scaled score each year and the distribution of students by quartile.

Findings: The data across the years demonstrates that students showed continued growth in mathematics achievement as measured by the test. The mean scaled score of the students in the study increased from 70th percentile in 1995 when students were in grade 3 to 80th percentile in 1998 when students were in grade 6. The grade-equivalent of the mean score rose from 4.7 for grade 3 to 10.2 for grade 6. Most importantly, the percent of students in the lowest quartile decreased from 12% to 6%, while the percent of students in the top quartile increased from 41% to 59%.

SAT Measures of Mathematics Achievement: Class of 2004

Means	May 1995 Grade 3 N=525	May 1996 Grade 4 N=524	May 1997 Grade 5 N=515	May 1998 Grade 6 N=522
Scaled Score	617	653	679	693
Percentile	70	77	82	80
Grade equivalent	4.7	6.5	8.5	10.2

This was a successful district with students on average performing above grade level to start with. The findings show they further increased achievement for all segments of the student population.

Title of evaluation: A Three-Year Longitudinal Study Assessing the Impact of the SuccessMaker Program on Student Achievement and Student and Teacher Attitudes in the Methacton School District Elementary Schools

Author: Leroy J. Tuscher, Ph.D., Lehigh University

Relation of author to model: Independent evaluator for Methacton School District

Date evaluation report completed: 1998

Months/years covered by evaluation: Three years (September 1995 to June 1998)

Number of schools involved: Four elementary schools, students in grades 3-5

Demographics of sample population: 6% of students free and reduced lunch; 7% minority

Purpose of study: The study was designed to measure the impact of using SuccessMaker mathematics on student achievement, to assess whether teacher and student attitudes towards using technology changed over time, and to determine whether technology changes the way instruction is delivered.

Methodology: Students worked a minimum of 12 minutes a day in SuccessMaker in addition to regular math class time. Student achievement patterns in SuccessMaker were compared with achievement patterns on the standardized test. The Stanford Achievement Test-9 was administered each year, to determine student achievement levels. To assess changes in attitudes between the beginning and end of the study, surveys were administered to students and teachers. The student surveys measured attitudes to 1) Mathematics and 2) Technology and Mathematics. The teacher surveys measured attitudes to 1) Computer Technology and 2) Computer Technology Impact on Instruction, Teaching, and Learning.

Measures/indicators used to assess effectiveness: An analysis of Stanford-9 results was completed for each year, using the previous year's test results as the pretest. Test results for two years prior to SuccessMaker use (1992-94) were compared to the results for the years of SuccessMaker use (1995-98). Scores used were SAT-9 Mathematics Total, and subtests of Calculation and Application. Student attitudes were collected on an 8-item, seven point Semantic Differential bi-polar adjective scale. Analysis was conducted to measure any changes in student attitudes toward mathematics, language arts, mathematics and technology, and language arts and technology. Teacher attitudes toward use of technology for instruction and how the implementation changed their instructional processes were measured on a five-point Likert scale.

Findings: The comparison of SAT-9 quartile rankings for pre-SuccessMaker (1992-94) and post-SuccessMaker program implementation (1995-98) reveals that the percentile rankings averaged over three years of the study exceed the average percentile rankings of four years prior to the study for students in grades 3, 4, and 5.

The teacher attitude survey revealed that teachers were more comfortable with computers in 1998 than in 1995, and agreed more that computers would change how schools educate children. In both 1995 and 1998 they agreed that technology helps meet the needs of both high and low ability students" and "encourages students to become self-directed learners." In 1998 they were more strongly in agreement that "Computer technology has forced me to rethink the ways in which I have traditionally delivered instruction" than in 1995.

Title of evaluation: Technology Literacy Challenge Grant Application, Pamlico County Schools

Author: Judy Humphries, Director of Technology, Pamlico County Schools, NC

Relation of author to model: Implementer

Date evaluation report completed: June 1997

Months/years covered by evaluation:

Number of schools involved: Students in eleven classrooms in grades 3 and 4 in Fred Anderson Elementary and in grades 5 through 8 at Pamlico County Middle School

Demographics of sample population: 56% of students have free/reduced lunch; 34% are minority. Prior to this program 1/3 of district students in grades 3-8 were below grade level in reading.

Purpose of study: This study was designed to measure achievement gains in reading for grade 3 and 4 students at Fred Anderson Elementary, and grades 5 through 8 students at Pamlico County Middle School. Prior to this program, in tests from 1993 through 1996, only about 66% of students reached proficient level in reading on the North Carolina End-of-Grade reading tests

Methodology: . A pilot program from December 1995 to May 1996 with Priority I Title I students in grade 4 demonstrated gains on the North Carolina End-of-Grade tests, with half the students showing gains of two years or more in six months. In this program, teachers integrated the technology into the classroom. The model was designed to integrate technology into the teaching and learning process, and included intensive staff development, integration strategies, analysis of student performance data, incentives for students, and meaningful parent involvement.

SuccessMaker was part of a combined strategy that included other initiatives as part of the instructional program. Volunteers were assigned to weaker students to give additional support. Elementary students use *Reading Readiness*, *Initial Reading*, *Reader's Workshop*, and *Reading Adventures* as appropriate for their level. Middle School students used the literature-based reading courses *Reading Adventures* and *Reading Investigations*.

Measures/indicators used to assess effectiveness: North Carolina's End-of-Grade Testing Program was the assessment, reported as a performance model (percent of students proficient in reading, writing, and math) and a growth model (requiring all students to make significant growth each year). The goal was to increase the percentage of students demonstrating proficiency in reading each year, with a goal of 82% proficiency by 1999-2000.

Findings: Students made significant gains since the 1995-96 school year. The number of students at proficiency level was above 75% by 1998 for all grades except grade 6, which reached 70%. Both Fred Anderson Elementary and Pamlico Middle School were recognized by the North Carolina State Board of Education as Exemplary Schools for the 1997-98 school year as a result of student growth.

In 2000-2001, Fred Anderson Elementary had nearly 81% of students at proficiency level and earned Exemplary growth and School of Distinction honors. Pamlico Middle had more than 84% proficient and earned School of Distinction honors.

Title of evaluation: An Interim Evaluation of Operation Safety Net, A Five Year Project

Author: Jerome L Levitt, Ph. D., Office of Evaluation and Research, Miami-Dade County Public Schools

Relation of author to model: Implementer - District evaluation department

Date evaluation report completed: April 2000

Months/years covered by evaluation: 1996-97 to 1998-99

Number of schools 44 treatment schools in the District Evaluation compared to 12 district Title I schools with similar demographics.

Demographics of sample population: 93% poverty, large percentage of Hispanic and Haitian students.

Purpose of study: Evaluation of the effectiveness of the Operation Safety Net (OSN) Program implemented in some district Title I schools. Analysis of teacher attitudes about the program and school climate; and parent and student perception of school climate, based on surveys. Analysis of data from each of the programs used in Operation Safety Net; Success-for-All, Jostens Learning (now Compass Learning) and Computer Curriculum Corporation (now part of Pearson Education Technologies). Recommendations made for the last two years of the program.

Methodology: This evaluation examined the first three years of implementation, focusing on reading achievement as measured by Stanford-8 reading comprehension subtest. Survey results for teachers, students, and parents are included in the report, using rank order choices, converted to numerical data.

Measures/indicators used to assess effectiveness: Longitudinal analysis of the Stanford Achievement Test 8, reading comprehension subtest data.

Findings:

- Reading comprehension achievement over the period showed an upward trend for both OSN and other (non-OSN) Title I schools used for comparison. However, an initial gap in performance did not appear to be closing in subsequent years.
- An analysis of achievement scores for schools using [SuccessMaker] software compared to scores for schools using Jostens software showed greater achievement gains for the [SuccessMaker] schools.
- Teacher survey results, obtained after the first year of program implementation, showed both negative and positive reaction to the Success-for-All curriculum, in about equal proportions.
- Teacher climate survey results (selected items), which were obtained to illustrate the perceived environment in which program teachers worked, showed that OSN teachers were experiencing far greater challenges than staff at non-OSN schools.
- Parents with children in OSN schools had more favorable attitudes toward their children's computer training and critical thinking/reasoning than did parents from the comparison schools.
- Students at OSN schools had the lowest scores regarding their feelings about school and reading in comparison to those students in other non-OSN schools.

Comparisons of data:

“...for students who were in the third, fourth and fifth grades, the [SuccessMaker] participants appeared to be improving at a slightly higher rate.”

Quotations from the study:

“Unfortunately, students at some schools learned that there was no penalty for wrong answers [in the Jostens Learning Company software] and they could just try each of the answers in turn until they ‘hit’ the correct one... Due to this problem, schools implementing the JLC software quickly reached the goals and in fact ‘ran out’ of software. That is students went through the material too quickly.”

“ [SuccessMaker] software appeared to handle students’ random attempts with more sophistication....Thus, this system may have been more accurate in gauging student performance.”

“Computer Curriculum Corporation (Pearson Education Technologies) met its shared accountability reporting requirements and met its shared accountability criteria. In addition, there was a significant positive trend for OSN students using [SuccessMaker] software compared to OSN students using Jostens Learning Corporation’s (Compass Learning) software.”

Recommendations (paraphrased and condensed):

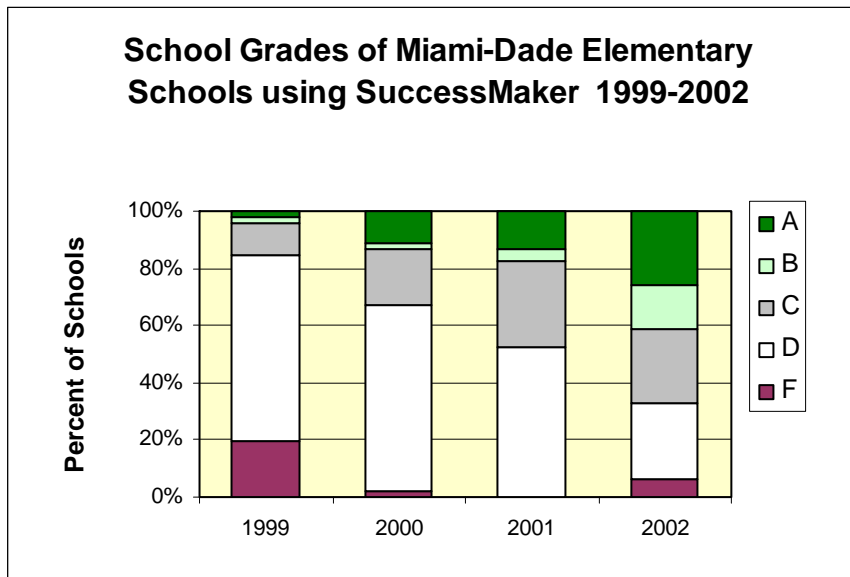
Contracts with computer-assisted instructional providers should be renewed. The results of this evaluation show promise for at least one of the companies, based on Stanford Achievement Test results, even though students were exposed to the computer software for only twenty minutes a day. Parents valued the use of computers in the schools. The role of that instruction should be expanded and be made available for parents.

Note: The first three years of Operation Safety Net were focused on student gains, and only foundations courses were used. In the last two years of the program, other courses proposed in the instructional plans are being used, such as the English language development, literature-based and tools-based courses. Pearson Education Technologies would like to complete its own study of the SuccessMaker Operation Safety Net Schools and include examination of the implementation in different years. The state assessment data (SAT and FCAT) show that achievement in math has improved significantly over the life of the program, and reading has improved in most schools with students moving out of the lowest quintile, particularly in the last two years. Only one school of the original 19 SuccessMaker schools is still an F school. (Other indicators are also measured in addition to student test performance.) The district has supported the use of the SuccessMaker program, setting up programs in another 19 schools in 2001.

Analysis of Growth 1999-2002 for SuccessMaker Schools in Miami-Dade County

In Miami-Dade County, 6 originally low-performing elementary schools used SuccessMaker in 1995, designated the Model Schools program. In 1996, 19 schools in Operation Safety Net began using SuccessMaker. By 2000, an additional 21 low-performing schools were added (designated the Plus schools), for a total of 46 schools. School grade results in the Florida School Accountability program over time are shown below.

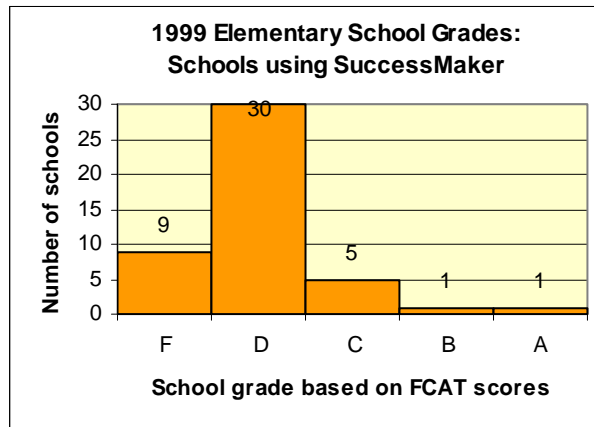
In 1999, 85% of the Miami-Dade schools using SuccessMaker in 2002 (39 out of 46) had a school grade of D or F based on Florida Comprehensive Assessment Test (FCAT) scores. By 2002, only 15 schools (33%) were still at D or F, and 19 (41%) earned an A or B. Changes over the years show steady growth, with particularly positive results for this group of schools in 2002.



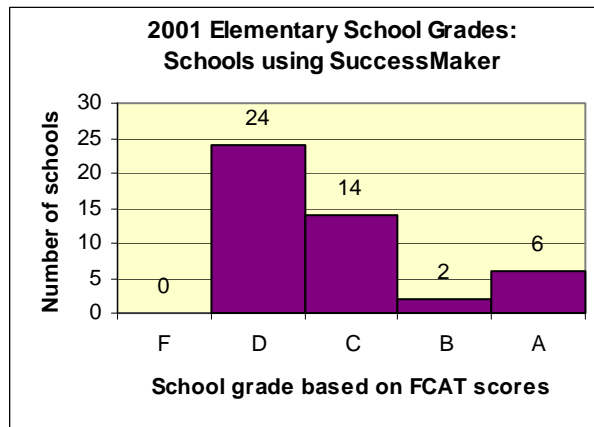
SuccessMaker Evidence of Effectiveness

Improvement in School Accountability rating of Miami-Dade elementary schools using SuccessMaker 1999-2002

Miami-Dade's Operation Safety Net and Model Schools program focuses on improving the lowest-performing schools. Among the 46 Miami-Dade elementary schools using SuccessMaker in Operation Safety Net, Plus, and Model Schools program in 2002, in 1999 only 7 of the 46 schools (15%) earned grades of C or better in the Florida School Accountability program, based on distribution of FCAT scores.



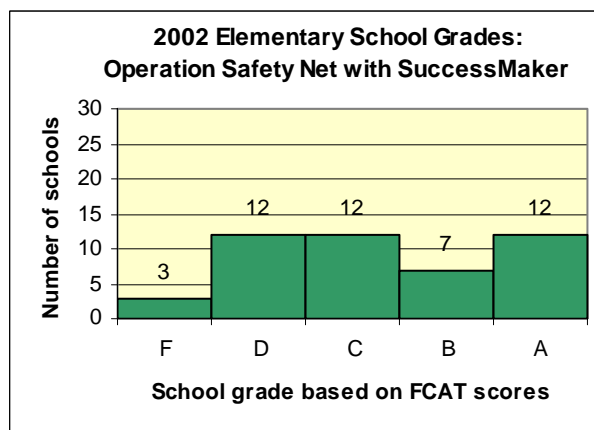
Two years later in 2001, these schools show marked improvement. (Some schools started using SuccessMaker in 2000.) In 2001 22 out of the 46 schools (48%) earned grades of C or better.



Between 2001 and 2002, the school grades improved for a high proportion of the schools. In 2002, 31 out of 46 schools (67%) earned a school grade of C or better, under stricter criteria. The number of schools earning a grade of A doubled (from 6 to 12).

Changes for Safety Net Schools in 2002 compared to 2001

Grade changes	Number of schools
Down one grade	4
Maintained grade	22
Up one grade	14



Title of evaluation: Seminole County Public Schools Relationship Study for 2000-2001

Author: Chuck Simon and Barbara Tingey, Pearson Education Technologies

Relation of author to model: Research department of Developer

Date evaluation report completed: September 28, 2001

Months/years covered by evaluation: 1 school year

Number of schools involved: 12 schools, grade 4 students and grade 5 students

Demographics of sample population: All the schools are Title I schools. They average 56% free and reduced lunch and 42% minority.

Purpose of study: The study addressed the issue of efficacy of SuccessMaker courseware in improving student achievement. The study was designed to describe reading achievement for grade 4 students and math achievement for grade 5 students as measured by the Florida Comprehensive Assessment Test (FCAT) and to summarize the relationship of FCAT scores and SuccessMaker levels. The study also was designed to test the accuracy of the SuccessMaker target levels set in the previous (1999-2000) evaluation report based on current student data.

Methodology: SuccessMaker data were collected at the time the FCAT was administered. Test scores were matched to student's SuccessMaker data. Students were included who met SuccessMaker usage requirements and had matched FCAT scores. Linear regression analysis was applied to matched scores and SuccessMaker levels. Additionally, logistic regression was applied to determine target SuccessMaker levels set for specific test achievement levels.

Measures/indicators used to assess effectiveness: In the previous year, courseware targets were designed to ensure that about 80% of students reaching the target would meet or exceed the designated FCAT level. The current study examines the updating of the previous targets to meet current population data. Finally, SuccessMaker target levels set in the previous year were evaluated with current data.

Findings: In reading, 257 students out of 611 reached the courseware target level in reading for FCAT achievement level 3. Of these 257 students, 86% (211 students) met or exceeded FCAT achievement level 3 for 2001. By way of comparison, among the remaining 354 students who did not reach the courseware target, only 36% met or exceeded achievement level 3..

In mathematics, 126 students out of 459 reached the courseware target level for FCAT level 3 attainment.. Of these 126 students, 90% (113 students) met or exceeded FCAT achievement level 3 for 2001. But among the remaining 333 students who did not meet the courseware target, only 35% met or exceeded achievement level 3..

Thus the previously set courseware targets appear to be quite reliable. They were designed to assure an 80% success rate, as compared to actual success rates of 86% and 90% in reading and mathematics, respectively.

Title of evaluation: Integrated Learning Systems in UK Schools, 1994; Integrated Learning Systems: A report of Phase II of the pilot evaluation of ILS in the UK, 1996

Author: (Phase I) J. Underwood with S. Cavendish, S. Dowling, K. Fogelman, and T. Lawson. (Phase II) J. Underwood with S. Cavendish, J. Gardner, C. Harrison, A. Lewis, S. Rodrigues, D. Passey, C. Fitz-Gibbon, N. Defty, and S. Dowling

Relation of author to model: independent evaluator, Leicester University, subsidized by the British Ministry of Education

Date evaluation report completed: 1996

Months/years covered by evaluation: 2 years

Number of schools involved: 23 schools

Demographics of sample population: Most schools were in low socio-economic areas in England, Scotland and Wales, with 50% free school meals.

Purpose of study: Phase I of the evaluation focused on “whether Integrated Learning Systems could improve pupils numeracy and literacy skills.” At each school, a control group of peers was tested as well as those using SuccessMaker. Phase I of the evaluation lasted 6 months.

Phase II focused on whether the learning gains found in Phase I were repeatable, what factors affected gains, whether gains were sustained for students continuing use of the software and for those that were no longer using it, and which groups of students particularly gain value from the use of SuccessMaker.

Leicester University evaluated three different treatments: GLOBAL software, SuccessMaker software and schools without integrated learning systems. A small study of Jostens software is described separately in the report. Small scale investigations of the use of the software for particular kinds of learners was also undertaken.

Methodology: In the initial study, nine schools used SuccessMaker, four used GLOBAL, and six schools agreed to act as control schools. The schools were chosen prior to the study, independently of the evaluators. In the second phase, fourteen schools were used, all having groups which used ILS interventions and groups which did not. Appropriate statistical techniques allowed the evaluators control of the intervention. The ILS groups had periods on and off the software to determine if learning gains were maintained

Measures/indicators used to assess effectiveness: All children in the pilot received standardized tests of non-verbal ability and pre and post standardized tests of numeracy, reading, spelling and self esteem. Profile of Mathematical Skills (France) for mathematics and Gap (McLeod & Unwin)/Gapedol (MacLeod & Anderson) for reading, standardized tests used in the UK.

Findings: Phase I results showed significant learning gains in numeracy for pupils using SuccessMaker for six months, with an effect size of 0.4, equivalent to a twenty-month gain in six months.

Leicester University evaluators found improved results in numeracy for students in five of the eight SuccessMaker schools they studied. In six schools using GLOBAL software they found no evidence of learning gains. Note: There were system problems with this software.

In Phase II, pupils continuing to use SuccessMaker over an 18-month period continued to maintain gains against the control groups. Those pupils who made gains in Phase I and then came off the system largely sustained their mathematical gains over one year as compared with the original control group.

Pupils who spent an adequate time in SuccessMaker, used it regularly, and were appropriately supervised did better than others using SuccessMaker

In literacy, implementation factors made a clear difference. Phase I results showed no significant difference in literacy for the ILS group and control groups, probably because of the short time frame, researchers speculate. In Phase II, students achieved significant gains with the program in schools where teachers supervised and gave instructional support. In one school pupils gained 8.4 months compared with only 2.7 months by the control group. In another school, pupils made a gain of 7 months, while the control group gained less than one month over the same period.

The evaluation showed that all students benefited in proportion to their abilities and that there was no difference due to gender. Most pupils with special needs (SEN) made good progress in both mathematics and reading and showed a shift towards more independent and self-regulated learning. Some SEN pupils made accelerated progress, diminishing the gap between them and their mainstream peers.

Most pupils learning English as a second language benefited from the program (especially mathematics and their teachers reported greater independence in their learning.

Underachieving pupils made progress, with the most able and the weakest students gaining the most benefit. Disruptive behavior was reduced and teachers reported significant improvements in motivation, self-esteem and confidence in previously under-achieving pupils.

Negative comments included content that was viewed as American and some tendency for pupils to have diminished interest after several months of continual use. [Note: British, Canadian, Australian, and New Zealand versions of the mathematics software are now available and a version of the reading software now is available with British spelling and vocabulary.]

Title of evaluation: Community School District 6 [New York City Schools] Integrated Technology Reading Support Project: First, Second and Third Year Evaluation Report[s]

Author: Metis Associates, 80 Broad Street, Suite 1600, New York City, NY 10004

Relation of author to model: independent evaluator

Date evaluation report completed: October 1966, March 1998, and September 1999

Months/years covered by evaluation: School years 1995-96, 1996-97, and 1997-98.

Number of schools involved: Third grade students in 13 schools in SY 1995-96; second and third grade students in 15 schools in SY 1996-97; and second to fifth grade students in 16 elementary schools and sixth grades in 7 middle schools in SY 1997-98. Computer usage was mainly on two computer stations in each classroom.

Demographics of sample population: 95% free breakfast and lunch; 88% Hispanic, 9% African American, remaining 3% divided among other ethnic groups; 40% limited English proficient.

Purpose of study: Metis Associates was retained to evaluate the project, focusing on four questions:

- How successful is the program in increasing students' reading achievement and improving their attitudes towards reading?
- What are students' perceptions of the program?
- What are teachers' perceptions of the program?
- How successful is the professional development in enabling teachers to utilize the program?

Methodology: Results were analyzed mainly from cross-year comparisons. Results also compared with other district students not in the SuccessMaker program in year one, but in subsequent years only two district elementary schools were not in the program. Students with less than one hour of instructional time in SuccessMaker courses were removed from the study. The recommended instructional time was 25 hours, but typically few students received that amount of instruction. Students averaged 15.9 hours the first year and 16.8 hours the second year. The third year second graders averaged 22.7 hours, third graders 22 hours, and fourth graders about 11 hours. Sixth graders averaged about 7 hours in the literature-based course Reading Adventures, the course where they had the most time.

Measures/indicators used to assess effectiveness: Degrees of Reading Power (DRP), CBT Reading Test, Language Assessment Battery for ESL students in years 2-3, and teacher and student surveys. Second grade students do not usually take the reading tests. Measures in SuccessMaker courses are used for correlative purposes and for improving the next year's implementation. Foundations reading courses were used in all years. Discover English was added in year two. Reading Adventures was added in year three.

Findings: In year one, "as a whole, third grade students demonstrated dramatic improvement on the DRP. " The percentage of students at or above the state reference point for reading was higher for students participating in the SuccessMaker program than for students who were not. Student perceptions of the program were positive, agreeing with positive statements and disagreeing with negative statements about 85% or better of the time. Teacher's experiences and competencies with the computer improved substantially.

In year two, second grade was added to the project as a result of a petition from the teachers. "The results of the annual standardized tests suggest that the project was quite successful in improving the reading achievement of third grade students." "In 1997, an increase of

approximately 7 percentage points of students who scored at or above the SRP [state reference point] from the previous year.” LAB scores increased for second grade participants and third-grade participants made gains over the previous year’s third graders. Students’ levels were more closely related to test performance than time taking instruction. Students’ and teachers’ attitudes were extremely positive. The addition of a lead teacher coordinator this year made professional development and site support more effective.

In year three, students’ results were equivalent to the time spent taking instruction. The percentage of students reaching the SRP with less than ten hours of SuccessMaker instruction declined slightly, the percentage with from 10-25 hours stayed the same, and the percentage of students who reached the SRP with more than 25 hours increased by 8.5 percent. Quartile movement analysis also indicated that the majority of students were able to maintain or advance their achievement quartile from Spring 1997 to Spring 1998. LEP students with less than ten hours increased performance by 6.8 percent; those with 1-25 hours increased 13.5 percent; those with more than 25 hours increased 12.8 percent.

Teachers need additional professional development to understand the levels in the courseware and to use reporting information for diagnosis of reading needs. Overall SuccessMaker has had a positive impact on student achievement. Even more improvement could be seen with more consistent instructional time.

Title of evaluation: Meta-Analytic Studies of Findings on Computer-Based Instruction

Author: James A. Kulik (in *Technology Assessment in Education and Training*, editors Eva Baker and Harold O'Neil, Lawrence Erlbaum Associates, New Jersey 1994)

Relation of author to model: independent evaluator

Date evaluation report completed: 1994

Months/years covered by evaluation: Studies ranged from a few weeks to two years, with the most common a year-long study.

Number of schools involved (or number of classrooms/students, if appropriate): 22 separate evaluation studies reported in journals or dissertations. (Nine studies were in published journal articles.) All grade levels were represented from 1 to 12. The breakdown of number of studies at particular levels is as follows: 3 primary, 14 elementary (with 5 including primary), 4 middle school/junior high, 3 high school, one with levels not stated. Courseware was reading, language arts and mathematics.

Demographics of sample population: Does not state.

Purpose of study: "For more than ten years, my colleagues and I have been organizing and summarizing the evaluation literature to identify representative results."

Methodology: First Kulik categorized the type of software program. Kulik drew from Slavin's terminology to define the SuccessMaker program (then called Stanford-CCC for products developed by Patrick Suppes and Richard Atkinson of Stanford University at Computer Curriculum Corporation) as a level III approach, defined as having common instructional materials and training procedures. Levels I and II had less structure. Then Kulik calculated the effect size of programs of each type. "A meta-analysis is calculated by subtracting the average score of the control group from the average score of the experimental group and then dividing the remainder by the standard deviation of the measure."

Measures/indicators used to assess effectiveness: The measures used were at least one of several standardized tests for each study.

Findings: Only the [SuccessMaker] program was defined as level III. "No other program of computer-based instruction has been the object of so much scrutiny. " "The average effect size is .40; the median is .39 and the standard deviation is .23." "Thus, knowledge that a school is using the [SuccessMaker] materials allows us to make clear and accurate predictions about what to expect. Gains of 1.4 years on a grade-equivalent scale are likely with a year-long program, whereas students who are conventionally instructed would gain only 1.0 year on the same scale. Gains of nearly 2.0 years are also quite possible, whereas gains of less than 1.0 years are highly unlikely." "Based on the evaluation findings, we can state with confidence that this program produces positive effects. It will probably be a long time before we can state with equal confidence that other specific programs work equally well."