4-1-2003

Assessment of the Effectiveness of the “Success for All” Reading Program in Hartford Public Elementary Schools

Luke Forshaw

Follow this and additional works at: http://digitalrepository.trincoll.edu/theses

Recommended Citation
Forshaw, Luke, "Assessment of the Effectiveness of the “Success for All” Reading Program in Hartford Public Elementary Schools". Senior Theses, Trinity College, Hartford, CT 2003.
Trinity College Digital Repository, http://digitalrepository.trincoll.edu/theses/271
Assessment of the Effectiveness of the
“Success for All” Reading Program in Hartford Public Elementary Schools

Luke Forshaw

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Bachelor of Arts Degree

Trinity College

May 2003
Abstract

This research project investigated the effectiveness of the “Success For All” reading program in Hartford public schools. SFA is a “whole school” reading intervention program that purports to prevent literacy problems in disadvantaged public school students (Slavin & Madden, 2001). SFA has been implemented in the Hartford public school system in order to raise student achievement. This research involves a mixture of secondary analysis of archived school level data and interviews with participants involved in SFA implementation, and assessment. CMT scores were examined over years before and after SFA implementation. Significant gains around the time of SFA implementation were found, however because of the significant amount of confounds surrounding implementation, it is difficult to argue for the effects of SFA.

Despite President George W. Bush’s strong sentiment, failure is not only an option; in many school districts across the county, it is routine. Numerous school systems are riddled with low morale, low esteem and subsequently, low academic achievement (Thomas B. Fordham Foundation, 1999). In the worst of cases, students are filtered through a system, learning just enough to warrant a grade-to-grade promotion. In the best of cases, students are taught basic skills which effortlessly scaffold into advanced cognitive processes. In order for this superior education to occur, students must grasp fundamental underlying cognitive processes at a young age. A chief underlying skill, which must be continually built upon, is literacy (Thames & York, 2003). Reading achievement is paramount in establishing educable individuals.

What happens to those students who never fully develop their literacy abilities? Without literacy, is there much hope for their academic and personal futures? More importantly, are there any interventions that can promote literacy development in schools where it has traditionally failed before? The purpose of this paper is to examine the effectiveness of an intervention on reading development in disadvantaged students. However, before attempting to see if there is a program that can help, one must understand where and why an intervention would be appropriate.

If there were a standardization of wealth and resources evenly spread across America, then there would be a greater chance for equity in reading ability in our nation’s students. However, this is not the case in our country (Thomas B. Fordham Foundation,
It is in the economically disadvantaged, often urban, sections of this country in which children do poorest in school (Thomas B. Fordham Foundation, 1999). There are a plethora of reasons proposed as to why this is the case; however, three major reasons are cited as the crux of the problem.

The first problem that affects the academic achievement of urban school children has to do with poor nutrition. The international reading association simply states “Good nutrition leads to better learning.” (http://www.reading.org/publications/brochures/brochures.html). In their various publications, the international reading association affirms the intuitive by explicitly saying that there is a link between health and learning ability. Shankar and Klassen (2001) find that many urban parents are not meeting their children’s proper nutritional needs. They have identified a host of factors that explain why minority children are not receiving the sufficient quantity of proper nutrition. The major reasons their study point to are cost issues, e.g. fruits and vegetables are too expensive, and many urban parents have a lack of education surrounding the importance of fortifying their child’s diet with healthy food (Shankar & Klassen, 2001).

Another factor that affects urban children’s academic achievement is exposure to violence. Delaney-Black et al. (2002) found that there was a significant link between exposure to violence and trauma and reading deficits among urban children. In this study, children who were exposed to more community violence, and had more instances of self-reported violence, were observed to have lower IQ scores and lower standardized reading achievement test performance (Delaney-Black et al., 2002). These researchers assert that the high levels of distress and anxiety, which are associated with violence, might be one
of the reasons that exposure is associated with low reading achievement. Their findings also indicate that violence exposure, independent of trauma-related stress, is also associated with impairments in academic ability (Delaney-Black et al., 2002). Community violence and self-violence are much more likely to occur in high crime, high poverty areas, which include urban areas.

In addition to the external factors of nutrition and community violence, urban children, who are often members of minority groups, may suffer from a generalized sense of disenfranchisement, which leads to school failure (Ogbru, 1990). Ogbru argues that two types of minorities exist in America: the voluntary minority and the involuntary minority. Each group is defined by its historical passage into America, the voluntary group being immigrants who themselves, or their ancestors, moved of their own free will in order to reap greater freedom. The involuntary group is defined as cultures of people who were assimilated into America against their will, via slavery or colonization (Ogbru, 1990). Ogbru continues to state that involuntary minorities feel a greater sense of disconnect from the American majority culture, which breeds high levels of educational futility, which eventually leads to lower levels of academic achievement (Ogbru, 1990). This issue compounds itself into a vicious cycle when adult members of involuntary minority groups transfer their sense of futility onto their children. The children and the parents might share an implicit feeling of disengagement from the majority culture of American education, which would lead to continued academic failure (Brookover et al., 1979).

Clearly an intervention of some kind is needed in order to ensure that our nation’s urban school children can rise above these common pitfalls. The intervention that is currently hailed as a means to address the academic, as well as social, issues that burden
urban schools are known as “whole school” reform (Adler & Fisher, 2001). The creators say when done correctly these reform programs combine focus on academics while tackling issues involving home life, school culture, and the surrounding community in order to ensure intellectual and personal growth for all students (Adler & Fisher, 2001).

The specific intervention that will be analyzed in this paper is the Success for all Program (SFA; Slavin & Madden, 2001). This program attempts to address many of the issues urban, socioeconomically disadvantaged children are faced with. SFA claims to prevent students from developing learning problems, via intensive intervention early in elementary schools (Slavin & Madden, 2001). More specifically SFA breaks into five major components that reorganize an entire school, the first component being curriculum restructuring.

The SFA curriculum involves a systematic reading program that emphasizes phonemic awareness, vocabulary enrichment, sound blending and story telling and retelling using cooperative techniques (Slavin & Madden, 2001). These curriculum adaptations are rigorously adhered to by the use of scripting in order to achieve-system wide continuity. The second major component of SFA is eight-week reading assessments. Every student’s reading performance is tested on an eight-week rotation. The students with reading difficulties are referred to academic tutoring, and alternative teaching methods and strategies are encouraged. The third chief component is extensive tutoring. In eight-week cycles, a certified teacher, or paraprofessional, work one-on-one for twenty minutes daily. At the end of the eight weeks, these students are assessed, and if they require more assistance, they continue this cycle until they no longer require assistance. The fourth component of SFA whole school reform is its assemblage of a family support
system. Every school in an SFA school district has a family support team, which is composed of at least one parent and a social worker, and is open to whoever else is interested in joining. This team is focused on alleviating both academic and social problems, ranging from malnutrition to behavior problems. Finally, the fifth component of a SFA school is the facilitator and the advisory committee. Every school appoints a teacher to act as a facilitator who works with teachers and staff in making sure that the program is being implemented in the best possible manner. The advisory committee is a small group comprised of the school principal, facilitator teacher and parent representatives which meets on a regular basis to review the program’s progress.

SFA has been called everything from “the most promising development in education reform in the 21st century” (Slavin & Madden, 2001, p. 38) to, a “Failure” (Pogrow, 2002). In order to understand whether SFA is a success, or a failure, one must look at both sides of the argument and continue to pursue independent research. The current research is broken into two camps, the Slavin and Madden group, along with many of their subsidiaries and independents that in the spirit of academia hold Slavin and Madden’s research under scrutiny. Of these independent reviewers, Pogrow (2000,2002) and Walberg and Greenberg (1999) stand out as the program’s largest detractors.

“The results of evaluations of dozens of SFA schools in districts in all parts of the United States clearly show that the program increases student reading performance.” (Slavin & Madden, 2001, p.44). Slavin and Madden’s own research began in the developmental stages of SFA, starting in 1987 when SFA began in its inaugural school. In its successive years, SFA boomed from the reading program begun at Johns Hopkins University into a full-fledged non-profit organization, whose program is implemented in
over eleven thousand schools nationwide (Slavin & Madden, 2001). In all instances, members of the SFA team proclaim that matched against control schools, SFA schools will show an improvement in reading scores (Slavin & Madden, 2001). However the SFA foundation does not claim SFA alone to be a cure for all urban schools. They also labor to make the point that “achievement outcomes are closely related to quality of implementation.” (Slavin & Madden, 2001, p.45). If SFA is implemented fully, then these gains will be present, but if there are snags in the implementation of SFA, the effectiveness of the program will suffer. In Slavin’s own words, “No program works everywhere, and outcomes of any program depend on the quality, completeness, and appropriate application of the program. However, it would be astonishing if Success for All were not effective when fully implemented.” (Slavin, 2002, p. 46). SFA claims to maintain its effectiveness by having a program created on the pillars of research and being constantly accountable for student gains. Therefore, Slavin and Madden (2001) contend that only programs, which sustain a valid research base and evidence of their effectiveness, should continue to receive federal funding.

Independent researchers, who have carefully evaluated the results that Slavin and Madden present, represent the other side of this argument. Much of the skepticism regarding SFA has to do with the doubt that it is “possible to solve the many substantive problems facing schools in one fell, comprehensive swoop.” (Pogrow, 2000, p.3). From this primary doubt, many specific presumptions of SFA’s vulnerability have arisen.

Walberg and Greenberg (1999) initially questioned the effectiveness of SFA. They said that SFA is not adequately supported by sufficient independent research to warrant the notice of “success” it has garnered (Walberg & Greenberg, 1999). They also
assert that the positive increments in reading achievement reported by Slavin and Madden are actually the result of experimenter bias, and that Slavin and Madden simply reported what they had desired to observe (Walberg & Greenberg, 1999). Specifically, Walberg and Greenberg say that the individually administered, SFA designed tests are intended to show only results that Slavin and Madden can use to bolster their program (Walberg & Greenberg, 1999). This criticism asserts that SFA have unknown reliability and validity, when compared against more standardized, state administered measures of student growth. To this criticism, SFA supporters respond by stating that their tests measure every relevant aspect of literacy without bias (Joyce, 1999).

While Walberg and Greenberg were the initial pair of skeptics, SFA’s harshest criticism comes from Stanley Pogrow (2000, 2002). Throughout a series of articles in Phi Delta Kappan, Pogrow accuses SFA and Slavin and Madden of everything from sloppy research to self-motivated bias skewing the results. One of the chief complaints Pogrow levels against SFA has to do with the various confounds in comparing schools who use SFA to schools who do not use SFA. Pogrow argues that the schools are not matched well enough to one another, and often special education students were removed from the testing pool in SFA schools, thus seemingly bolstering the effect the program had on reading achievement scores (Pogrow, 2002).

Pogrow also states that in Slavin and Madden’s research, they do not do an adequate job of parsing out extraneous causal variables, which might have affected the reading achievement scores they report (Pogrow, 2002). For example, Pogrow cites SFA’s claim of credit for Texas minority students’ score increase on the Texas Assessment of Academic Skills (TAAS). Pogrow points out that upon further inspection,
there were substantial reductions in class size, along with a vast test-preparation effort, which he asserts were the actual cause of the rise in TAAS scores, not necessarily SFA (Pogrow, 2002).

Pogrow’s criticisms of Slavin and Madden’s research have made current research efforts much more aware of the potential confounds that exist when trying to draw inferences of causation from a school intervention program. In any non-experimental or quasi-experimental study, there exist a host of confounds which must be controlled for as well as possible.

In the case of school assessment, there must first be a standardized measure of achievement that is equivalent across schools. In this study, the unit of analysis is a single school, and the standardized measure of achievement is the fourth grade reading score on the Connecticut Mastery Test (CMT). The CMT is a test administered across the state of Connecticut, at grades four, six and eight. Because of SFA’s emphasis on early intervention, it is thought that most, if any effects of the program will be seen in the fourth grade CMT scores. This logic is derived from the fact that SFA asserts that it is an early intervention program, and thus results from this intervention would be seen in the fourth grade.

Another major confound that can be controlled for is student demographics from school to school. These demographics consist of various characteristics, such as student body size, level of poverty, and grade configuration of school. However, all students in these three districts in Connecticut all take the CMT in October of each year.

There exists a set of extraneous potential confounds, that are much more difficult to control for in a non-experimental design. The first of such factors is teacher and
administrator turnover rates. If a school is a revolving door for staff, that might in itself create some effect on the reading achievement scores on the CMT. Another extraneous factor that will be difficult to control for is the level of funding available at different schools. That is to say, if one school receives much more funding than its comparison school, the ability to definitely pronounce SFA as having or not having the effect on reading scores becomes much more difficult. Finally, the level of SFA training administered in a school is also a confound that must be addressed. Slavin and Madden themselves preach about the importance of solid implementation in order for SFA to thrive, hence if one school has had more extensive training for the program, it is not of equivalent status to the school which has had less training.

A series of district specific challenge exist in this analysis of Hartford Public School. Hartford is a large urban district in central Connecticut. Hartford Public Schools currently have twenty-four thousand-four hundred and seventy-nine students enrolled in one of their thirty-three schools. (http://www.hartfordschools.org/about_us/facts.html). In 2002 only seventeen percent of students in grade four in Hartford Public Schools achieved at or above state goal for reading on the CMT compared to the state average of fifty-two percent. Hartford is a district in need of measures to help raise the reading achievement levels of their entire student population (http://www.state.ct.us/sde/c23_press_release.pdf). In response to this need for improvement, the Hartford Public School district adopted the Success for All program in the summer of 1999. All but one school agreed with SFA implementation. Aside from this one school, all other schools in the district are currently using SFA.
A challenge in trying to measure effectiveness of SFA in Hartford public schools is that there are a limited number of schools that suffice as control comparison units. Hartford public schools are overwhelmingly using the SFA program, meaning that there are few schools to act control, non-SFA schools. For this reason, comparison schools will be schools located in other, highly similar Connecticut cities (Bridgeport and New Haven) that share similar school demographics. Schools in Hartford will be compared to analogous schools in New Haven and Bridgeport in order to have a control against SFA.

A final confound that will be addressed is Hartford’s creation of the “Amato Plan”. Anthony Amato was appointed superintendent of Hartford Public Schools in April 1999 (Chedekel, 1999). He The “Amato Plan” is named after the one-time superintendent of Hartford Public Schools, and it is a series of programs meant to cope with the common pitfalls of urban education. The grandest initiative brought on, SFA was implemented and teachers were trained in August of 1999. The central component of this program is SFA, however, there also exist other programs that might be impacting on student reading achievement scores. Some other aspects of this plan included a standardization of all basal readers in the fourth grade. Prior to 1999, students were reading a variety of different reading textbooks, but in 1999, this curriculum became standardized across the district. Another major component of this plan was a larger emphasis placed on reading non-fiction texts. A large component of the Connecticut Mastery Test expects students to read and understand pieces of non-fiction. Prior to Anthony Amato’s introduction in 1999, many schools were placing a tremendous emphasis on fiction reading, but after 1999 began incorporating more non-fiction text into their reading instruction. In order to
make any definite claims regarding the effectiveness of SFA, these other aspects of the “Amato Plan” must be controlled for.

The overall effectiveness of SFA is highly suspect. However, theoretically SFA accounts for and attempts to deal with many of the problems that plague urban schools. It has a strong parent – teacher link, in its family support team, which attempts to deal with issues of nutrition, community violence and increasing the sense of belonging in school culture. The SFA program’s attempt to provide extensive one-on-one instruction when needed combined with a strong, standardized reading curriculum might also help in raising overall student achievement. However, because of the highly tumultuous and political nature of Hartford public schools, I hypothesize that SFA will have no significant effect on raising reading scores. Hartford public schools have a high faculty turnover rate, which most likely impedes the implementation process (Green, 1999). Policy changes and support in Hartford public schools are highly connected to changes in political infrastructure, meaning that the constant power struggles in the school system often mean discontinuity in a program. With out the proper support for implementation, any of the claims of success the Slavin and Madden state will not have a chance (Slavin & Madden, 2001). If SFA were implemented to its highest capacity, I would still be suspect in believing that it would raise reading achievement across an entire district. Certainly, when the program is not implemented fully, it will have a significantly smaller chance of raising reading achievement scores.

The goal of this paper is to examine the question of if there is a single program that can intervene on enough levels to produce substantial change in urban schools. By using Hartford Public schools as the backdrop, this paper hopes to understand the
limitations and successes of whole-school reform in an urban school system. More specifically, this thesis hopes to shed light directly on the status of Hartford Public schools, and to validate or refute the claims of victory that SFA supporters make.
Method

Participants

This study used a quasi-experimental design to look at the effectiveness of SFA in raising reading achievement. Eighty-three schools over three districts (Bridgeport, Hartford and New Haven) were involved in this study. The districts of Bridgeport and New Haven were used as comparison, control districts because they do not use SFA. Additionally, these districts are classified in the same Educational Reference Group (ERG) as Hartford (http://www.csde.state.ct.us/public/der/ssp/terms.pdf).

Materials

Archived data published by the Connecticut State Department of Education over the six-year period of 1996 – 2001 was used to create an extensive database. The Connecticut State Department of Education and the Connecticut State Library provided extensive assistance in identifying and acquiring the appropriate measures of school level data needed for this analysis. The results of the Connecticut Mastery Test (CMT) for reading achievement were used as the dependent variable to assess reading comprehension.

Several interviews with officials from the Connecticut State Department of Education, Hartford Public Schools and the Connecticut State Library provided data and focused the variables I controlled for. These interviews also provided a host of data regarding Hartford Public School district initiatives in curriculum and in possible confounds in the data gathered.
Results

Looking at the eighty-three schools, over the six-year period, many significant findings were discovered. The first repeated-measures MANOVA found year effects for the percentage of students on or above fourth grade reading on the Connecticut Mastery Test, $F(5,66) = 2.73, p <.05$. In a follow up analysis, a paired-sample t-test revealed the increase from 1998 to 1999 to be significant $T,(79) = -2.41, p <.05$ (See Figure 1).

The years were then aggregated into two, three-year clusters. The first cluster was made up of the years prior to the introduction of SFA (1996, 1997 and 1998) and the second cluster was made up of the three years after SFA introduction (1999, 2000, 2001). After making this adjustment, a main effect for years before and after SFA introduction became evident, $F(1,78) = 15.56, p <.001$. Additionally a main effect of year by district also proved to be significant, $F(2,78) = 4.99, p <.05$ (See Figure 2).

The next step was to look at the effects of year and district on district wide percentages of limited English proficiency (LEP). There was a significant effect of year on LEP, $F,(5,65) = 2.32, p <.05$. There was also a significant effect of district on LEP, $F,(2) = 6.99, p <.05$. Of this district difference, Hartford was significantly higher than either Bridgeport or New Haven (See Figure 3).

This finding led to the next analysis. In order to find the effects of year (broken into pre and post SFA clusters) on grade four reading CMT, while controlling for LEP, a repeated measures ANOVA with LEP as a covariate was performed. The results of this analysis shows a substantial improvement in CMT by district after 1999, even when LEP is controlled for, $F,(2) = 9.56, p <.05$ (See Figure 4).
An analysis of the grade six CMT scores by district provided highly similar trends to those seen in the fourth grade analysis. Additionally, attempts to control for poverty were also made, however the measure of school-level poverty was too unreliable to be used as a covariate (See Figure 5). When the trends of poverty rates were initially examined, there were suspiciously large shifts over time. After speaking with people from the State Department of Education, it was explained that reporting poverty is highly variable. Apparently, standards for reporting poverty levels in schools varied considerable from year to year from school to school.
Discussion

There is a sudden increase of seven percent in fourth grade reading CMT scores from 1998 to 1999 among Hartford students. This finding seemingly supports SFA’s claims of effectiveness in raising reading proficiency. However, SFA was just introduced as a district initiative in the summer of 1999, and the CMT was administered in October of 1999. It is highly unlikely that one month of exposure to SFA is enough to boost scores in such a dramatic fashion. Furthermore, after this initial jump in 1999, the average reading CMT scores dip and then level off. One would expect to see the results of a new district initiative to begin slowly, and increase over time. The data does not support this assumption. The improvement is consistent with SFA, but no sustained growth is surprising.

Over this six-year period, the New Haven trend in CMT reading scores looks relatively stable. However, the Bridgeport CMT reading scores, similar to those of Hartford show gains. It is important to note that Bridgeport, a district that does not use SFA also showed gains in fourth grade reading CMT scores over this six-year period.

One important factor to note is that this research is a field study with many different confounds. It is almost impossible to disentangle SFA from other initiatives made by the Hartford Public School system. One key confound here is that HPS systematically realigned its reading curriculum with the CMT at the same time SFA was implemented. Specifically, HPS standardized its basal reader series across the district. At the same time, HPS developed a more CMT friendly reading curriculum by increasing the number of non-fiction texts for students to read. As a result of these confounds, there
is no real way to know which weight to give to each initiative in the raising of CMT scores.

In this analysis we wanted to control for student poverty levels, but discovered the indicator of student poverty is unreliable so we did not include it. The means for determining student poverty is variable from district to district and in some cases, from school to school. This issue speaks to a larger problem which one experiences when using this type of research model. In relying on data that is created by someone else, you rely on their definitions. One problem occurs when these definitions change over a given period of time. If in one year, eligibility for free and reduced priced lunch meant that your parents had to make less than sixteen thousand dollars a year and the next year it means that your parents have to make less than ten thousand dollars a year, it is hard to ascertain a sense of consistency. Also, by relying on some archived data that was gather by someone else, you end up relying on someone else’s criteria for assessing what ever you are examining. That is to say, if you consider poverty to be making less than thirty thousand dollars a year, yet the source of your data has defined poverty as making less than sixteen thousand dollars a year, your hopes for evaluation and what the data actually evaluate are two very different things. For these very reasons, adjustments for poverty we not included in the final analysis of the data.

One additional piece of data, which would have helped provide a more definitive assessment of the Success For All program, were the quality of implementation scores. These scores are generated by SFA evaluators, who assess how well each school has implemented the program. I requested information regarding degrees of implementation for SFA from Hartford Public Schools. However, as a result of no response from the
school system, the implementation data was not included in this study. This data would have helped to separate schools that have implemented SFA “well” from those who have not. This control would have either strengthened the affect of SFA or would have made it seem less important. However, without the direct assessment of degree of implementation, I am not in a strong position to argue for the effects of SFA.

The implications of this research are far reaching. As a result of federal, “No Child Left Behind” legislation it has become even more apparent the emphasis our society places on student literacy achievement

http://www.ed.gov/offices/OESE/esea/execsumm.html. “No Child Left Behind” also impacts “Failing” schools in a number of financial and managerial ways. All of the districts evaluated in this study would certainly be classified as “failing” and whole school reform efforts, like SFA, would be put into place in order to raise test scores. Whole school reform efforts like SFA proliferate when more districts are in need of restructuring. This issue also speaks largely to the broader question of “How much importance should be placed on standardized tests”. It is important to attempt to have an objective view as to whether or not a program of this nature actually does result in student gains on these standardized tests.
References


Electronic brochure regarding proper childhood nutrition. (2002). Retrieved August, 7


http://www.hartfordschools.org/about_us/facts.html

State of Connecticut Department of Education press release discussing CMT scores in the

case study of beating the odds. The Reading Teacher, 54, 6, 616-619

Social Systems and Student Achievement: Schools Can Make a Difference,
New York: Praeger.
Chedekel L. (1999) A Driven Man, A Dire Setting

The Hartford Courant

Green, R. (1999) Teaching With Heart

The Hartford Courant


ERIC Digest, ED429988


Acknowledgements

I would like to thank and acknowledge first and foremost my advisor, David Reuman of the Psychology Department at Trinity College. Without his willingness, enthusiasm and support, I would not have been able to take this project on. I would also like to thank those who helped in the initial development of this research, specifically, Jack Dougherty, Barbara Henriques and Dina Anselmi. I also thank those students who have given me support in this project; Ashley Bunn, Emily Doerr, Alex Blanchard and Christian Sterling. Much of the direction for where to look and what to examine came from Barbara Beaudin from the State Department of Education. Other members from the State Department of Education who offered considerable help were; Marcus Rivera, Michelle Leon and Thomas Murphy. Additionally, Marta Diez from Hartford Public Schools helped in developing my understanding of the Amato Plan. Finally I would like to acknowledge and thank Lee and Valerie Forshaw and Carole and Carolyn DeLuca for their constant support and encouragement throughout this entire process.
Figure 2

Unweighted mean of grade 4 CMT score by district by year (1998-1999, 1999-2001)
Figure 3
Figure 5

Percentage of Students Eligible for Free or Reduced Price Lunch


Averages by Districts for Poverty

- New Haven
- Bridgeport
- Hartford