

**Abstract Title Page**  
*Not included in page count.*

**Title:** Understanding variation in implementation of SFA in the i3 scale-up project

**Authors and Affiliations:**

Janet Quint, MDRC  
Pei Zhu, MDRC  
Fred Doolittle, MDRC

[janet.quint@mdrc.org](mailto:janet.quint@mdrc.org)  
[pei.zhu@mdrc.org](mailto:pei.zhu@mdrc.org)  
[fred.doolittle@mdrc.org](mailto:fred.doolittle@mdrc.org)

## **Abstract Body**

*Limit 4 pages single-spaced.*

### **Background / Context:**

Enhancing the reading performance of at-risk elementary students is one of the thorniest problems in American education. In 2009, fourth-grade scores on the National Assessment of Educational Progress were only slightly higher than they were in 1992 (NCES, 2010), and in fact have changed little since 1980. Further, there are substantial gaps in reading proficiency according to social class and ethnicity. Among fourth graders not eligible for free lunch, 45% scored at or above proficient, in comparison to only 17% among fourth graders eligible for free lunch. The percent proficient rates were 42% for white fourth graders, but only 16% for African Americans, 17% for Hispanics, and 20% for American Indians.

Success for All (SFA), a whole-school turnaround model for elementary schools that focuses primarily on ensuring that every child succeeds in learning to read throughout the elementary grades (Slavin, Madden, Chambers, & Haxby, 2009)., received one of four Investing in Innovation (i3) scale-up grants. Its goal is to help hundreds of struggling schools by reaching directly into the heart of practice -- the interactions between teachers and students -- to improve daily lessons and school functioning.

### **Purpose / Objective / Research Question / Focus of Study:**

The main confirmatory research question guiding the design of the impact evaluation is: *What is the impact of SFA on elementary school students' reading achievement, compared to students in non-SFA schools?* An answer to this question will determine our assessment of whether SFA is successful at turning around low-performing schools. In addition to the main confirmatory question, this evaluation will address exploratory questions such as subgroup impacts and impact on non-cognitive outcomes. These analyses are intended to deepen our understanding of the overall average impact of SFA.

The implementation study has three principal goals: to measure the fidelity with which the program model is put in place (and to determine whether the identity of the SFA coach affects fidelity), to assess the treatment-control contrast in educational experience, and to document the implementation process and lessons for scale-up and replication.

### **Setting:**

During the 2010-11 school year, the SFA team recruited 37 Title I schools from five school districts located in the Southwestern, Mid-Atlantic, and South Central regions of the country to participate in the evaluation.

### **Population / Participants / Subjects:**

The 37 schools in the evaluation have been randomly assigned to treatment (n = 19) or control (n = 18) conditions. About 3,000 kindergarten students in these schools were assessed on reading

skills in the fall of 2<sup>nd</sup> the 2011-12 school year; they will then be followed for three years, with assessments taking place each spring.

### **Intervention / Program / Practice:**

The main elements of SFA are:

- School staff buy-in (usually indicated by a faculty vote).
- Coaches who work with administrators and teachers to implement SFA.
- The presence at each school of an SFA facilitator who helps all teachers with program implementation, ongoing professional development, and school-wide assessments.
- Extensive professional development for all school staff to help them understand and use research-proven approaches to reading instruction, cooperative learning, classroom management, motivation, teaching of metacognitive skills, and assessment.
- A K-6 reading program that uses extensive cooperative learning in pairs and small groups to build phonemic awareness, phonics, comprehension, vocabulary, and fluency.
- Frequent criterion-referenced, instruction-based formative assessments to make sure that all students are on track toward success, along with quarterly benchmark assessments to track progress in meeting grade-level expectations.
- One-to-one or small-group tutoring for students who are found to be falling behind grade-level expectations.
- A Solutions Team, which works to prevent or solve problems that go beyond academics.
- Leadership development that engages the principal and school leadership team in a continuous improvement process based on data analysis, goal -setting, and achievement monitoring.
- Implementation self-assessment checklists for teachers and school leaders and implementation benchmarks completed by coaches, which provide data for monitoring the quality of program implementation and formative outcomes.

### **Significance / Novelty of study:**

The study goes beyond previous evaluations of SFA in two key respects. First, it brings to bear a wide variety of implementation data that will enable us to draw more informed conclusions about how implementation takes place and the extent to which SFA produces changes in classroom practice. Second, the study tests the potential of a lower-cost version of SFA in which local district coaches rather than coaches employed by SFA assume responsibility for working with participating schools. The study will examine whether high-quality implementation is maintained with the use of these local personnel, an issue with important implications for policy and practice since the lower-cost strategy makes the intervention more readily affordable to economically hard-pressed school districts.

### **Statistical, Measurement, or Econometric Model:**

Our basic impact estimate will be a two-level HLM model with students nested in schools. Students cannot be nested in classrooms, as students are regrouped every quarter and may have several reading teachers in a year. Blocking will account for any stratification in the school lotteries should districts request this. Covariates in the impact model will include key student

characteristics, such as percentages of ELL, special education, and free/reduced price lunch students and baseline student reading achievement test scores. This model will provide an intent-to-treat estimate of the impact of providing access to the intervention on students in the average school in the sample.

### **Usefulness / Applicability of Method:**

The methodology, which relies on randomization of schools, is justified by the whole-school nature of the SFA reform. Preliminary data indicate the baseline equivalence of students at T and C schools on measures of early literacy.

### **Research Design:**

In the 37 T and C schools, we will examine two groups of students:

- Fall 2011 kindergarten students in the randomly assigned schools will comprise the *primary study sample*, so designated because they will have been exposed to SFA from the beginning of their formal educational experience. They will be followed for three years through the end of the 2013-14 school year, when they should have completed second grade. Since the analysis focuses on the schools in the sample, we will not follow students who move away from their original study school, but will include “in-movers” who join the target grades over time. We will collect annual data on the composition of students in both the treatment and control schools to check for any unexpected effects on student mobility and, if there are none, we will also be able to examine impacts for a “stable sample” of students who remain in the SFA and control schools over time.
- Students in third, fourth, and fifth grades will constitute an *auxiliary study sample*. We consider this to be an auxiliary sample because students will have been exposed to SFA but will not have received it since the beginning of their school experience. They would be assessed in the spring of the 2012-13 and the spring of the 2013-2014 school year, at which point students in these grades would have had two and three years of exposure to the program, respectively.

### **Data Collection and Analysis:**

The Impact Study: The primary student outcome is students’ achievement in reading. For the *primary sample*, we plan to use different developmentally-appropriate *measures of reading achievement*. In the fall of 2011, kindergarten students were individually pretested on the Peabody Picture Vocabulary Test (PPVT) and on Woodcock-Johnson III Letter-Word Identification. Their scores will be used as covariates in the impact analysis. In the spring of 2012, we will field individually-administered follow-up tests using the Woodcock Letter-Word Identification and Word Attack scales. In the spring of 2013 (when students are completing first grade) and 2014 (when they are completing second grade), we will field individually-administered Woodcock Letter-Word Identification, Word Attack, and Passage Comprehension assessments, and the DIBELS, a reading fluency measure.

Two group-administered measures of reading achievement would be taken for the *auxiliary sample*. The first is the Gates-MacGinitie Reading Test, which provides a nuanced indicator of

students' reading levels and is therefore proposed as the primary measure. Because of the policy importance of state assessments, we will also make arrangements with the study districts to obtain state reading test data for older-grade students and will treat these as a secondary outcome to see if our findings for older students based on the Gates-MacGinitie are confirmed using another measure. To deal with the variation in tests across states we will place the different tests on the same metric by converting them to z scores, as suggested by May (2009). Students' third-grade scores on these measures would be used as covariates in the analysis of their fourth-grade and fifth-grade scores.

The Implementation Study: Data from multiple sources will inform the implementation study. The School Achievement Snapshot is the key source of information on *implementation fidelity*. The Snapshot, which will be completed for each school at the conclusion of the school year by the school's SFA coach, contains 40 items tapping the presence or absence of schoolwide structures associated with the program model, along with 20 items describing classroom processes (a construct that includes both instructional practices and student engagement). We will further ascertain whether implementation fidelity is affected by whether the SFA coach is an employee of SFA or an employee of the school district trained by SFA.

The Snapshot will also suggest questions to be asked in surveys of principals and teachers at treatment and control schools in order to understand the *contrast in treatment* that implementing SFA entails. The surveys should provide information about the presence or absence in the control schools of structures and processes akin to those that SFA schools and teachers are expected to adopt.

Teacher logs will constitute a further source of information on differences in instructional practices at the two sets of schools. We plan to administer a log developed by Brian Rowan and his colleagues at the University of Michigan that has been shown in prior research to differentiate effectively between instruction in SFA schools and in schools that adopted two other special reading programs (as well as schools where no special reading intervention was in place). Logs will be collected from first- and second-grade reading teachers in each of the 37 study schools in the spring of 2012, 2013, and 2014, with an expected sample of approximately 48 logs per school.

The principal and teacher surveys will foster our understanding of the *implementation process*. In addition, MDRC researchers plan to conduct site visits to the program schools, in the course of which they will interview the principal and teachers providing reading instruction to elicit their perspectives on SFA and its implementation, the support school personnel received, challenges that arose, and responses that were developed to address them.

### **Findings / Results:**

The evaluation is ongoing and at this point we do not have any findings

### **Conclusions:**

Not applicable, given the absence, at this juncture, of impact and implementation findings.

## **Appendices**

*Not included in page count.*

### **Appendix A. References**

*References are to be in APA version 6 format.*

**Appendix B. Tables and Figures**  
*Not included in page count.*