

July 2005

The Challenge of Scaling Up Educational Reform Findings and Lessons from First Things First

Final Report

Janet Quint, Howard S. Bloom, Alison Rebeck Black, and LaFleur Stephens

with

Theresa M. Akey

Introduction

This report on First Things First — a major comprehensive school reform — arrives at an opportune moment, when President George W. Bush, the nation’s governors, and business and foundation leaders have announced a renewed commitment to reforming American high schools. Now operating in more than 70 schools in nine districts across the country, First Things First (FTF) seeks to improve low-performing schools by strengthening relationships between teachers and students and by making classes more engaging and rigorous. FTF was initially launched in Kansas City, Kansas, and subsequently tested in 12 middle schools and high schools in four additional districts through the Scaling Up First Things First Demonstration, a five-year research and demonstration project supported by the Institute of Education Sciences in the U.S. Department of Education. The project was a collaboration of two organizations: the Institute for Research and Reform in Education (IRRE), which developed the program model and provided support and technical assistance to partner schools and districts, and MDRC, which evaluated the initiative.

This report, the last of four produced by MDRC, describes the implementation and effects of the program model in these five districts, all serving high proportions of minority and economically disadvantaged students. It complements and updates a report on the FTF program in Kansas City, Kansas.¹ In summary, the key findings of this study are:

- Middle and high school students in Kansas City, Kansas, registered large gains that were sustained over several years and were pervasive across the district’s schools; similar gains were not present in the most comparable schools in the state. The improvements occurred over the course of eight years of substantial

¹Michelle Gambone, Adena Klein, Jean Summers, Theresa Akey, and Cynthia Sipe, *Turning the Tide: The Achievements of the First Things First Education Reform in the Kansas City, Kansas, Public School District* (Philadelphia: Youth Development Strategies, Inc., 2004).

effort by the school district and IRRE to implement FTF as the district’s central educational reform.

- It is not yet clear whether the expansion sites, which had operated FTF for two or three years at the time of the research follow-up, will replicate the impressive findings for Kansas City.

What Is First Things First?

FTF entails major changes in school structure, instruction, and accountability and governance policies. The model includes three components:

- **Small learning communities.** In this initiative, small learning communities (SLCs) contain groups of up to 350 students and their core-subject and other key teachers who remain together for several years. They are organized around broad themes (for example, “Science and Technology” and “Performing Arts”) that are meant to inform instruction and provide the SLCs with unique identities.
- **Family Advocate System.** Each student is paired with a staff member — generally a teacher in the student’s SLC — who is expected to meet regularly with the student and monitor his or her academic, social, and emotional progress. The advocate is responsible for assisting the student, creating a more positive relationship between the school and the student’s family, and working with parents to promote their child’s academic success.
- **Instructional improvement efforts.** Teachers work with their colleagues to align curricula with state and local standards, and they participate in professional development activities designed to help them learn, practice, and regularly use strategies that make classroom instruction rigorous and engaging.

How Was FTF Evaluated?

To assess program implementation, the report draws on a combination of quantitative data from teacher and student surveys and qualitative findings from classroom observations and interviews with administrators, teachers, students, and others. While a random assignment design — considered the “gold standard” for evaluating program impacts — was not feasible for this study, MDRC used a rigorous research method, called a “comparative interrupted time-series analysis,” to estimate the effect of FTF. In principle, the impact of FTF on a student outcome equals the *difference* between what that outcome was after the school reform was under way and what it would have been without the reform (the “counterfactual”). In practice, one can estimate this difference by comparing the change over time in a student outcome for schools that adopted FTF with the corresponding change for similar schools that did not adopt the reform, and variants of this approach were used for each of the five sites in the evaluation. Ideally, the evaluation de-

sign that is used to produce impact estimates should comprise data on consistently measured student outcomes for multiple pre-intervention baseline years, multiple post-intervention follow-up years, multiple FTF schools, and multiple comparison schools that are closely matched with the FTF schools. Unfortunately, this set of ideal conditions did not exist in any of the study sites; instead, the evaluators had to make the best of the data that were available, while remaining mindful of the limitations of the resulting analyses.

How Was FTF Implemented?

FTF began operations in Kansas City, Kansas, in 1998-1999 in one of the district's four comprehensive high schools, along with that high school's feeder middle and elementary schools. The three remaining comprehensive high schools and their feeder schools were added to the program over the next two years. The "expansion," or "scaling-up," sites discussed in this report were phased in over a two-year period, beginning in 2001-2002. They include three high schools and four middle schools in Houston, Texas; a high school and its two feeder middle schools in the Riverview Gardens School District in suburban St. Louis County, Missouri; and two high schools in Greenville and Shaw, Mississippi.

FTF is a complex reform whose implementation requires change at every level. It demands much both of personnel in the schools and districts mounting the reform and of the staff of IRRE, who are responsible for guiding and assisting local efforts. The following key findings emerge from MDRC's analysis of the initiative's implementation in these various sites.

- **FTF has evolved continuously not only at the implementation sites but in the minds of its developers, as IRRE personnel have learned from both successes and challenges.**

For example, the Family Advocate System was not added to the mix of program elements until 2000-2001. IRRE's role in providing technical assistance in the area of instructional improvement increased considerably, and that assistance became more comprehensive and systematic over time.

- **Predictably, changes in structure took hold more quickly and more easily than changes in instruction; the instructional improvement efforts associated with FTF were implemented most fully in Kansas City, Kansas.**

While the creation of small learning communities was relatively easy and popular among teachers and students alike, changing teachers' instructional practices proved challenging. Central office support for instructional improvement in Kansas City helped ensure progress in this area. Over time, teachers at the expansion sites followed their Kansas City counterparts, moving forward in aligning curricula and assessments with state standards and in making greater use of active engagement strategies in their lessons.

- **District and school leadership and outside technical assistance were the key determinants of implementation success at the expansion sites.**

Implementation progressed further in settings where district and school leaders provided consistent support for the initiative, where the principal and School Improvement Facilitator (a school district employee working at each school to guide implementation of the reform) had a cooperative and mutually respectful relationship, and where IRRE staff offered intensive technical assistance.

Did FTF Make a Difference for Student Outcomes?

MDRC looked at the impact of FTF both in its original site — Kansas City, Kansas — and in the later, scaling-up districts. The key findings follow.

- **In Kansas City, Kansas, high school and middle school academic outcomes improved substantially as FTF was implemented, while similar trends were not observed in comparison schools, pointing to the initiative’s central role in improving academic performance.**

These academic outcomes included increased rates of student attendance and graduation, reduced student dropout rates, and improved student performance on the Kansas state tests of reading and mathematics. As Table ES.1 shows, the estimated effects on student test scores reflect double-digit increases in the percentage of students who scored at levels deemed “proficient” or above by the state and double-digit reductions in the percentage of students who scored at levels deemed “unsatisfactory.”

For example, on the most recent state reading test (for spring 2004), FTF high schools experienced an 11.1 point relative gain in the percentage of student scores that were proficient or above. In other words, the increase in the percentage from its initial level three years earlier was 11.1 points greater for FTF high schools than for their comparison schools. Even larger relative improvements were observed for the percentage of student scores that were unsatisfactory. In spring 2004, this percentage had dropped by 15.5 points more for FTF high schools than for comparison schools. Findings for FTF middle schools indicate a relative increase of 13.7 points in the percentage of student scores that were proficient or above and a relative decline of 13.6 points in the percentage of scores that were unsatisfactory. Thus, overall, there was a pronounced and consistent pattern of relative improvements in the reading performance of FTF high school students and middle school students.

Findings for math in Table ES.1 reveal correspondingly large and consistent improvements for middle schools. By spring 2004, FTF middle schools had experienced a 9.6 point increase (relative to their comparison schools) in the percentage of math scores that were proficient or above and a 9.0 point relative decrease in the percentage of scores that were unsatisfactory. Math scores for FTF high schools also showed signs of improvement (with relative declines

The First Things First Evaluation

Table ES.1

**Estimated Impact of First Things First on Student Test Scores:
Kansas City, Kansas**

	Spring 2002 ^a	Spring 2003	Spring 2004
Impact on Percentage Proficient^b			
<u>High schools</u>			
11th-grade reading test	6.9	10.2 **	11.1 **
10th-grade math test	1.2	3.4	-4.4 *
<u>Middle schools</u>			
8th-grade reading test	3.0	23.1 ***	13.7 ***
7th-grade math test	5.0	11.0 ***	9.6 **
Impact on Percentage Unsatisfactory^b			
<u>High schools</u>			
11th-grade reading test	-5.4	-11.1 **	-15.5 ***
10th-grade math test	-10.8 ***	-6.7 **	-5.2
<u>Middle schools</u>			
8th-grade reading test	-5.4	-22.3 ***	-13.6 ***
7th-grade math test	-7.3 *	-13.1 ***	-9.0 **

NOTES: Sample includes students from four FTF high schools and eight FTF middle schools.

The "impact" was calculated as the difference between the change from the quasi-baseline level for FTF schools in spring 2001 and the corresponding change over time for comparison schools. Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

^aSpring 2002 is the fourth year of implementation for three schools, the third year of implementation for three schools, and the second year of implementation for six schools.

^b"Proficient" is defined as the sum of the top three (of five) performance categories on the Kansas state test. "Unsatisfactory" is defined as the bottom performance category on the Kansas state test. Improved student performance is represented by a relative increase in the percentage proficient and a relative decrease in the percentage unsatisfactory.

in the percentage of scores that were unsatisfactory). However, there was no clear pattern over time in changes in the percentage of scores that were proficient or above.

These positive effects — plus those for other high school and middle school outcomes — reflect improvements that, in many cases, were sustained over time and occurred at numerous schools. Therefore, despite the inability to select comparison schools that closely matched Kansas City's exceptionally low-performing schools (an issue that is discussed further in the report), the multiplicity and magnitude of the improvements that were observed for FTF schools in this district support the conclusion that the reform model was critical to causing them.

- **It is not yet clear whether the expansion sites will replicate the robust findings for Kansas City.**

Findings for the FTF replication districts are less pronounced or less consistent than those for the reform model's original district. As Table ES.2 illustrates, estimates of impacts on state test scores are almost never as large as those for Kansas City; they vary markedly across districts; and they are seldom statistically significant. The lack of statistical significance reflects, in part, the small number of schools in the replication study from each district and, thus, the limited statistical precision of the study to detect impacts at these schools.

One of the largest urban high schools where implementation was most complete has registered positive effects on student achievement. There are some suggestive signs of success at other schools as well; however, the overall pattern of findings leaves considerable uncertainty about how much improvement in student performance was produced by the reform.

It is important to note several limitations of the analyses presented in this report, which make them a conservative test of the program's effects. First, as noted, the statistical precision for measuring impacts at a single school or a few schools (as is the case in most of the expansion sites) is often too limited to identify with confidence impacts other than those that are exceptionally large. Second, in the current educational environment, there are strong pressures on all schools to improve, so that the outcomes of FTF are measured against those of comparison schools that may also be trying to change. Third, because expansion sites began their efforts recently, there is only a brief window of time through which to view their success. Finally, because at some sites the benchmark used to gauge improvement is a "quasi-baseline" year after implementation had already begun, any impacts produced before or during this year are "netted out" of the analysis and thus are not attributed to the initiative.

What Are the Policy Implications of This Study?

This report tells a complex story about a complicated initiative. The implementation findings indicate that mounting the intervention is hard; doing it well requires commitment, persistence, and effort. The positive effects of FTF in its home district of Kansas City, Kansas, were sizable, pervasive, and sustained. So far, the schools participating in the scaling-up demonstration, with one exception, have not registered similar effects.

What does this say about FTF — and about school reform efforts more generally? The experience in Kansas City, Kansas, points to four conditions that were *sufficient* to produce

The First Things First Evaluation

Table ES.2

Estimated Impact of First Things First on Student Test Scores:
Replication Sites

	Follow-Up Year 1	Follow-Up Year 2	Follow-Up Year 3
Impact on Percentage Passing^b			
Houston, Texas^a			
<u>High schools</u>			
10th-grade reading test	-1.1	6.6	8.8 *
10th-grade math test	-3.3	4.2	7.0
<u>Middle schools</u>			
8th-grade reading test	-1.6	-5.1 **	-1.9
8th-grade math test	2.5	1.8	6.5
Delta Region of Mississippi			
<u>High schools</u>			
10th-grade English test	12.9	8.4	
9th-grade algebra test (Shaw High School)	-10.0	-3.3	
9th-grade algebra test (Greenville-Weston High School)	15.8	-15.6	
Impact on Percentage of Low Performers^c			
Riverview Gardens, Missouri			
<u>High school</u>			
11th-grade communication arts test	5.6	-7.1	1.5
10th-grade math test	-2.8	-7.6	-10.0
<u>Middle schools</u>			
7th-grade communication arts test	3.7	6.5	0.0
8th-grade math test	-4.5	-9.5	-7.1

NOTES: Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

The "impact" was calculated as the difference between the change from the baseline or quasi-baseline level for FTF schools and the corresponding change over time for comparison schools.

^aIn follow-up Years 1 and 2, impacts presented are the average of impacts for three high schools and four middle schools. In follow-up Year 3, impacts presented are for one high school and one middle school.

^bImproved student performance is represented by a relative increase in the percentage passing.

^cLow-performing students are those scoring in the bottom two proficiency categories established by the State of Missouri. Improved student performance is represented by a relative decrease in the percentage of low performers.

meaningful impacts on a wide array of outcomes in secondary schools serving disadvantaged populations. Whether these conditions are also *necessary* remains an open question.

1. A districtwide focus, with the district's staying the course for many years in its provision of pressure and supports for the reform's changes

From the outset, the Kansas City, Kansas, school district took ownership of FTF as its major school reform initiative. In contrast, at the scaling-up sites, the school districts did not provide similarly consistent support and oversight to the schools mounting the initiative. The

experience of the successful Houston high school suggests that lack of strong district support may be offset if exceptionally strong school-level leadership is in place.

2. Schools that had operated FTF for many years when their impacts were measured

The findings of this report are consistent with other research indicating that comprehensive school reforms are more effective when they have been in place for at least five years. At the time of the research follow-up, FTF had been in operation for a much longer period in Kansas City than in the expansion sites. Even the latest-starting Kansas City schools had been in operation for four years when the last impact data were collected. On the other hand, the follow-up data at the expansion sites reflect a maximum of only two or three years of experience operating the intervention.

3. Balancing a need for more personalized learning environments with a comprehensive and intensive approach to improving instruction that emphasizes alignment, rigor, and student engagement

The FTF experience suggests that striking such a balance is not easy. For instance, the expansion schools were able to implement small learning communities quickly, during the demonstration's planning year. By all accounts, they enabled teachers and students to develop closer relationships with each other. In contrast, early instructional improvement efforts in the scaling-up sites were much less systematic than in Kansas City.

4. Intensive and responsive technical assistance from providers who are willing to make midcourse adjustments where needed

IRRE was on the scene in Kansas City, Kansas, throughout the planning period and early implementation years of FTF in the district. But IRRE's capacity was stretched by the addition of so many new sites at once and by the subsequent expansion into two more large urban districts that are not part of this report.

The experiences of operating FTF in Kansas City, Kansas, and in the expansion sites suggest that school reform is too difficult to expect results without long-term support from high-level leaders and without sufficient technical assistance. The success of the model in Kansas City points to the critical role that districts play in providing a unified message and the pressure and support that all educators need to keep their eyes on the prize: better teacher-student relationships and improved teaching and learning in the classroom.