

# Appendix

## Appendix A1 Study characteristics: Quint, Bloom, Black, & Stephens, 2005—Houston study (quasi-experimental design)

Characteristic	Description
<b>Study citation</b>	Quint, J., Bloom, H. S., Black, A. R., & Stephens, L. (2005). <i>Scaling up First Things First: The challenge of scaling up educational reform</i> . New York, NY: MDRC.
<b>Participants</b>	<p>The study examined the performance of students attending schools in the Houston Independent School District. The main analysis sample included students from three high schools and four middle schools implementing <i>First Things First</i>, as well as students from a set of matched comparison schools. The middle school study did not examine outcomes relevant to WWC dropout prevention reviews. Therefore, the results in this intervention report pertain only to the high school analysis.</p> <p>The high school sample consists of all students enrolled in the study schools during the study period: three years prior to <i>First Things First</i> implementation and one year after implementation. This included a total of 7,891 high school students in the <i>First Things First</i> schools. The study authors did not report the number of students in comparison high schools. Comparison schools in the Houston school district were matched to each <i>First Things First</i> high school based on overall student performance at the schools on standardized achievement tests during the baseline period. High schools in the district whose mean combined reading and math scores on the Stanford Achievement Test, ninth edition (SAT-9), were within 0.25 standard deviations of the mean for the <i>First Things First</i> school were selected as comparison schools. This methodology resulted in 10 comparison schools for 2 of the <i>First Things First</i> high schools and 11 comparison schools for the other. A high school could be chosen as the comparison school for more than one <i>First Things First</i> school, and this often occurred. A total of 13 Houston high schools served as comparison schools for the <i>First Things First</i> Houston study.</p> <p>The high school study examined two cohorts of students. Cohort 1 consisted of one intervention high school that implemented <i>First Things First</i> in 2001 and its matched comparison schools. Cohort 2 consisted of two intervention high schools that began implementing <i>First Things First</i> in 2002 and their comparison schools. The WWC used results for cohorts 1 and 2 combined to rate the effectiveness of <i>First Things First</i>.</p> <p>In the years prior to the implementation of <i>First Things First</i>, the three <i>First Things First</i> schools and their comparison high schools had, on average, similar attendance and promotion rates and served students who had similar tenth-grade passing rates on math and reading tests. The study authors indicated that the <i>First Things First</i> high schools in Houston were heavily Hispanic, while other low-achieving high schools in Houston were heavily African-American. Therefore, it was not possible to match high schools closely on both their racial and ethnic composition and their student performance.</p>
<b>Setting</b>	<i>First Things First</i> was implemented in three high schools in the Houston Independent School District: Lee High School, Sam Houston High School, and Sharpstown High School. During the study period, these three high schools served student populations that were primarily Hispanic. Lee High School began implementing <i>First Things First</i> in the fall of the 2001/02 school year, and Sam Houston and Sharpstown began implementing the program the following school year, 2002/03.

(continued)

**Appendix A1 Study characteristics: Quint, Bloom, Black, & Stephens, 2005—Houston study (quasi-experimental design) (continued)**

<b>Characteristic</b>	<b>Description</b>
<b>Intervention condition</b>	<p><i>First Things First</i> targeted three high schools in Houston with low-achievement levels. Each of these schools implemented the key features of the three components of the whole school reform:</p> <p><i>Small learning communities.</i> <i>First Things First</i> reorganizes schools into small learning communities of up to 350 students and their teachers. The small learning communities in Houston were focused around a central theme (for example, technology), and students remained in the same communities throughout high school. <i>First Things First</i> also recommends that schools reduce student-teacher ratios in math and language arts classes to increase the amount of individualized attention that students receive. Class sizes in the Houston study schools were reduced from an average of 26 students to 20 students between the planning year and the second implementation year.</p> <p><i>Family and student advocate system.</i> Advocates met one-on-one with students weekly and contacted the students' families at least monthly to discuss their academic and personal progress. In at least one Houston high school the advocates became counselors for the students, helping them to schedule the classes they needed to progress toward graduation. Consistent with the model's guidelines, advocates in Houston high schools met weekly with their students as a group. These group sessions were held as either a daily "homeroom" period or a weekly class meeting.</p> <p><i>Instructional changes and supports.</i> <i>First Things First</i> contracted with Kagan Cooperative Learning, Inc., to train teachers on cooperative learning strategies intended to complement the small learning communities. It also called for increased instructional time for math and language arts courses.</p>
<b>Comparison</b>	<p>Matched comparison schools were Houston high schools that did not implement <i>First Things First</i>. The study identified comparison schools from the districts that were similar in average performance on standardized achievement tests in the three years preceding program implementation. High schools in the district whose mean scores on the Stanford Achievement Test, ninth edition (SAT-9), for reading and math combined were within 0.25 standard deviations of the mean for the <i>First Things First</i> school were selected as comparison schools.</p>
<b>Primary outcomes and measurement</b>	<p>The outcome of interest from this study was a measure of the percentage of ninth graders who attended at least one day of school during the following year, referred to as the "one-year persistence rate." The study also examined <i>First Things First's</i> effects on attendance and standardized test scores. These outcomes do not fall within the three domains examined by the WWC's review of dropout prevention interventions (staying in school, progressing in school, completing school) and are not included in this report.</p>
<b>Staff training</b>	<p>Teachers at <i>First Things First</i> high schools were regular teachers employed by the Houston Independent School District. As part of the instructional changes, teachers in <i>First Things First</i> schools were trained in cooperative learning strategies by Kagan Cooperative Learning. Each school employed a <i>First Things First</i> director to serve as a liaison between the district and the <i>First Things First</i> schools. Staff from IRRE provided ongoing professional development for teachers in the implementation of the family advocate system.</p>

## Appendix A2 Outcome measures in the staying in school domain

Outcome measure	Description
<b>One-year persistence rate</b>	The percentage of ninth-graders who attended at least one day at a district school at any point during the following school year. Students who were recorded in district records as having transferred to another district at the end of the ninth grade are coded as having persisted in school.

## Appendix A3 Summary of study findings included in the rating for the staying in school domain<sup>1</sup>

Outcome measure	Study sample	Sample size (schools) <sup>2</sup>	Authors' findings from the study					
			Mean outcome		WWC calculations			
			<i>First Things First</i> group	Comparison group <sup>3</sup>	Mean difference <sup>4</sup>	Effect size <sup>5</sup>	Statistical significance <sup>6</sup> (at $\alpha = 0.05$ )	Improvement index <sup>7</sup>
<b>Quint, Bloom, Black, &amp; Stephens, 2005—Houston study (quasi-experimental design)<sup>8</sup></b>								
One-year persistence rate	Ninth graders (cohorts 1 and 2)	16	76.3	77.3	–1.0	–0.03	ns	–1
<b>Domain average for staying in school<sup>9</sup></b>						–0.03	ns	–1

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the improvement index. The one-year persistence rate was measured during the second year of implementation of *First Things First* in Houston.
2. Quint, Bloom, Black, & Stephens (2005) used individual student data for their analysis; however, the relevant outcomes are school-level measures. In the study, each block of schools consists of a *First Things First* school matched with a group of 10 or 11 comparison schools. A school could serve as the comparison school for more than one *First Things First* school. The WWC confirmed with the study authors that the total number of unique comparison schools was 13.
3. Quint, Bloom, Black, & Stephens (2005) reported only baseline persistence rates and the difference between the baseline to follow-up changes of the intervention and comparison groups. The study authors provided the WWC with the follow-up means for both groups. The WWC generated the adjusted comparison group means reported here using the following transformation: adjusted comparison group mean = follow-up comparison group mean + (baseline intervention group mean – baseline comparison group). Stated differently, the adjusted comparison group mean equals the follow-up intervention group mean minus the impact, since, under the comparative interrupted time-series technique used in the Quint, Bloom, Black, & Stephens (2005) study, impacts are calculated as: impact = (follow-up intervention group mean – baseline intervention group mean) – (follow-up comparison group mean – baseline comparison group mean).
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of effect size calculation, please see the [Technical Details of WWC-Conducted Computations](#). The effect size for the dichotomous variable “one-year persistence rate” was computed using the Cox Index.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between groups. The level of statistical significance was reported by the study authors.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
8. The level of statistical significance was reported by the study authors, or where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the [WWC Tutorial on Mismatch](#). For the formulas the WWC used to calculate the statistical significance, see [Technical Details of WWC-Conducted Computations](#). In the case of the Houston *First Things First* study, the study authors provided details of their two-level analysis model, which adjusted for clustering within the school, and thus no additional corrections for clustering were necessary.
9. This row provides the study average, which in this instance, is the same as the single measure for the persistence rate.

## Appendix A4 *First Things First* rating for the staying in school domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.<sup>1</sup>

For the outcome domain of staying in school, the WWC rated *First Things First* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects because it had only one study, and that study showed no statistically significant or substantively important outcomes, either positive or negative, in this domain.

For explanations of how *First Things First* fared on the criteria for these ratings, see below.

### Rating received

**No discernible effects:** No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

**Met.** The single study of *First Things First* showed no statistically significant or substantively important effects in this domain.

### Other ratings considered

**Positive effects:** Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

**Not met.** No study of *First Things First* showed a statistically significant positive effect in this domain.

### AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

**Met.** No study of *First Things First* showed statistically significant or substantively important negative effects in this domain.

**Potentially positive effects:** Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

**Not met.** No study of *First Things First* showed a statistically significant or substantively important positive effect in this domain.

### AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

**Met.** No study of *First Things First* showed statistically significant or substantively important negative effects in this domain.

(continued)

## Appendix A4 *First Things First* rating for the staying in school domain (continued)

**Mixed effects:** Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

**Not met.** No study of *First Things First* showed a statistically significant or substantively important effect, either positive or negative, in this domain.

**OR**

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important *important* effect.

**Not met.** No study of *First Things First* showed a statistically significant or substantively important effect in this domain.

**Potentially negative effects:** Evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

**Not met.** No study of *First Things First* showed a statistically significant or substantively important negative effect in this domain.

**AND**

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

**Met.** No study of *First Things First* showed statistically significant or substantively important positive effects in this domain.

**Negative effects:** Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

**Not met.** No study of *First Things First* showed a statistically significant negative effect in this domain.

**AND**

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

**Met.** No study of *First Things First* showed statistically significant or substantively important positive effects in this domain.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the [WWC Intervention Rating Scheme](#).

## Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence <sup>1</sup>
		Schools	Students	
Staying in school	1	16	nr	small
Progressing in school	0	na	na	na
Completing school	0	na	na	na

na = not applicable or not studied

nr = not reported

1. A rating of “moderate to large” requires at least two studies and two schools across studies in one domain, and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is “small.”