

2007–2008 FIRST THINGS FIRST EVALUATION: YEAR ONE IMPLEMENTATION



Austin Independent School District  
Department of Program Evaluation

November 2008  
Publication Number 07.79

## **EXECUTIVE SUMMARY**

Austin Independent School District (AISD) has undertaken a transformation of secondary education across the school district. The Office of Redesign was established to facilitate and support improvement across all middle and high schools in the district. In this effort, Office of Redesign staff worked collaboratively with district high schools and national experts to develop systems and program implementation plans that will enable the district to build its internal capacity to address deep-seated challenges to student success. The High School Redesign Initiative focused on four major goals:

- Closing achievement gaps between all student groups
- Increasing 4-year high school completion rates for all students
- Ensuring that all high school graduates are well prepared for college and career success
- Increasing the college and career readiness rates of English language learners (ELLs)

To attain these major goals, Office of Redesign staff supported campus staff by assisting them in making structural changes in their schools and in implementing instructional improvement systems based on research-based reform models. These structural changes and support systems included the following:

- Dividing large comprehensive schools into smaller learning communities (SLCs) for students
- Creating and implementing a student advisory/family advocacy program within every high school
- Establishing and facilitating professional learning communities (PLCs) for teachers
- Constructing a secondary literacy model for ELLs
- Improving the teaching and learning of mathematics in all AISD high schools
- Providing intensive technical assistance and support for Johnston High School to create the conditions for instructional improvement that can lead to maximum student performance

These structural changes and support systems provide high schools with the tools needed to prepare all students for graduation and college. One model of school restructuring adopted by AISD is the Institute for Research and Reform in Education's (IRRE) First Things First (FTF) initiative at the LBJ, Reagan, and Travis high schools. For the 2007–2008 school year, the program evaluation of the FTF initiative assessed the following outcome measures: (a) self-reported student attitudes and evaluations of their family advocate, (b) math and reading Texas Assessment of Knowledge and Skills (TAKS) performance in FTF and non-FTF campuses, (c) disciplinary referral patterns across FTF and non-FTF campuses, and (d) results

from the engagement, alignment, and rigor (EAR) observation protocols during the 2007–2008 school year. This evaluation generated several key findings:

- The responsibility for conducting EAR protocol visits was not evenly shared by staff within and outside campuses. Often, only a handful of observers on campuses accounted for the majority of classroom visits. Distributing observation duties across members on and outside campuses will bolster the reliability of the findings generated by the data collection tool.
- As the school year progressed, a greater percentage of classrooms met the EAR thresholds. However, among classrooms selected for observation, fewer than 30% satisfied the rigor requirement.
- Overall, approximately 85% of students surveyed at FTF schools could identify their family advocate.
- More than half of the students surveyed at LBJ and Travis indicated they had not participated in a conference with their parents/guardians and family advocate, whereas only 22% of Reagan students reported no conferences had taken place.
- One-third of students at FTF campuses reported having met with their family advocate at least once a week, outside of formally scheduled class periods.
- Students who felt more comfortable discussing personal issues with their family advocate also were more likely to interact frequently with their advocate.
- The percentages of students at FTF campuses who met the passing standard on the math TAKS increased sharply and significantly from 2006–2007 to 2007–2008. However, these gains were not symmetric across ethnic groups, nor did they remain when controlling for student- and school-level characteristics.
- Both FTF and FTF-comparison campuses experienced dramatic declines in the percentages of students with discipline referrals and in the per student discipline referral rate from 2006–2007 to 2007–2008. At FTF campuses, high student attrition, particularly among students who had chronic disciplinary problems and who were predisposed to dropping out, may have contributed to these declines.
- After controlling for student and staff characteristics, the decrease in the discipline referral rate at FTF campuses was not measurably greater than the declines at FTF-comparison campuses.

## TABLE OF CONTENTS

Executive Summary .....	i
List of Figures .....	iv
List of Tables .....	v
Program Overview .....	1
Methods .....	2
<i>Evaluation Objective</i> .....	2
<i>Data Sources and Research Design</i> .....	2
Results.....	4
<i>Description of the SLCs on FTF Campuses</i> .....	4
<i>Teacher Perceptions of Professional Learning in their SLCs</i> .....	4
<i>Summary of FTF Engagement, Alignment, and Rigor Protocol Observations</i> .....	5
<i>Summary of FTF’s Family Advocacy Implementation in 2007–2008</i> .....	11
<i>The Impact of FTF on TAKS Performance</i> .....	16
<i>The Impact of FTF on Disciplinary Referrals</i> .....	24
Discussion.....	33
Conclusion and Recommendations.....	36
References.....	38
Appendix A: FTF Student Survey Profile and Sampling Methodology.....	39
Appendix B: Selection of Comparison Campuses.....	40
Appendix C: Technical Material for TAKS Multivariate Analyses Section .....	41
Appendix D: Technical Material for the Disciplinary Analysis Section .....	42

**LIST OF FIGURES**

Figure 1: Monthly EAR Protocol Observations, by Campus, 2007–2008 .....	7
Figure 2: EAR Protocol Observations, by Campus, 2007–2008 .....	8
Figure 3: Classrooms Meeting EAR Thresholds, by Campus, 2007–2008.....	9
Figure 4: Classrooms Meeting EAR Thresholds, by Campus and SLC, 2007–2008.....	10
Figure 5: Students Reporting They Knew Their Family Advocate, by Campus, 2007–2008	13
Figure 6: Students Reporting an Advocate Met With a Parent/Guardian, by Campus and Frequency of Meeting, 2007–2008.....	13
Figure 7: Students Reporting an Advocate Met With a Parent/Guardian, by Campus and Grade Level, 2007–2008 .....	14
Figure 8: Student Reports of Frequency of One-on-One Contact With an Advocate, by Campus, 2007–2008 .....	15
Figure 9: Student Reports About Whether They Felt Comfortable Talking With a Family Advocate .....	16
Figure 10: Students Meeting Math and Reading TAKS Standards, by Campus, 2006–2007 and 2007–2008 .....	18
Figure 11: Students Meeting Math and Reading TAKS Standards, by FTF Status, 2006–2007 and 2007–2008.....	19
Figure 12: Students Meeting Math TAKS Standards, by FTF Status and Ethnicity, 2006–2007 and 2007–2008.....	20
Figure 13: Students Meeting Math TAKS Standards, by FTF Status, Ethnicity, and Economic Disadvantage Status, 2006–2007 and 2007–2008 .....	21
Figure 14: Predicted Probability of Earning a Proficient Score on Math or Reading TAKS Between 2006–2007 and 2007–2008, by FTF Status .....	23
Figure 16: Students With at Least One Referral, by Campus and FTF Status, 2006–2007 and 2007–2008 .....	26
Figure 17: Students With Multiple Referrals, by Campus and FTF Status, 2006–2007 and 2007–2008 .....	27
Figure 18: Relationship Between Staff Perceptions of Campus Safety and Percentage of Students With Multiple Discipline Referrals, 2006–2007 and 2007–2008 .....	28
Figure 19: Per Student Discipline Referral Rate, by Campus, Cohort, and FTF Status, 2006– 2007 and 2007–2008.....	30
Figure 20: Predicted Discipline Referrals, by FTF Status, 2006–2007 and 2007–2008 .....	31
Figure 21: Predicted Discipline Referrals, by FTF Status and Student and School Characteristics, 2006–2007 and 2007–2008.....	33

**LIST OF TABLES**

Table 1: SLC Enrollment Data, 2007–2008 ..... 4  
Table 2: TAKS Performance and Economically Disadvantaged Status, by FTF Status, 2006–  
2007 and 2007–2008..... 17

## **PROGRAM OVERVIEW**

Austin Independent School District (AISD) has undertaken a transformation of secondary education across the school district. The Office of Redesign was established to facilitate and support improvement across all middle and high schools in the district. In this effort, Office of Redesign staff worked collaboratively with district high schools and national experts to develop systems and program implementation plans that will enable the district to build its internal capacity to address deep-seated challenges to student success. The High School Redesign Initiative focused on four major goals:

- Closing achievement gaps between all student groups
- Increasing 4-year high school completion rates for all students
- Ensuring that all high school graduates are well prepared for college and career success
- Increasing the college and career readiness rates of English language learners (ELLs)

To attain these major goals, Office of Redesign staff supported campus staff by assisting them in making structural changes in their schools and in implementing instructional improvement systems based on research-based reform models. These structural changes and support systems included the following:

- Dividing large comprehensive schools into smaller learning communities (SLCs) for students
- Creating and implementing a student advisory/family advocacy program within every high school
- Establishing and facilitating professional learning communities (PLCs) for teachers
- Constructing a secondary literacy model for ELLs
- Improving the teaching and learning of mathematics in all AISD high schools
- Providing intensive technical assistance and support for Johnston High School to create the conditions for instructional improvement that can lead to maximum student performance

One model of school restructuring adopted by AISD is the Institute for Research and Reform in Education's (IRRE) First Things First (FTF) initiative. This model was used at the LBJ, Reagan, and Travis high schools. Grounded in a research-based reform framework, FTF strives to transform the traditional setting of large, comprehensive schools into an environment of smaller sub-groupings of students (i.e., SLCs) organized around career-oriented themes. Furthermore, students are assigned a family advocate who monitors their academic performance; regularly contacts family members to communicate student information and to encourage parental involvement in school and at home; and builds strong, durable relationships with students. Finally, the FTF model promotes instructional improvement within schools by

ensuring students are engaged in class, the curricula are aligned with state and district standards, and the classroom instructional strategies are rigorous. These interconnected goals were monitored through frequent classroom observations using the engagement, alignment, and rigor (EAR) protocol, a rubric that allowed instructional coaches, school administrators, and district staff to identify and quantify the hallmarks of a successful classroom, according to the FTF initiative.

## **METHODS**

### **EVALUATION OBJECTIVE**

The Department of Program Evaluation (DPE) staff provided information for decision makers about program participation and outcomes to facilitate decisions about program implementation and improvement.

### **DATA SOURCES AND RESEARCH DESIGN**

The evaluation of FTF examined four student outcomes: (a) math and reading Texas Assessment of Knowledge and Skills (TAKS) performance in FTF and non-FTF campuses, (b) self-reported student attitudes and evaluations of their family advocate, (c) disciplinary referral patterns across FTF and non-FTF campuses, and (d) results from the EAR observation protocols during the 2007–2008 school year.

Data for these outcome variables were collected from various district sources. Math and reading TAKS performance data were pulled from the district records for the 2007 and 2008 test administrations. For students tested more than once, the highest score for a particular subject test and grade was taken. Only the scores for students who were classified as active according to district records were analyzed; thus, the student-level data were unbalanced due to student attrition.

In Spring 2008, students at the three FTF campuses (i.e., Reagan, Travis, and LBJ) were surveyed using an instrument developed by IRRE. The entire populations of students at FTF campuses were surveyed. Of the 2,970 students actively enrolled during Spring 2008, valid surveys were completed by 1,085 students, yielding a response rate of 36.5%. Response rates varied by campus (see Appendix Table A 1).

Disciplinary referral data were taken from district data sources. Referrals were collapsed down to the individual student level for a given school year. Thus, unique incidents for a student in a particular school year were summed, providing an interval measure of the total number of disciplinary referrals a student received during the 2006–2007 and 2007–2008 school years. Importantly, the change in the severity of offenses that prompted referral was not analyzed due to formidable data collection challenges. Last, data from the EAR observations were extracted to analyze patterns in EAR protocol performance throughout the school year and across FTF campuses.

Whether a particular school was chosen as a site to implement FTF was not determined randomly. Thus, comparing student outcomes across FTF and non-FTF schools would generate unreliable inferences and undermine the validity of the recommendations stemming from the quantitative analyses because much of the variance in student performance across these schools could be attributable to underlying student characteristics within these campuses (e.g., the percentage of economically disadvantaged students enrolled, the pervasiveness of disciplinary problems, the ethnic composition of the school, and students' past scores on standardized tests). To avoid this pitfall, the research design adopted for this evaluation was quasi-experimental, whereby the outcomes of FTF schools over time were compared with outcomes for schools with comparable enrollment profiles. As Quint, Bloom, Black, Stephens, and Akey (2005) noted, these comparison schools represent the "counterfactual" scenario, allowing us to isolate the programmatic impact of FTF assignment on student outcomes across time. Because of the similarity between the FTF-comparison schools and the schools assigned the FTF intervention, gains in measured student outcomes at the FTF schools that were greater than those at the comparison schools can be attributed to FTF implementation. The selection process for FTF-comparison schools is discussed in a forthcoming section.

DPE staff interviewed a sample of teachers from the FTF high schools in May and June of 2008. An invitation to participate in an interview or focus group was sent to randomly selected teachers at each campus. Focus groups and interviews were conducted based on the teachers' preference and availability. The interviews and focus groups were designed to allow us to understand and describe the context in which the FTF initiative was implemented. Across the three FTF high schools, 23 teachers participated in an interview or focus group.

Several data analysis techniques were utilized for the evaluation. Basic comparative descriptive statistics were generated to provide a simplistic description of changes in student outcomes across time. Furthermore, multivariate inferential statistical techniques were used to estimate the effect of FTF participation on a range of outcome variables, while controlling for a host of student-level characteristics. Details concerning the statistical justifications for each estimation procedure are provided in the sections describing the results of these analyses.

## RESULTS

### DESCRIPTION OF THE SLCs ON FTF CAMPUSES

One imperative structural reform prescribed by FTF was the reorganization of large, comprehensive high schools into interdisciplinary, thematically interrelated SLCs. The establishment of SLCs within high schools was an attempt to counteract student anonymity and disconnection that is thought to be a pervasive problem in comprehensive high schools with large student enrollments (Connell, 2002). Shrinking the learning environment within which students and teacher function facilitates the creation of strong, durable bonds between students and instructors. As a result of these improved connections, according to Connell (2002), “Students, families, and staff will form more personal, long-standing and mutually accountable relationships around their work” (p. 6). To satisfy this FTF requirement, FTF schools were tasked with restructuring their schools into thematically organized SLCs, with enrollments both below 300 and roughly equally distributed. Table 1 provides an overview of the SLC enrollment patterns at Reagan, Travis, and LBJ during the 2007–2008 school year.

Table 1: SLC Enrollment Data, 2007–2008

Campus	SLC	Enrollment
Reagan	Art of Learning	284
	Leadership, Engineering, Architecture, and Design	293
	Medical, Arts, Science, and Health	288
Travis	Visual/Performing Arts	287
	International Hospitality, Business	247
	Law, Criminal Justice, Government	278
	Media, Engineering, Technology	271
LBJ	Health/Science	290
	Performing and Visual Arts	273
	Technology and Business	294
	Health Sciences and Law	297

*Source.* ASTU

*Note.* This table only includes active students as of the last day of the 2007–2008 school year.

### TEACHER PERCEPTIONS OF PROFESSIONAL LEARNING IN THEIR SLCs

One of the major support systems for instructional improvement is the SLC. Each FTF school provided a common planning time during which teachers representing different content areas came together to focus on improving their instruction. Instructional improvement activities included classroom visits within and across disciplines, dialogue about the observations, examination of student work, and peer review of lesson plans. This time also allowed teachers to share strategies that were successful in their classrooms and to get ideas for improving their instruction. It was a forum for student data review, target setting, and action

plan development. We expected the results of the teachers' professional learning time to be evident in observed instructional practices and student outcomes.

In the teacher focus groups conducted in Spring 2008, teachers discussed the professional learning taking place within their SLCs and identified practices that benefited their instructional improvement. Teachers appreciated opportunities to share their lesson plans and to provide recommendations about strategies and content. Teachers reported the peer review was helpful and constructive with respect to improving their practice. They liked learning about the different content areas and learning from the different practices. They also reported that observing their peers provided them with ideas for their own classes and models to emulate. They expressed hopes that these practices would increase rigor in their individual classrooms and inter- and intra-departmental collaboration.

Teachers also expressed some concerns about the activities taking place in their SLCs. They reported little instruction from FTF staff about how the SLC time was to be used and few suggestions to help them use that time effectively. Teachers attended the SLC meetings because they were mandatory; however, they did not necessarily buy in to the idea of sharing and working across content areas/disciplines, thus preventing high levels of participation within the time set aside for professional learning. For example, one teacher said,

There is a lot of resistance. Many people do not bring their lesson books in or share what they are doing in our group. They come and listen and then leave without participating. They do not use the time effectively.

Additionally, teachers worried that the time spent in their SLCs infringed on their personal planning time and on classroom instruction time. For example, another teacher expressed her frustration:

This time takes away from the preparation for your own planning and classroom instruction time. I have four preps, and the group takes my lunchtime and conference time. Even on the days that I have time during lunch for preparation, there are other teacher groups meeting in my space.

Finally, some teachers reported they did not meet with or participate in their SLCs. Some of these teachers reported their assigned SLC met at a time when they were teaching. Other teachers, who taught specialized electives, did not meet with their SLCs because they did not feel the group addressed or supported their professional learning needs.

### **SUMMARY OF FTF ENGAGEMENT, ALIGNMENT, AND RIGOR PROTOCOL OBSERVATIONS**

In addition to the structural changes mandated by the FTF school reform initiative, FTF prescribes three overarching instructional goals designed to improve instruction at participating campuses:

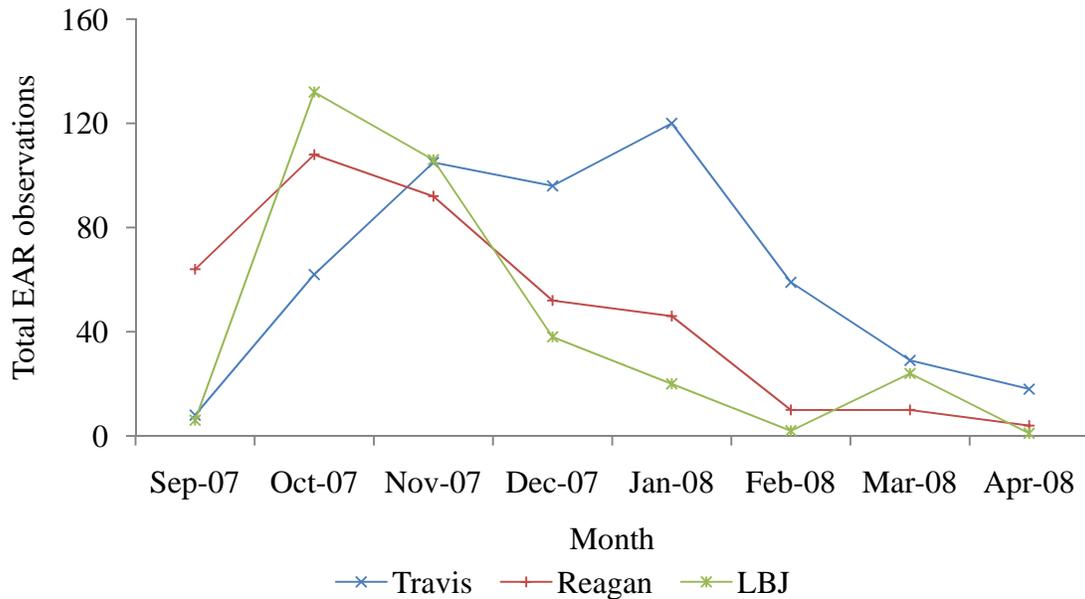
- Engagement: “Students being actively involved—emotionally, behaviorally, and cognitively—in their academic work”
- Alignment: “Students being asked to do and doing work that reflects academic standards deemed to be important by their district and state and having opportunities to master the methods used on their state’s high stakes assessments”
- Rigor: “Reflects the common sense notion that students will only achieve at high levels if that level of work is expected and inspected for all students”

To support these instructional improvements, district, campus, and external staff engaged in a rigorous series of training sessions called Measuring What Matters focused on assessing classroom instruction and student learning. Trained individuals periodically assessed classroom instruction and student learning at FTF schools throughout the school year. Classrooms were observed using the EAR protocol developed by IRRE. The instrument provided the observer with a detailed rubric containing a series of prompts asking observers to describe the degree of student engagement, curricular alignment, and academic rigor in the classroom. Observers determined whether the observed classroom met the required threshold for each of the components.

Frequent classroom observations were encouraged to ensure an accurate depiction of each classroom’s progress in meeting the instructional goals of the campus. Peer-led professional development sessions then were tailored to meet support needs determined through results from the EAR observation protocol.

The frequencies of classroom visits were highest for Reagan and LBJ during the first two months of the school year (Figure 1). As the school year progressed, these two schools showed a decline in classroom observations. However, the third school, Travis, was the most consistent in conducting its observations, and the number of classroom visits continued to increase at Travis through January 2008. Travis also had more observations taking place during the last months of the school year, compared with Reagan and LBJ.

Figure 1: Monthly EAR Protocol Observations, by Campus, 2007–2008



Source. IRRE, November 2008

Note. Classroom observations spanning October 1, 2007 to May 1, 2008 are included.

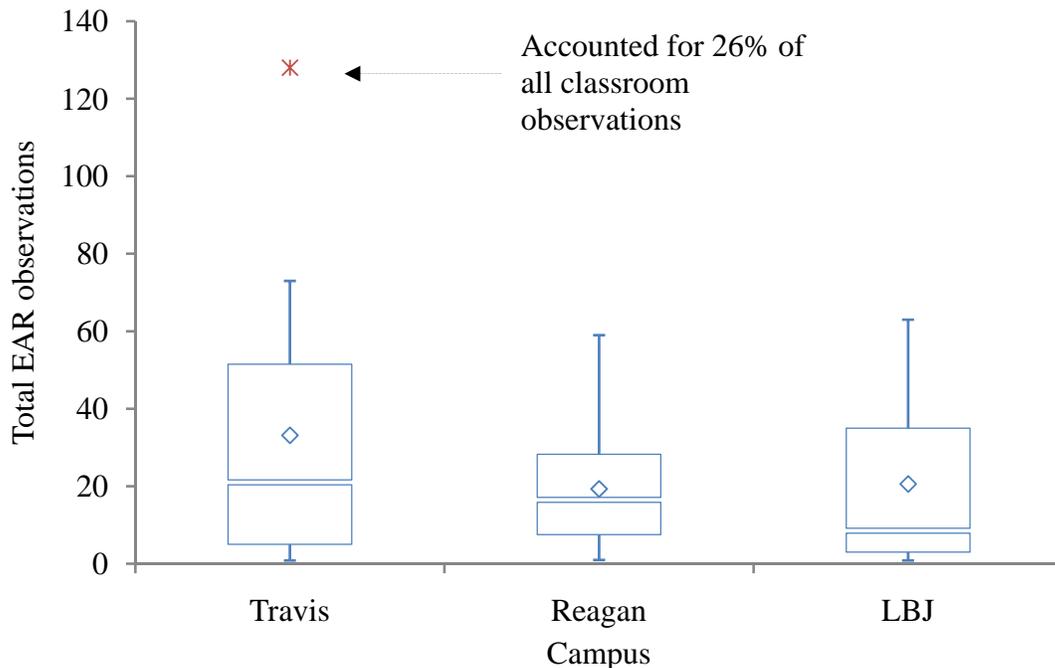
A variety of persons observed the FTF classrooms, including district curriculum and program administrators, campus-level instructional specialists, and campus-level administrators. Broad participation is encouraged by the MWM framework to avoid the introduction of systematic and chronic biases in EAR protocol completion. To assess fidelity to this objective, classroom observations from the 2007-2008 school year were disaggregated by campus and illustrated on a boxplot conveying several key pieces of information (Figure 2). For our purposes, we are interested in extreme characteristics in the data, particularly the existence of a few individual observers conducting a large proportion of the total number of classroom visits, in addition to determining the average number of visits per observer.

#### ***How do I interpret Figure 2?***

One method used to assess how broadly the duty of conducting classroom visits was shared by campus and district staff was to identify whether individual staff members conducted an inordinate portion of total observations. To visualize this, a boxplot was created which displays several key pieces of information, including the median (designated by the white space separating the boxes) and mean number (denoted by the blue diamond) of classroom visits per observer, and whether any individual observers were responsible for an inordinate share of classroom visits (denoted by a red asterisk). Simplistically, the more compressed boxes indicate classroom visits were shared more evenly among campus staff (i.e., a smaller difference between the 25<sup>th</sup> and 75<sup>th</sup> percentile values) where 50% of all observations lie. The taller boxes indicate the presence of extreme values (i.e., a larger greater dispersion between the 25<sup>th</sup> and 75<sup>th</sup> percentile values) or a wider difference between observations within this range.

Campus EAR protocol data show differing patterns of data collection by campus and district staff. For instance, at Reagan, four observers accounted for 43% of all classroom observations. At Travis, however, the four classroom visitors with the largest share of observations were responsible for 64% of the observations while one observer accounted for approximately one-quarter of all classroom observations (asterisk).<sup>1</sup> Ensuring that campus observations are collected widely by campus and district staff will bolster the reliability of the findings generated by the data collection process.

Figure 2: EAR Protocol Observations, by Campus, 2007–2008



Source. IRRE, November 2008

Note. Classroom observations spanning October 1, 2007 to May 1, 2008 are included.

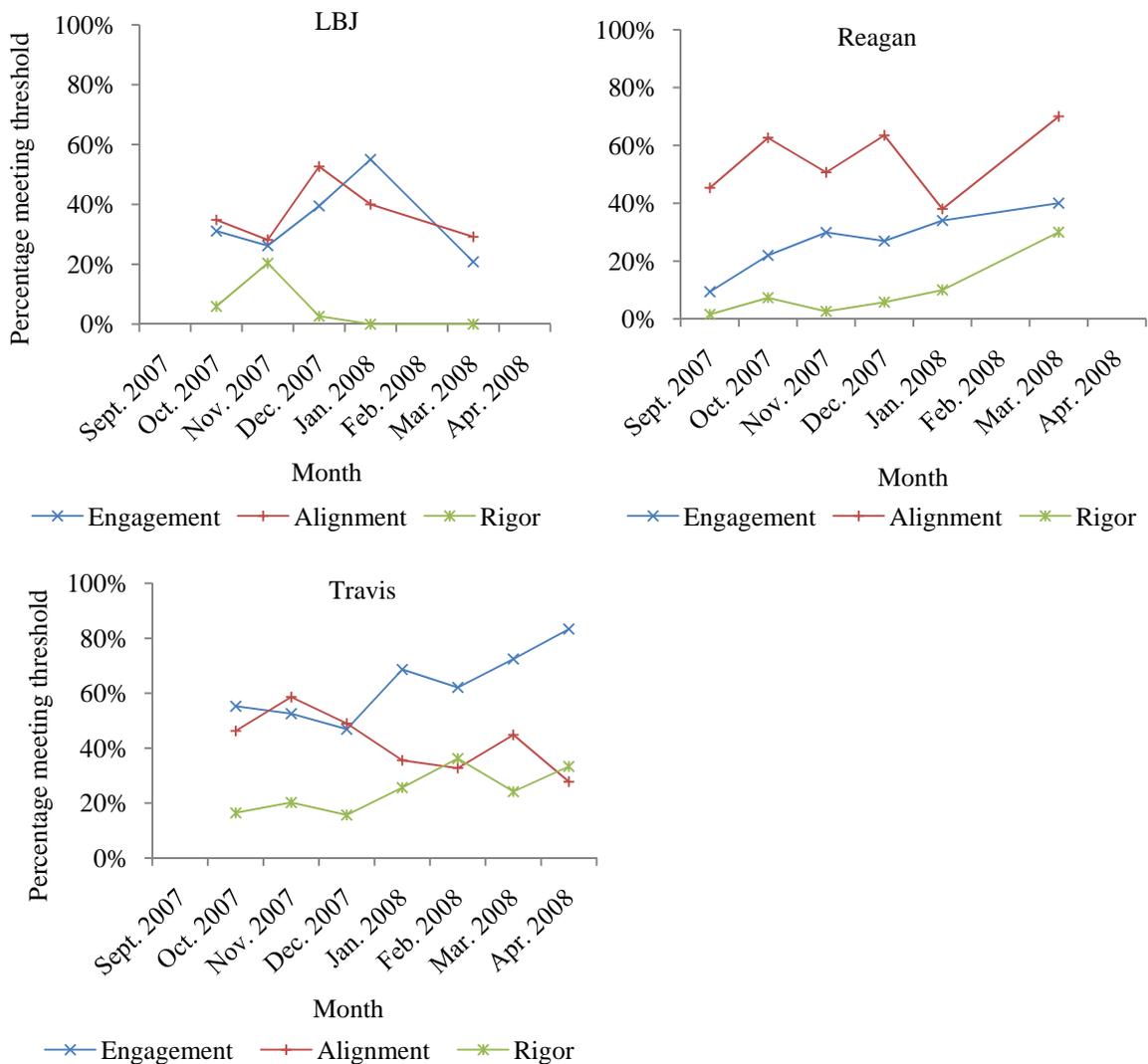
According to Connell and Broom (2004), instructional practice and strategies should be shaped, refined, and improved based on the collection and discussion of EAR protocol results. More specifically, “The protocol can be used to help instructional staff come to agreement about what good instruction looks like, better understand their strengths and challenges, and track their progress in improving their practice over time” (p. 29).

To examine whether campus instructional practices responded to this professional feedback mechanism, EAR protocol data were collected for each campus and disaggregated by month. As the percentages of classrooms that met the threshold for each of the EAR instructional components were plotted by month for the FTF campuses (Figure 3), several

<sup>1</sup> Outliers were determined by summing the 75<sup>th</sup> percentile value and (1.5\*interquartile range [IQR]), where the IQR is equal to the difference between the 75<sup>th</sup> and 25<sup>th</sup> percentile value.

discernable patterns emerged. At every FTF campus, fewer classrooms met the rigor instructional goal than met any other EAR component. However, the percentage of classroom visits satisfying this threshold increased over the course of the school year at Reagan and Travis. Reagan and Travis also exhibited an increase in the percentage of classroom visits that met the student engagement threshold through the school year. The classroom visit profile for LBJ was significantly different from that of the other two schools. The percentage of classrooms meeting the thresholds for each instruction component decreased sharply for LBJ after the midpoint of the school year.

Figure 3: Classrooms Meeting EAR Thresholds, by Campus, 2007–2008

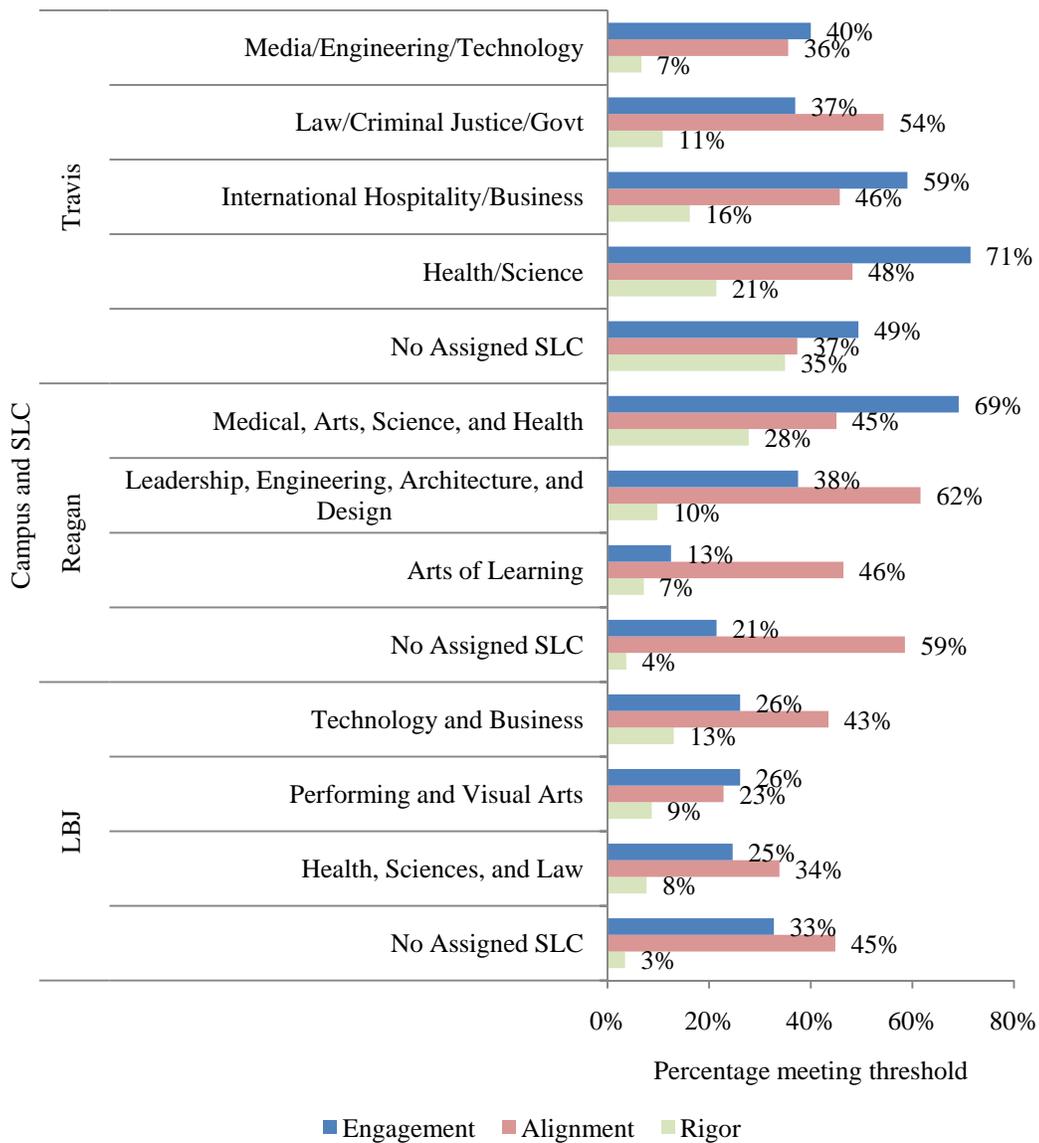


Source. IRRE, November 2008

Note. Only months in which at least 10 observations were conducted were reported.

To discern whether these trends were consistent within the schools, EAR threshold performance by SLC membership was examined (Figure 4). Within-school differences between the percentages of classrooms meeting threshold appeared across SLCs. However, fewer classrooms at each school met the rigor threshold than met any other instructional goal of the EAR protocol. At Travis, although the Health/Science SLC classroom visits were the least likely at that campus to satisfy the engagement threshold, they also had the highest percentages of classrooms meeting the rigor requirement.

Figure 4: Classrooms Meeting EAR Thresholds, by Campus and SLC, 2007–2008



Source. IRRE, November 2008

Note. Only SLCs in which at least 10 observations were conducted were reported.

**SUMMARY OF FTF'S FAMILY ADVOCACY IMPLEMENTATION IN 2007–2008****Teacher Perceptions of Family Advocacy Implementation**

During teacher focus groups, teachers discussed the implementation of the family advocacy component of the FTF initiative. Many of their comments highlighted the importance of family advocacy and its potential for positively influencing students. Teachers were excited about connecting with their students, maintaining contact with them for several years, and supporting their decision making along the way. They thought students liked the advocacy activities in which they learned about course choices, individual learning styles, and getting to know their classmates. Examples of teachers' comments are provided.

“The class time spent on selecting courses for the next year was productive. Students shared their experiences and advice on which courses to take. They talked a lot about the different electives and course requirement.”

“Goal setting was productive when students identified challenges to meeting their goals. The students helped each other by providing recommendations for overcoming the challenges to their goals.”

“The kids enjoyed finding out about the things that they had in common with one another and with me.”

Teachers also expressed concerns about the family advocacy component. They suggested that some of the activities be restructured because they did not always meet student needs or interests. This lack of student interest and engagement made it difficult for teachers to implement the existing lessons, and teachers reported that student attendance during their advocacy class was low or inconsistent. For example, a teacher said, “It is hard to get them to participate, much less really talk with you. The kids do not come to Advocacy, and there is no incentive or accountability.” Another said, “The curriculum is developed by FTF. Students do not relate to the curriculum. It does not spark interest from students, parents, or teachers.”

Teachers recommended improvements for the family advocacy curriculum. Teachers wanted more lessons that presented content to engage the students. For example, teachers found the STAR reports to be especially helpful for monitoring student progress in school. They requested that they be updated more frequently to supply the latest attendance, discipline, and grade reporting. Teachers recommended that the curriculum contain more college and career information for the students. They felt that the students needed more of these activities, and said that students liked to explore college and career information.

Teachers reported difficulty in consistently communicating and developing relationships with parents. They reported they found it difficult to find the time to contact parents when teachers had other demands on their time, when the parents had busy family

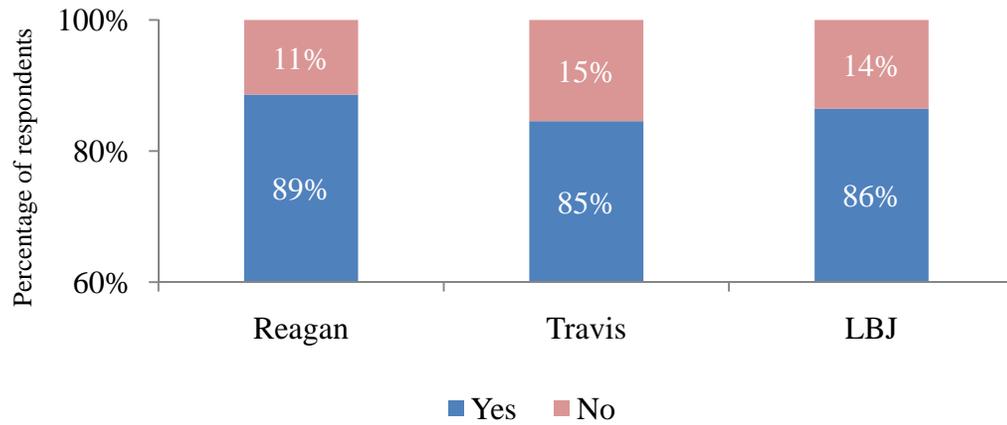
schedules, and when phone numbers and addresses in the student information system were out of date. Some teachers did not feel that it was their responsibility to be the primary school contact for parents and did not prioritize this task.

### **Student Self-Reports Concerning Family Advocacy**

Students at the district's three FTF high schools were surveyed during Spring 2008. The survey included a variety of questions pertaining to student perceptions of and attitudes toward their respective schools. Of the 2,970 students actively enrolled during the Spring 2008 semester, valid surveys were completed by 1,085 students, yielding a response rate of 36.5%. Reagan had the highest response rate (53.2%), while LBJ recorded the lowest (29.3%). Response rates are presented in the Appendix (Table A 1). For concision and to avoid overlap with other district surveys of students, only survey items relating to family advocates were analyzed.

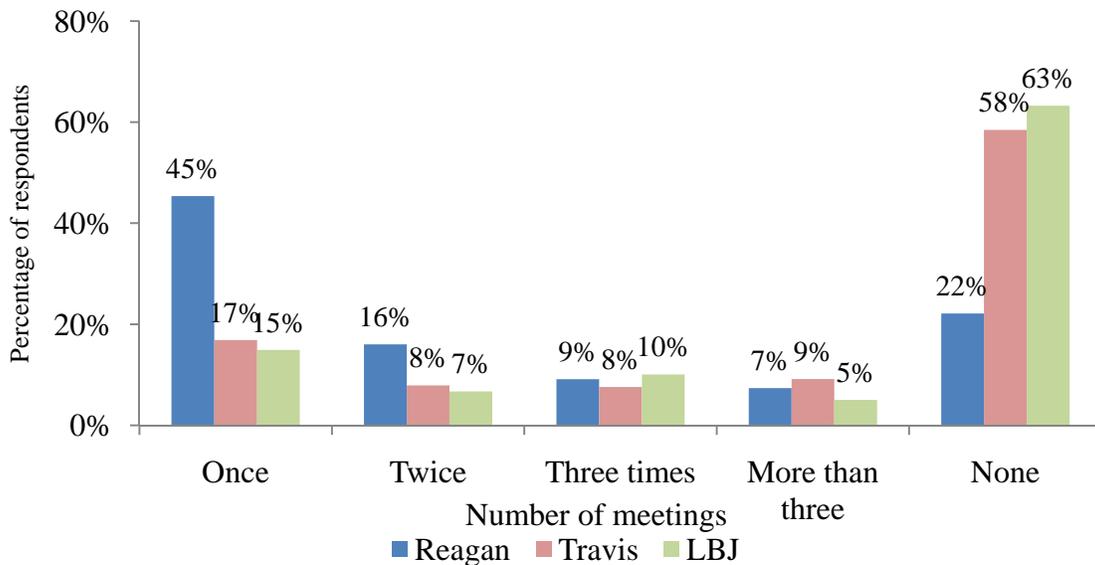
Given that strong, meaningful relationships between the students, their parents, and their family advocates are a cornerstone of the family advocacy component of the FTF reform initiative and that teachers report difficulty with parent outreach, student perceptions of the relationships between students, their families, and school staff were explored using questions from the FTF survey. Overall, 86% of students who responded to the FTF survey reported they knew who their family advocate was (Figure 5). Approximately 89% of students at Reagan answered "Yes" to this question, while 85% at Travis responded similarly. Despite this high level of family advocate recognition, 63% of respondents from LBJ and 58% of respondents at Travis claimed their family advocate had not conducted a parent/guardian and student conference; these percentages were dramatically higher than the percentage at Reagan (22%) (Figure 6). Forty-five percent of Reagan respondents, in sharp contrast with respondents at Travis (17%) and LBJ (15%), indicated one family conference had been organized by their family advocate. In an independent evaluation of the first year of implementation of FTF schools outside of Kansas City, Quint, Byndloss, and Melamud (2003) found that 32% of students reported that planned meetings with the student and the student's family had been conducted. Their findings align closely with the results at Travis (42%) and LBJ (37%), whereas the high percentage of respondents at Reagan reporting a meeting had been held (78%) is acutely elevated.

Figure 5: Students Reporting They Knew Their Family Advocate, by Campus, 2007–2008



Source. FTF Student Survey, prepared by the DPE, October 2008

Figure 6: Students Reporting an Advocate Met With a Parent/Guardian, by Campus and Frequency of Meeting, 2007–2008



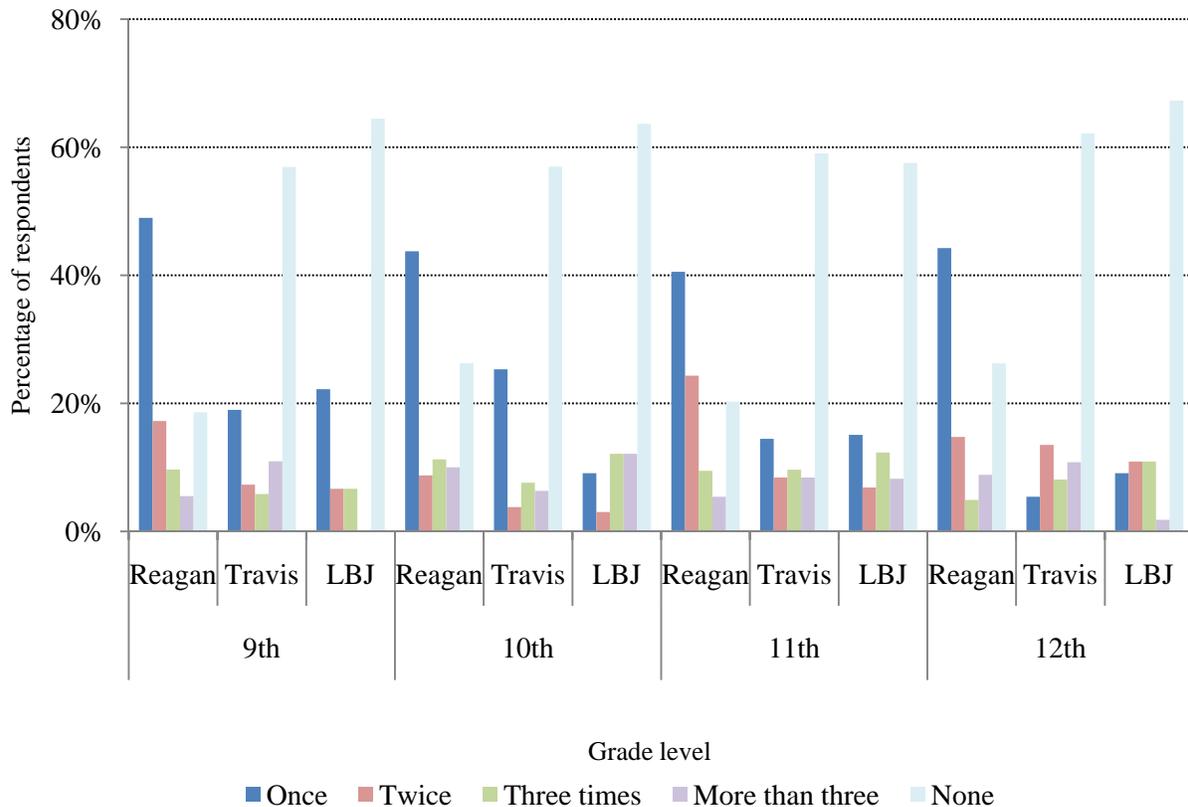
Source. FTF Student Survey, prepared by the DPE, October 2008

The patterns of parent/guardian, student, and family advocate meetings showed variation across grade levels within schools (Figure 7). At Reagan, freshmen respondents (49%) were the most likely of the student groups to report having participated in one family advocate meeting with a parent/guardian, while 11<sup>th</sup> grade respondents were the least likely (41%). Conversely, among Travis respondents, 25% of 10<sup>th</sup> grade respondents indicated their family advocate had conducted a parent/guardian conference once during the school year, compared with 19% of 9<sup>th</sup> grade respondents and 14% of 11<sup>th</sup> grade respondents.

Focus group interviews with advocates at FTF schools revealed that advocates had persistent difficulties reaching parents to secure greater involvement. According to one family

advocate at Travis, advocates “make phone calls and have difficulty contacting the parents or getting them to come to the school.” Citing time limitations, some family advocates acknowledged they had difficulty contacting parents proactively.

Figure 7: Students Reporting an Advocate Met With a Parent/Guardian, by Campus and Grade Level, 2007–2008

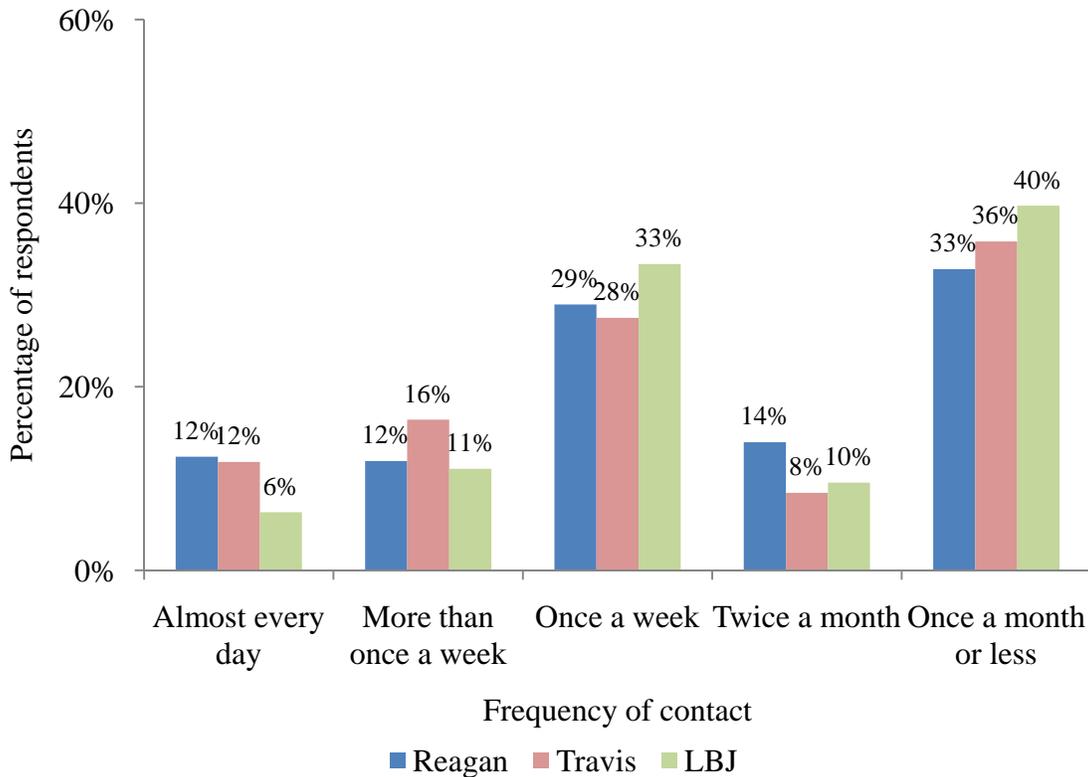


Source. FTF Student Survey, prepared by the DPE, October 2008

According to the FTF framework, frequent and substantive interactions between students and their family advocates are a critical mechanism for cultivating the conditions necessary to improve student performance (Klem, Levin, Bloom, & Connell, 2003). These relationships represent an important connection between students, their families, and school, and they provide a stable and open communication avenue between students and school staff. Thus, these relationships allow staff to respond quickly to students’ academic, behavioral, and social needs, while also providing a mechanism to monitor whether students are staying on track to accomplish their goals. Approximately 36% of respondents reported interacting with their family advocate outside of regularly scheduled class periods once a month or less, while 30% indicated they met with their family advocate once a week (Figure 8). Again, these findings are analogous to those published by Quint et al. (2003), who reported 31% of students

indicated meeting once a week during the first full year of implementation.

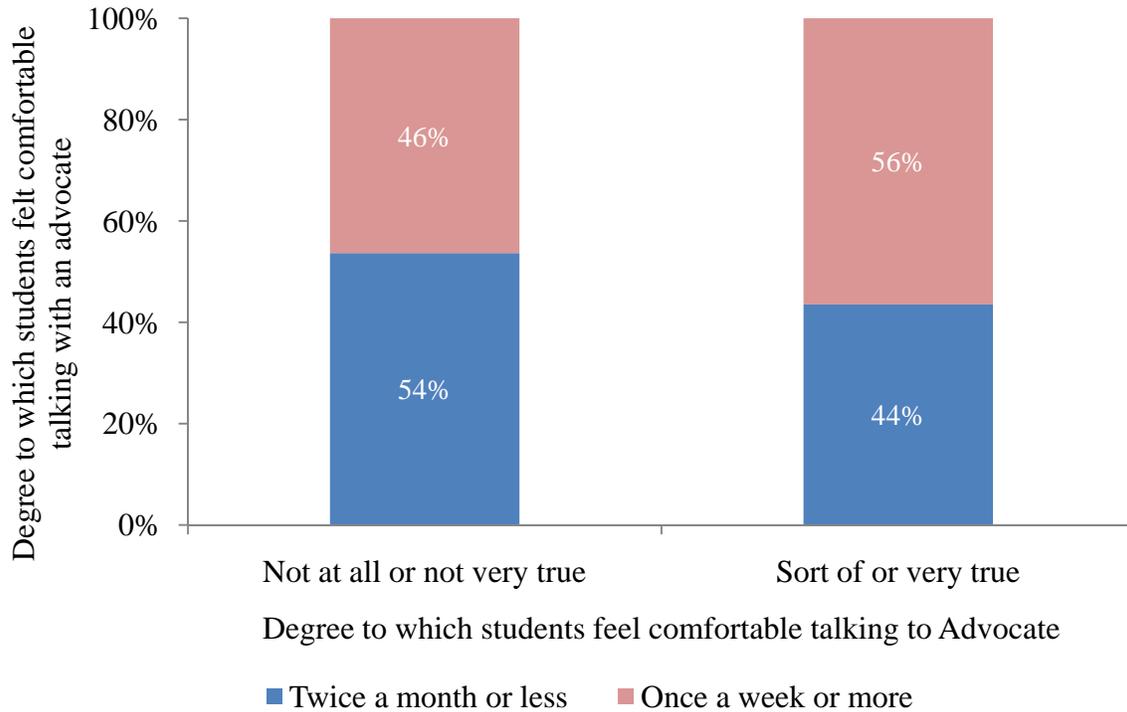
Figure 8: Student Reports of Frequency of One-on-One Contact With an Advocate, by Campus, 2007–2008



Source. FTF Student Survey, prepared by the DPE, October 2008

Students' willingness to discuss academic, social, and family difficulties may be conditioned by whether they feel comfortable interacting with their family advocate. Figure 9 provides tentative empirical support for this proposition. Fifty-six percent of respondents who considered the statement "I feel comfortable talking with my family advocate" to be "sort of or very true" reported having spoken with their advocate at least once a week outside of a formal advisory period setting. In contrast, students who did not agree with this statement were 10 percentage points less likely to report meeting once a week or more. These findings should be interpreted with care because the data did not indicate whether students were reluctant to approach an advocate because they felt uncomfortable discussing personal issues with an advocate, or whether they gradually become more comfortable as the frequency of interaction increased.

Figure 9: Student Reports About Whether They Felt Comfortable Talking With a Family Advocate



Source. FTF Student Survey, prepared by the DPE, October 2008

### THE IMPACT OF FTF ON TAKS PERFORMANCE

According to IRRE's (2003) "theory of change," structural reorganization through the creation of SLCs and family advocacy periods, in tandem with continuous instructional monitoring and refinements using the Measuring What Matters (MWM) toolkit, is expected to improve student outcomes. IRRE posited and investigated a direct link between FTF implementation and gains in student performance on state assessments. The comprehensive high school reform initiative advocated by IRRE was "designed to help schools raise their students' achievement to levels needed for postsecondary education (without remediation) and high-quality employment" (p.2). This section analyzes the impact of FTF initiation on one dimension of student achievement: TAKS performance.

According to a variety of student outcome measures and demographic characteristics, schools selected for FTF implementation differed considerably from those not chosen (Table 2). For example, at least 80% of the student body at each FTF campus qualified for the national free or reduced price lunch program. In 2006–2007, the highest average scale score on the TAKS math test (LBJ, 2093) was 35 points below the average scale score on the TAKS math test (Akins, 2128) at the lowest performing non-FTF campus not selected as a FTF-comparison campus. In addition, TAKS math and reading proficiency rates at FTF campuses were far below those of other campuses, with the exception of Johnston (Figure 10). Such stark

differences in student characteristics between FTF and non-FTF campuses presented difficulties in determining the effect of FTF implementation on student outcomes across time.

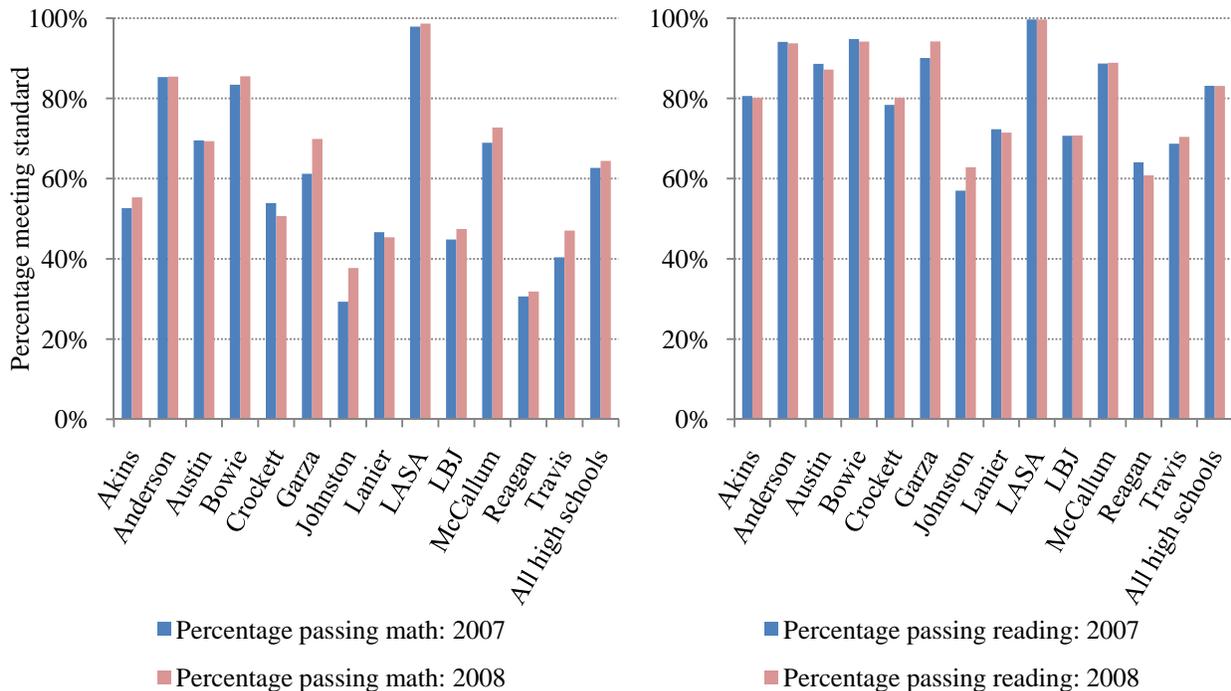
Table 2: TAKS Performance and Economically Disadvantaged Status, by FTF Status, 2006–2007 and 2007–2008

		2006–2007			2007–2008		
		Mean TAKS math	Mean TAKS reading	Percentage economically disadvantaged	Mean TAKS math	Mean TAKS reading	Percentage economically disadvantaged
FTF campus	Reagan	2034	2142	86%	2034	2139	89%
	LBJ	2093	2157	77%	2077	2169	82%
	Travis	2066	2159	83%	2100	2172	87%
FTF-comparison campus	Lanier	2088	2181	82%	2090	2194	88%
	Johnston	2023	2114	88%	2044	2147	92%
Other campus	Austin	2214	2281	29%	2224	2279	33%
	McCallum	2215	2281	34%	2237	2289	38%
	Crockett	2129	2212	52%	2120	2218	59%
	Anderson	2334	2326	17%	2342	2338	21%
	Bowie	2287	2312	7%	2306	2325	11%
	Akins	2128	2216	56%	2127	2217	62%
	LASA	2425	2382	25%	2460	2402	27%

Source. AISD student records, prepared by the DPE, October 2008

Note. TAKS scores represent the mean of the highest valid score received by a student on any test administration during the 2006–2007 and 2007–2008 school years.

Figure 10: Students Meeting Math and Reading TAKS Standards, by Campus, 2006–2007 and 2007–2008



Source. AISD student records, prepared by the DPE, October 2008

Note. TAKS proficiency represents the highest valid score deemed proficient that was received by a student on any test administration during the 2006–2007 and 2007–2008 school years.

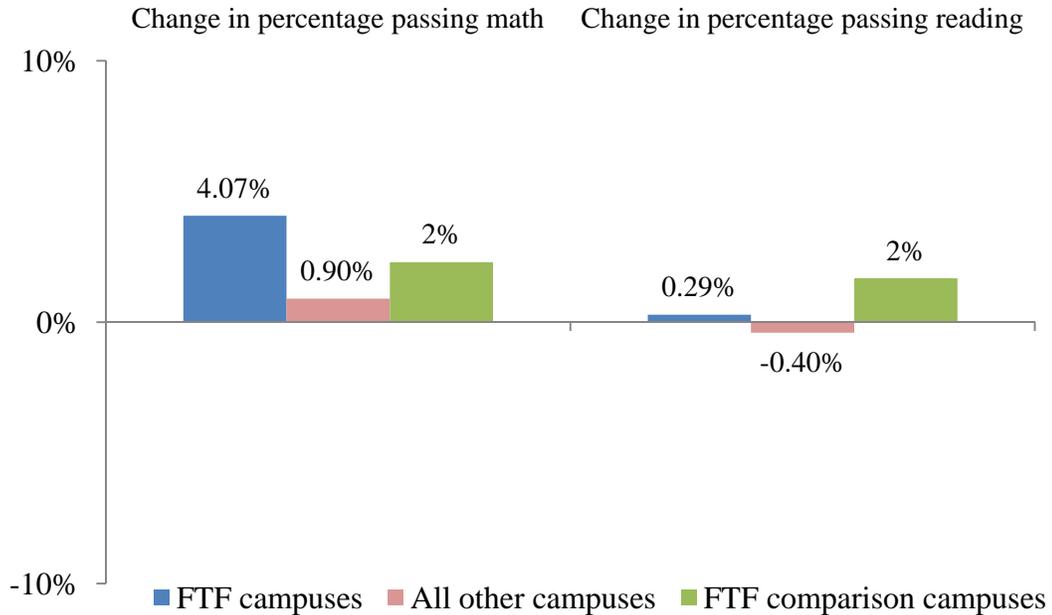
To attribute the impact of FTF intervention to student performance outcomes, we had to determine what student outcomes could have been in the absence of FTF program implementation. This “counterfactual” reasoning is essential to draw causal inferences from the data analyses. Despite its strengths, however, this approach is not without its hazards. The comparability of the comparison campus to the campus selected for intervention hinges on the variables used in the selection process. Using this approach requires the assumption that cases are identical in every sense, except for the program being implemented. This is a demanding requirement, particularly in the area of education research, in which interventions continually are being implemented and refined. Because of this, the robustness of the conclusions drawn using this strategy are qualified.

To formulate a “counterfactual” scenario, we followed the lead of Quint et al. (2005) and identified two schools in AISD that were comparable to LBJ, Reagan, and Travis: Johnston and Lanier. A more detailed explanation of the selection methodology is provided in Appendix B.

Aggregate changes in the percentages of students scoring proficiently on the TAKS math and reading tests between 2006–2007 and 2007–2008 were examined at FTF schools,

FTF-comparison schools, and all other AISD high schools. The increase in the percentage of students meeting the TAKS standard at FTF campuses, particularly for the math test, was sizeable (4.07 percentage points) and was approximately twice that at the FTF-comparison schools (2.29 percentage points). The remaining campuses improved minimally (<1 percentage point).

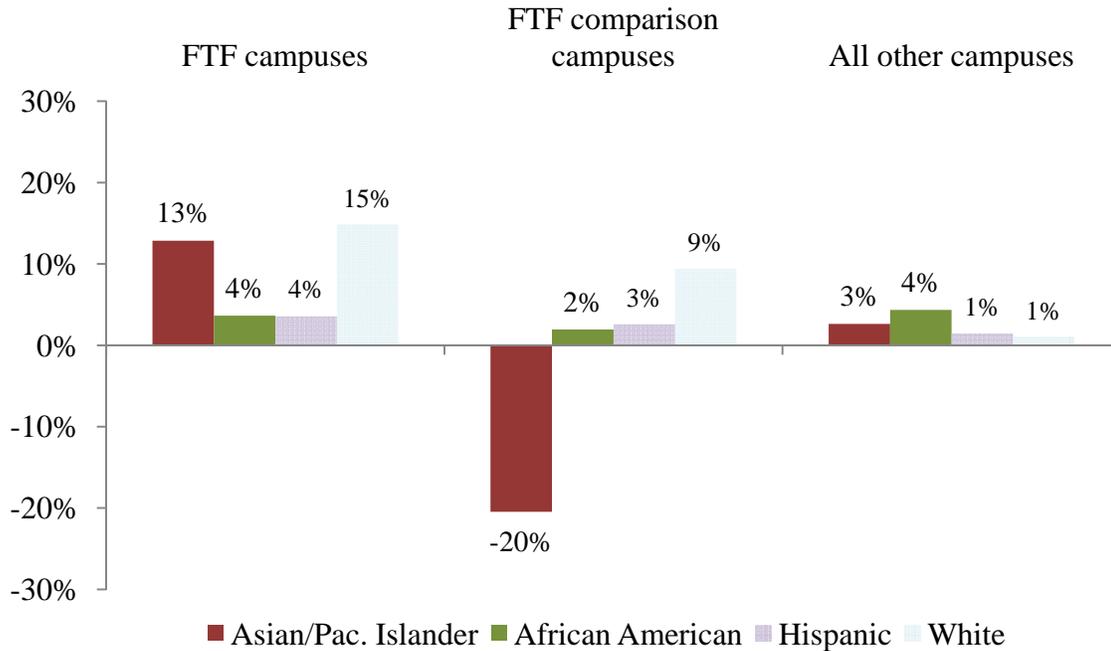
Figure 11: Students Meeting Math and Reading TAKS Standards, by FTF Status, 2006–2007 and 2007–2008



Source. AISD student records, prepared by the DPE, October 2008

A principal objective of national, state, and district educators is the narrowing of achievement gaps across students with different demographic backgrounds. Although FTF campuses experienced a sharp increase in the percentage of students satisfying the math TAKS passing standard between 2006–2007 and 2007–2008, the improvements were not symmetrically distributed across ethnic groups (Figure 12). For instance, White students at FTF campuses showed the most dramatic improvement in passing rates between 2006–2007 and 2007–2008 (15 percentage points), compared with gains for African American and Hispanic students (4 percentage points).

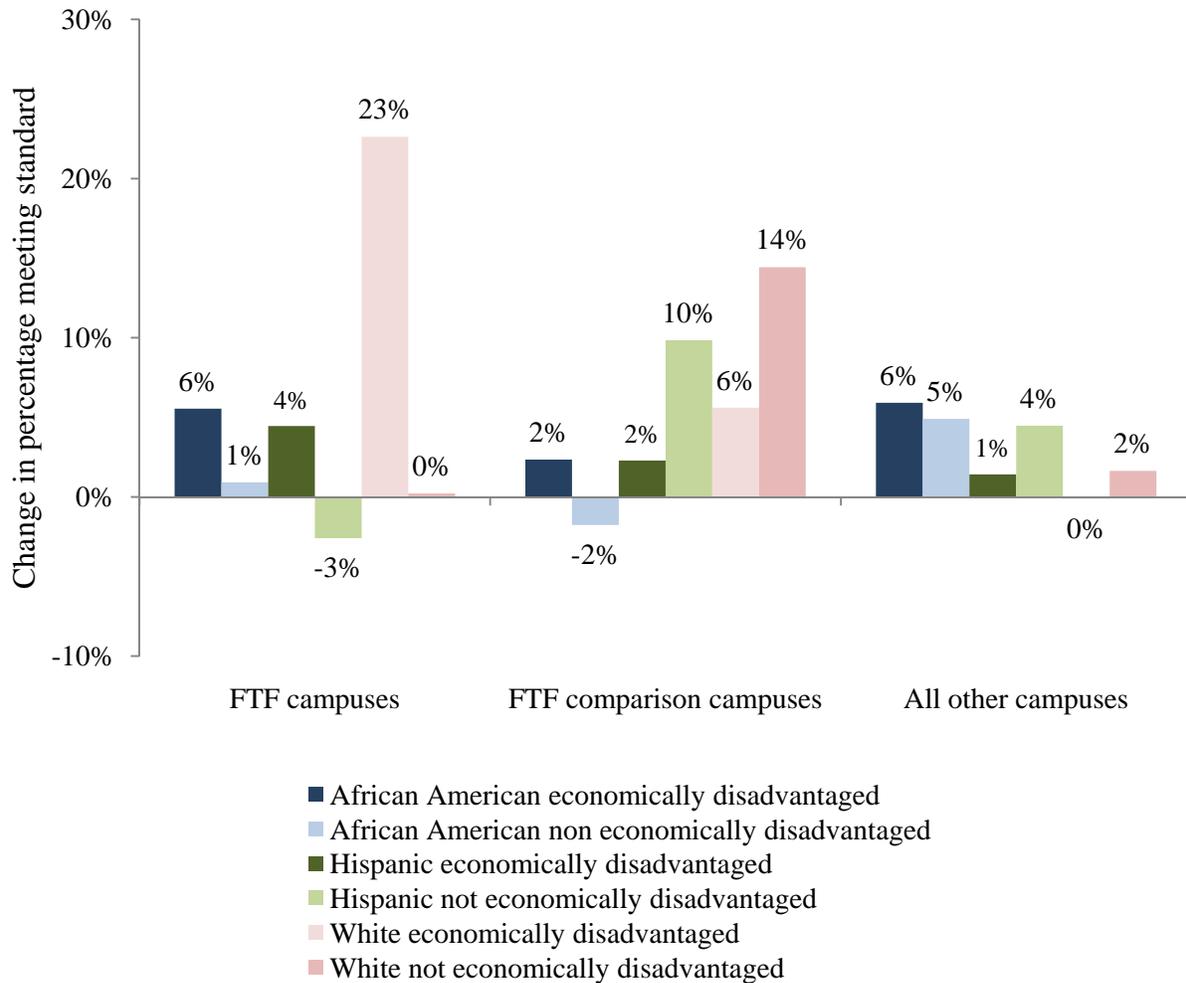
Figure 12: Students Meeting Math TAKS Standards, by FTF Status and Ethnicity, 2006–2007 and 2007–2008



Source. AISD student records, prepared by the DPE, October 2008

Grouping TAKS test performance by ethnicity assumed the student composition of each ethnic group was homogenous, and that students across the groups only differed in their ethnic classification. Figure 13 incorporates economic disadvantage status to account for student-level characteristics that varied across ethnic groups and reinforces the patterns identified in Figure 6. The math TAKS passing rates of economically disadvantaged White students at FTF campuses increased markedly from 2006–2007 to 2007–2008 (23 percentage points), compared with gains for student at FTF comparison campuses (6 percentage points) and a negligible increase for economically disadvantaged African American students at the remaining high schools. The gains at FTF campuses for economically disadvantaged students, irrespective of ethnicity, were greater than those at the FTF comparison and other campuses.

Figure 13: Students Meeting Math TAKS Standards, by FTF Status, Ethnicity, and Economic Disadvantage Status, 2006–2007 and 2007–2008



Source. AISD student records, prepared by the DPE, October 2008

To account for school- and student-level factors and to better determine whether the improvement resulting from FTF intervention was meaningful and not generated by sampling or measurement differences, multivariate logistic regression procedures incorporated a host of variables to explain the probability a given student would meet the required TAKS score during a given administration year (for a more technically detailed explanation about the multivariate estimation procedure, see Appendix C). The following variables were included: gender, economic disadvantage status, ethnicity dummy variables, special education status, grade point average (GPA), limited English proficiency (LEP) status, and campus dummy variables to adjust for intra-school clustering.

To separate the impact of FTF implementation at FTF and FTF-comparison schools, we constructed and estimated a series of models that progressively incorporated variables potentially related to whether a student satisfied the math TAKS standard. First, and most

naively, we estimated the change in the predicted probability from 2006–2007 to 2007–2008 that a student at a FTF campus or FTF-comparison campus met the math or reading minimum proficiency score (“unconditional model”). Statistically significant changes in Figure 14 are denoted by asterisks; the level of confidence in the accuracy of the estimated relationship is provided in the footnote. Second, we added student- and school-level variables to the statistical model (“conditional model”). This model included the following variables: gender, economic disadvantage status, ethnicity dummy variables, special education status, GPA, LEP status, and campus dummy variables to adjust for intra-school clustering. This model ensured that any observed changes in test performance across school years were not generated by factors unrelated to FTF implementation.

To illustrate the impact of specific variables on student outcomes, vertical bar graphs are displayed. The taller the vertical bar, the more decisive the impact of the factor on the graduate’s outcome. The height of the vertical bar is determined by comparing the difference in the likelihood of a student outcome between two students who are alike in most respects, but who are different on one critical characteristic. For our purposes, that critical characteristic was whether the student was enrolled in a FTF or non-FTF school during the 2007–2008 school year.

Beginning with the unconditional math TAKS model, which did not control for school- or student-level characteristics, students at FTF campuses showed statistically significant gains with respect to the likelihood of meeting the math TAKS standard between 2006–2007 and 2007–2008 (Figure 14). Students at Reagan, Travis, and LBJ were 4 percentage points more likely to satisfy the math standard in 2007–2008 than in 2006–2007. The difference across years was smaller and not statistically significant at the FTF-comparison campuses (2 percentage points). No statistically significant changes were detected at FTF or FTF-comparison campuses for the reading TAKS analysis.

In sum, after controlling for student- and school-level characteristics, FTF program

implementation did not have a measurable impact on student math or reading TAKS passing rates (Figure 14). The percentages of students at FTF campuses who met the passing standard on the math TAKS increased sharply and significantly from 2006–2007 to 2007–2008.

***How do I interpret the predicted probability graphs?***

To illustrate the impact of specific variables on student outcomes, vertical bar graphs are displayed throughout the report. The taller the vertical bar, the more decisive the impact of the factor on the graduate’s outcome.

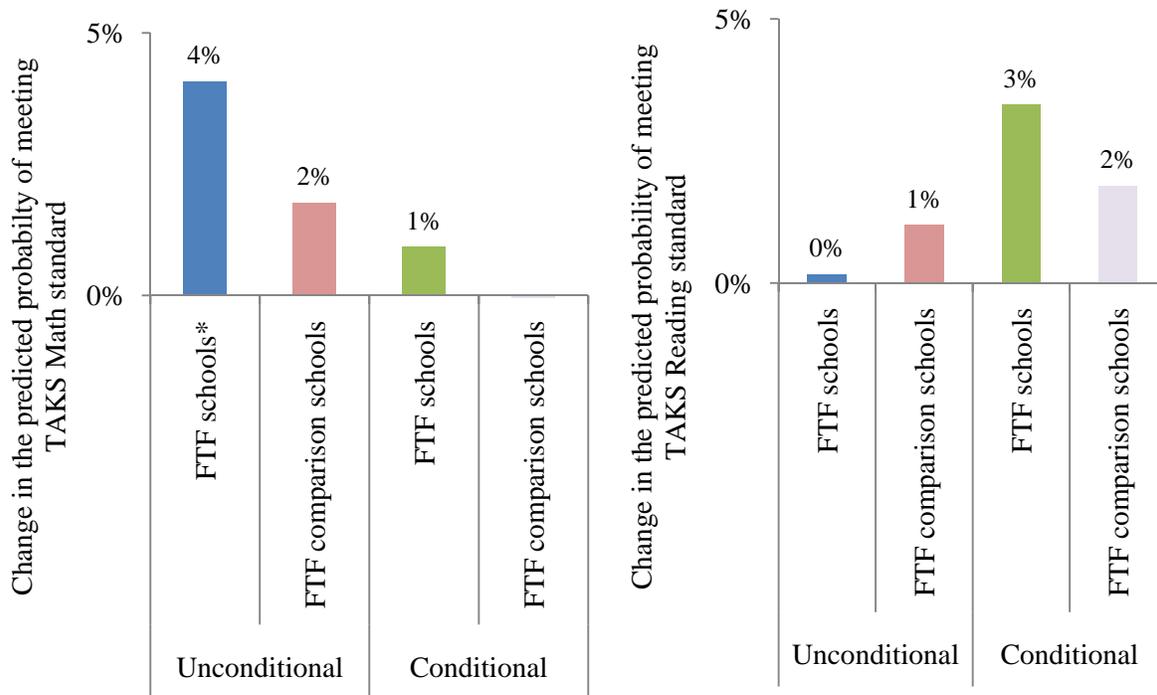
The height of the vertical bar is determined by comparing the difference in the likelihood of a student outcome between two students who are alike in most respects, but who show substantial differences in one characteristic. For our purposes, that substantial difference is time, which captures the improvement in student performance from the 2006-2007 and 2007-2008 academic years. Of primary substantive interest in this report is whether the improvement, or decline, in performance between school years was large enough to constitute a statistically significant difference from the prior year, particularly at FTF campuses.

Using an example from Figure 14 below, the probability that a student met the math TAKS standard in 2007-2008 showed strong improvement (4 percentage points) from 2006-2007. This difference was statistically significant. The strength of this improvement (1 percentage point) weakened considerably after student and school-level characteristics were taken into consideration.

However, these gains were not symmetric across ethnic groups, nor did they remain when controlling for student- and school-level characteristics.

Although students at FTF campuses demonstrated a 1 percentage point increase in the likelihood of meeting the math TAKS standard in 2007–2008, compared with 2006–2007, and this gain exceeded the increase observed for FTF-comparison campuses, this improvement was not statistically significant. It is important to note that Quint et al. (2005) did not find statistically significant differences during the first implementation year in scores on state-mandated reading or math tests for most school districts they examined. Moreover, the impact of FTF implementation varied across individual campuses, even after controlling for alternative explanations. This was not fully captured by the modeling strategy adopted. Thus, drastic revisions to the FTF intervention model based upon our results may not yet be warranted.

Figure 14: Predicted Probability of Earning a Proficient Score on Math or Reading TAKS Between 2006–2007 and 2007–2008, by FTF Status



Source. AISD student records, prepared by the DPE, October 2008

Note. Estimates were derived from a logistic regression with robust standard errors and unit-specific effects.

\* p <.01, two-tailed test

**THE IMPACT OF FTF ON DISCIPLINARY REFERRALS**

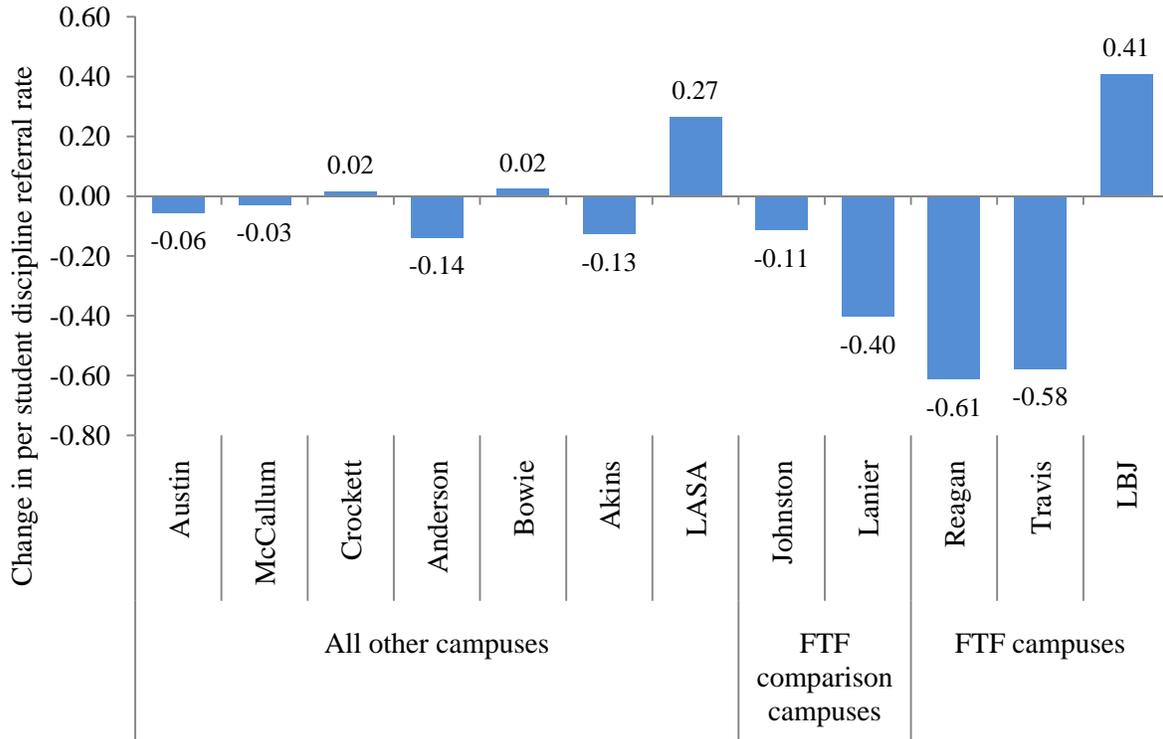
If the structural and instructional changes promoted by the FTF initiative are successfully implemented and maintained with fidelity, an important student outcome that should improve is the disciplinary referral rate (Klem et al., 2003). This reduction is achieved through the cultivation of strong student, parent, and teacher bonds. The Family Advocacy System establishes a channel through which

Students experience the regular contact necessary to feel that there is someone in the school to whom they go with problems, who cares about them and looks out for their best interests, who guides their decisions, and who strives to keep them on track academically and behaviorally. (p. 3)

In this section, we explore the link between FTF implementation and disciplinary referral patterns. We adopted a research design that mirrored the one used in the previous section (for more detailed explanations about the extraction and preparation of disciplinary data, in addition to coding judgments, see Appendix D).

Student discipline referral rates, defined as the average number of referrals per student including both discretionary and mandatory offenses, decreased markedly from 2006–2007 to 2007–2008 for two of the three (Reagan and Travis) FTF schools (Figure 15). The largest decline occurred at Reagan, where the number of per student referrals declined by .61 referrals. Travis witnessed a similarly dramatic fall in per student referrals (.58). These declines exceeded those at the two comparison schools. At LBJ, the per student discipline referral rate increased (.41) in 2007–2008.

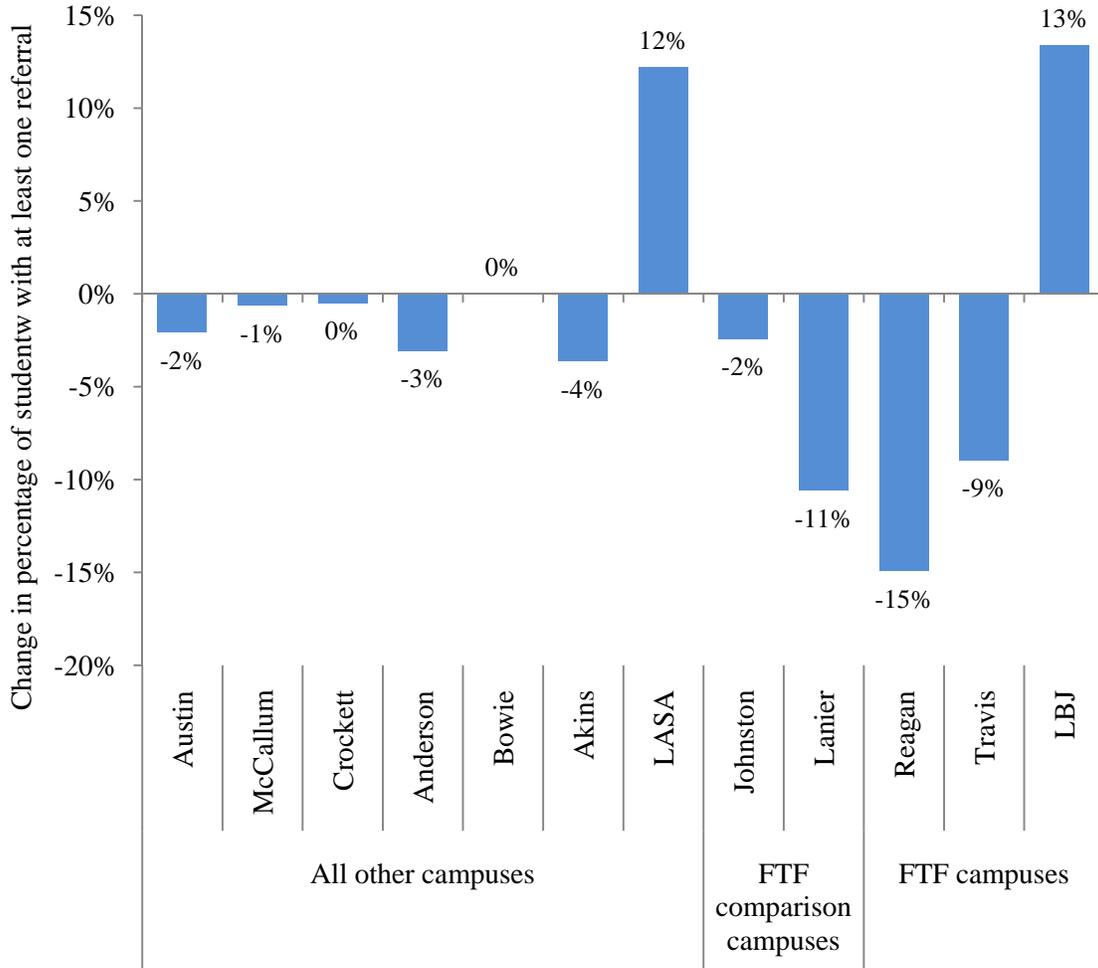
Figure 15: Per Student Discipline Referral Rate, by Campus and FTF Status, 2006–2007 and 2007–2008



Source. AISD student discipline records, prepared by the DPE, October 2008

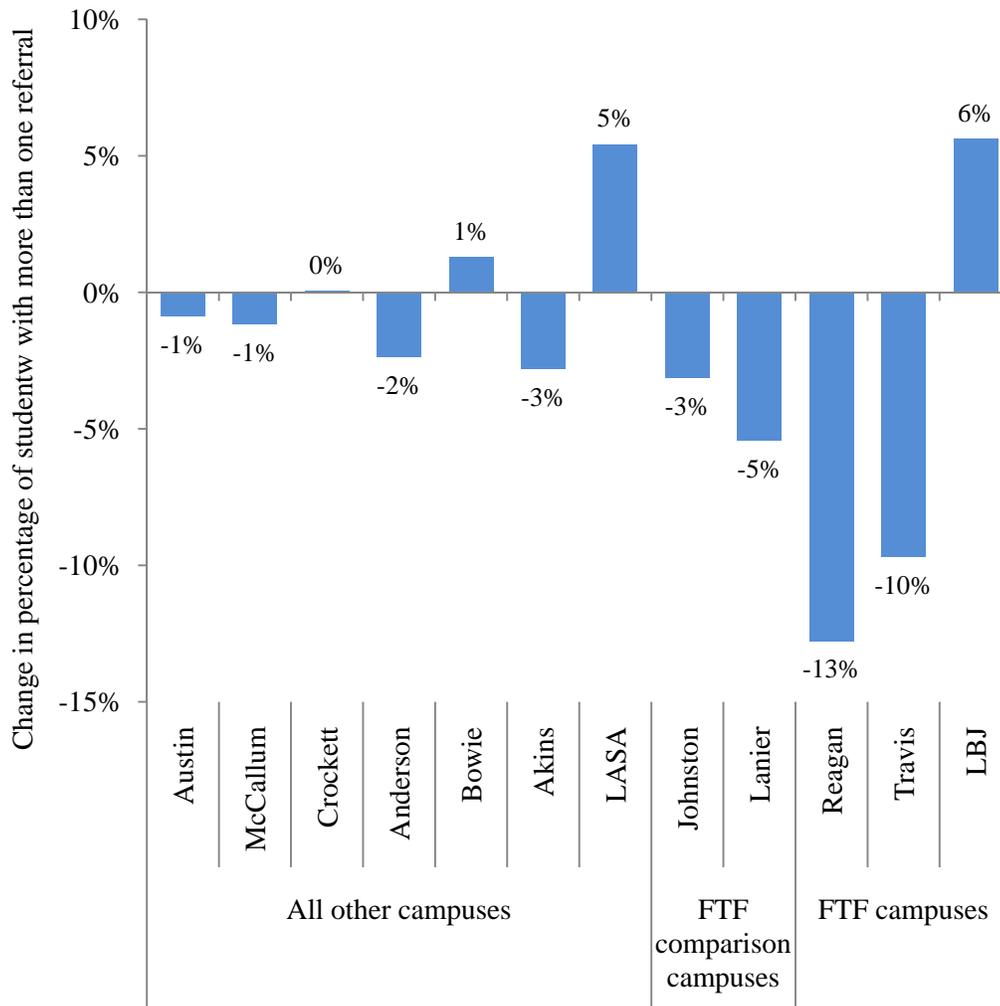
The reduction in the per student discipline referral rate at Reagan and Travis was accompanied by a decrease in the percentage of enrolled students who had at least one referral and who had multiple referrals (Figures 16 and 17). At Reagan, the share of students with at least one referral declined from 36.8% to 21.9% (15 percentage points). Travis also showed a sharp decline (9 percentage points), although this was less than the decrease at Lanier (11 percentage points). Furthermore, the percentage of students with multiple referrals decreased at Travis (10 percentage points) and Reagan (13 percentage points). Once again, these improvements exceeded those that occurred at the FTF-comparison campuses.

Figure 16: Students With at Least One Referral, by Campus and FTF Status, 2006–2007 and 2007–2008 Status



Source. AISD student discipline records, prepared by the DPE, October 2008

Figure 17: Students With Multiple Referrals, by Campus and FTF Status, 2006–2007 and 2007–2008



Source. AISD student discipline records, prepared by the DPE, October 2008

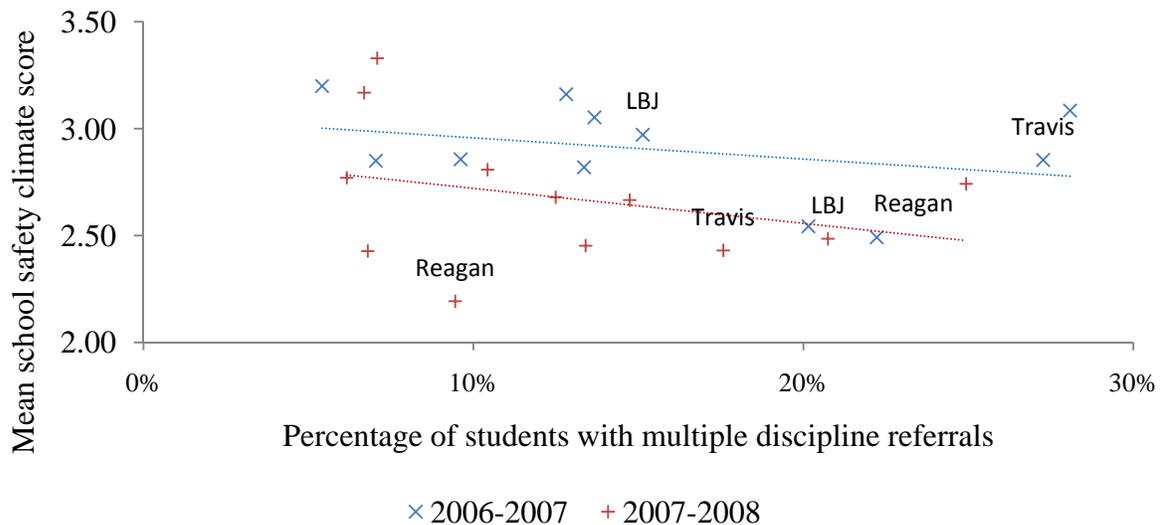
Whether the decreases in referrals represent an actual reduction in the type of behaviors precipitating referrals on campuses is unclear. The frequency of student misbehavior at school may have remained steady; however, campus staff may have become more tolerant of this behavior and less likely to initiate referrals that would have prompted disciplinary action in preceding years. Researchers examining disciplinary referral patterns continually grapple with this dilemma, including the challenges of identifying the disproportional disciplinary treatment of ethnic groups (Skiba, Michael, Nardo, & Peterson, 2002). One approach to address this problem is to control for staff perceptions of student misbehavior on campuses.

If referral rates declined in response to improvements in student behavior, staff perceptions of safety on campuses should reflect this phenomenon, with reports of campus safety becoming more favorable. School safety is a composite measure of subscale items

capturing staff perceptions of the type and frequency of undesirable student behaviors on school grounds (Schmitt, Bush-Richards, & Cornetto, 2008). By graphing the relationship between staff perceptions of campus safety taken from the 2006–2007 and 2007–2008 AISD Staff Climate Survey, several patterns were discernable (Figure 18). A weak, negative relationship was found between staff perceptions of school safety and the percentage of students with multiple referrals. A similar relationship was found when the school safety survey item was plotted against the percentage of students on campus with at least one referral. More directly, staff at schools with a high percentage of students with multiple referrals reported low perceptions of campus safety. However, at Travis and Reagan, two schools that achieved sharp declines in referral rates between 2006–2007 and 2007–2008, staff perceptions of safety declined.

One potential alternative explanation for this finding is that the erosion in perceptions of safety that corresponded with declining percentages of students receiving multiple discipline referrals might have been a function of shifts in the severity of offenses that occurred on campus. That is, if violent offenses remained prevalent or spiked, staff evaluations of campus safety would have been unaffected or become more negative. This potential explanation warrants deeper exploration in future analyses.

Figure 18: Relationship Between Staff Perceptions of Campus Safety and Percentage of Students With Multiple Discipline Referrals, 2006–2007 and 2007–2008



Source. AISD student discipline records, 2007–2008 and 2006–2007, and AISD Staff Climate Survey, prepared by the DPE, October 2008

The analyses presented in Figure 18 did not control for differences in student characteristics within schools and across school years. Because chronic disciplinary problems are correlated with student achievement and retention, and dropout rates vary widely across campuses, the decreases in referral rates may have been affected by the attrition among

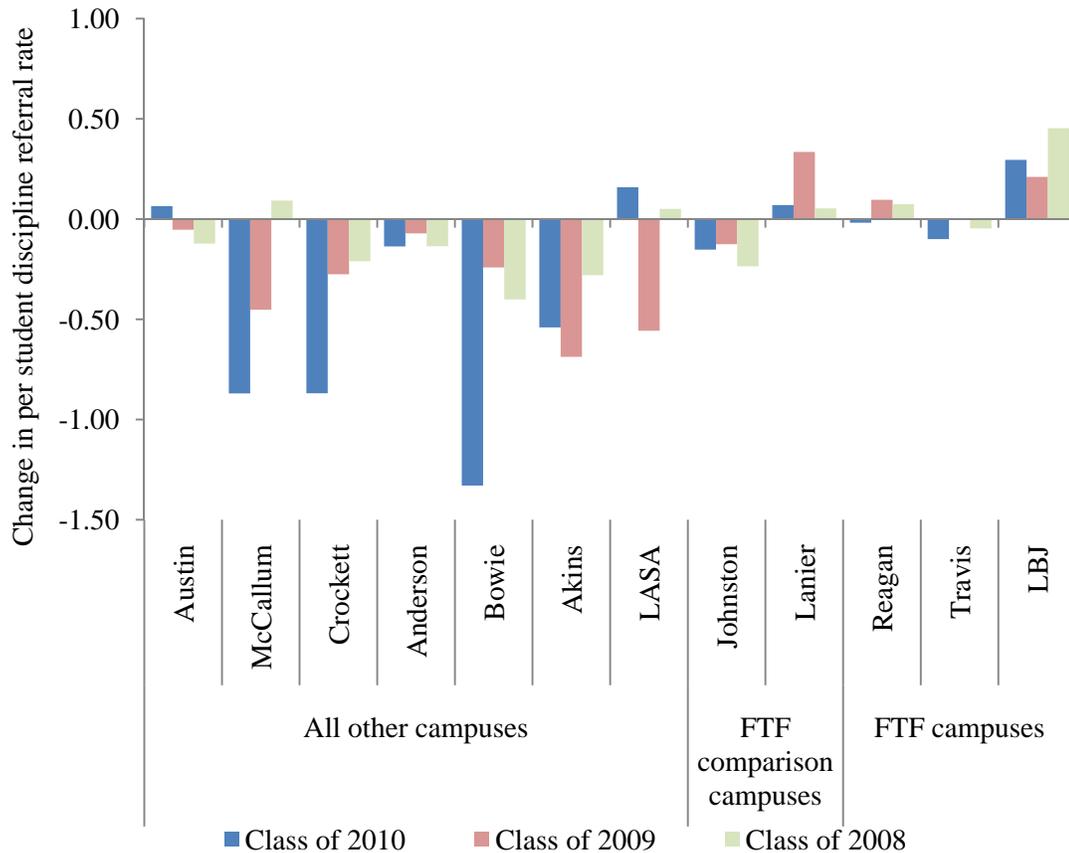
students with high referral rates (Ekstrom, Goertz, Pollack, and Rock, 1986). Ekstrom, Goertz, Pollack, and Rock (1986), for instance, found that dropping was strongly associated with frequent behavioral and attendance problems. Thus, improvements in referral rates might be an artifact of campus-level dropout rates.

To investigate the plausibility of this explanation, discipline referral rates were calculated and grouped by students' graduation cohort to eliminate the potentially differential impact of student dropout rates across campuses (Figure 19). Only students who remained enrolled and who advanced grade levels were included in the calculation to attain comparability across years. Each graduation cohort represents the difference in the per student referral rate between 2006–2007 and 2007–2008 school year. For instance, for the Class of 2010, who were in 9<sup>th</sup> grade in 2006–2007 and transitioned to 10<sup>th</sup> grade in 2007–2008, the average per student discipline referral rate declined by nearly 1.5 referrals per student at Bowie. Similar declines were found at McCallum and Crockett for the Class of 2010. After controlling changes in enrollment patterns, the frequency of discipline referrals at these campuses appears to decline as students advanced grade levels.

Among FTF campuses, the declines among members of Class of 2010 cohorts were considerably smaller, or in the case of LBJ, greater. Reagan and Travis, which experienced the sharpest decline in per student referral rates from 2006–2007 to 2007–2008, also had the highest percentage of students with more than five office referrals in 2006–2007 (9% and 11%, respectively). Of students with more than five referrals, only half remained enrolled. Students who did not return in 2007–2008 accounted for 33% of total referrals at Travis and 26% of total referrals at Reagan in 2006–2007.

In sum, the results do indicate a tentative connection between dropout rates and disciplinary referral rates, although additional exploration is warranted. The cohort-specific declines at the non-FTF campuses were markedly greater than those experienced at FTF. Moreover, students at Travis and Reagan who did not remain enrolled in the 2007-2008 school year accounted for one third to one quarter of all referrals in the previous school year. Furthermore, these results suggest that, in line with Ekstrom, Goertz, Pollack, and Rock (1986), excessive referral accumulation may be an important precursor to student attrition.

Figure 19: Per Student Discipline Referral Rate, by Campus, Cohort, and FTF Status, 2006–2007 and 2007–2008



Source. AISD student discipline records, prepared by the DPE, October 2008

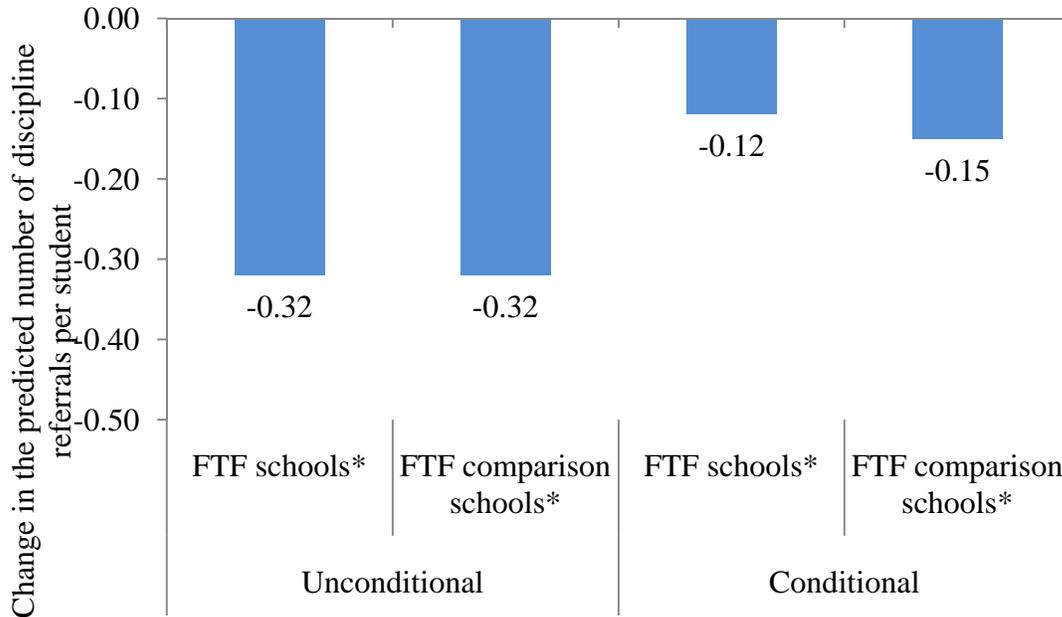
Note. Only students who were enrolled during both school years and who had advanced grade levels were included in the analyses.

To gain a better understanding of the impact of FTF implementation on discipline referral incidences, multivariate logistic regression procedures were used to control for potentially confounding student- and school-level variables. Both an unconditional model, with only the FTF implementation measure included, and a conditional model, incorporating a series of school- and student-level control variables, were estimated to determine the impact of FTF on student disciplinary incidents while controlling for confounding student and school-level characteristics.

Both FTF campuses and FTF-comparison campuses experienced statistically significant declines in the per student discipline referral rate from 2006–2007 to 2007–2008 (Figure 20). This finding reinforces the results presented in Figure 19, in which the per student discipline referral rates at Reagan and Travis declined markedly and in excess of the declines calculated for FTF comparison schools. The estimated declines withstood the inclusion of control variables and remained statistically significant. However, after controlling for school and

individual characteristics, the magnitude of the change between 2006–2007 and 2007–2008 was modestly larger at FTF comparison schools compared to FTF schools (.03 student referrals) at FTF-comparison campuses. These results are encouraging for district programs because the findings indicate high schools experienced a widespread decline in referral frequency even in schools not chosen for FTF implementation.

Figure 20: Predicted Discipline Referrals, by FTF Status, 2006–2007 and 2007–2008



*Source.* AISD student records, prepared by the DPE, October 2008

*Note.* Estimates were derived from a negative binomial regression with robust standard errors and unit-specific effects.

\*  $p < .01$ , two-tailed test

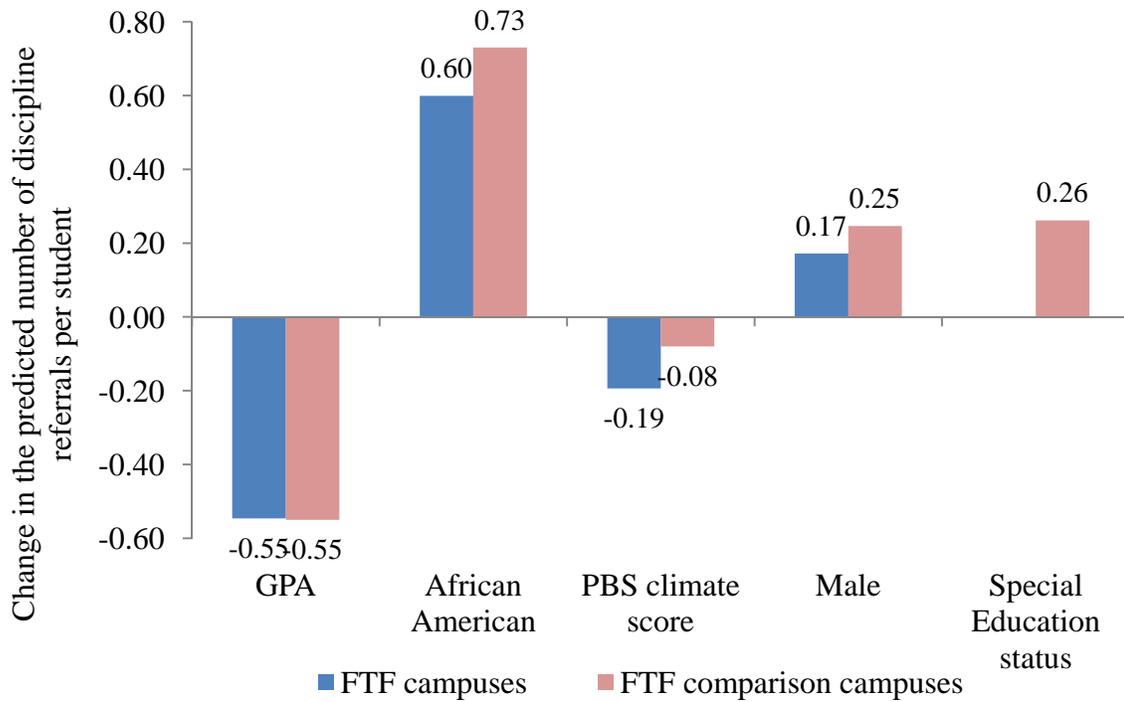
Several other control variables exerted sizeable and statistically meaningful effects on the number of discipline referrals, and differed across FTF and FTF-comparison campuses. At FTF-comparison campuses, African American students received .73 more discipline referrals per student than did Hispanic students; at FTF campuses, this disparity narrowed to .60. Importantly, disproportionate discipline referrals for African American students have been identified and explored in other studies. According to findings by Skiba et al. (2002), referral disproportionality between ethnic groups does not dissipate after student-level characteristics are incorporated. Rather, over-referral of African Americans, their findings suggest, is driven by differences in school staffs' punitive responses to African American students' misconduct.

The AISD Positive Behavior Support (PBS) model establishes behavior support systems at different levels of student intervention to promote pro-social behaviors and to limit disruptive student behaviors. Although implementation has been concentrated among middle school campuses, six district high schools were designated to receive, and have accepted, PBS support. Even schools that have not received formal district support for PBS implementation may have incorporated many of the tenets and strategies of the model, or may have preexisting staff behaviors that are hallmarks of the PBS philosophy.

Students in schools with higher PBS implementation scores have lower per student discipline referral rates, and this negative relationship was stronger at FTF campuses than at FTF-comparison campuses. This climate measure, taken from the AISD Staff Climate Survey, captures the prevalence of staff reinforcement of desirable student behaviors, as well as the existence and frequency of desirable student behaviors on campus. Two of the three campuses selected for PBS implementation (Travis and Reagan) also were chosen for FTF intervention, which may help explain why the relationship is stronger among FTF campuses.

To summarize the multivariate results from the discipline referral analyses, both FTF and FTF-comparison campuses experienced dramatic declines in the percentages of students with discipline referrals and in the per student discipline referral rate from 2006–2007 to 2007–2008. At FTF campuses, high student attrition, particularly among students who had chronic disciplinary problems and who were predisposed to dropping out, may have contributed to these declines.

Figure 21: Predicted Discipline Referrals, by FTF Status and Student and School Characteristics, 2006–2007 and 2007–2008



*Source.* AISD student records and AISD Staff Climate Report, prepared by the DPE, October 2008

*Note.* Estimates were derived from a conditional negative binomial regression with robust standard errors and unit-specific effects. GPA and PBS indicate the change generated from a one standard deviation increase. African American, Male, and Special Education status denote the difference in number of referrals relative to a corresponding reference group. For African Americans, the reference group is Hispanics.

## DISCUSSION

The partnership between AISD and IRRE to implement the FTF framework in a selection of AISD high schools has shown promising results after the first year of implementation. Despite overall positive findings for FTF schools, some differences were found across and within schools. Thus, the findings in the report merit deeper investigation and contemplation.

The classroom observation protocol is a valuable tool for monitoring the improvement of instructional practices and student learning. Abiding by program implementation expectations, this tool was to be used consistently to provide timely information to teachers and administrators for instructional improvement purposes throughout the school year.

Although classroom observations took place frequently in all of the schools at the beginning of the year, they decreased significantly as the school year progressed. The factors causing the decreasing numbers of classroom observations were not explored. Commitment of campus and district administrators, time for completing the observations, ongoing training needs, and TAKS testing calendars may have influenced the frequency of observations. In addition to the decreasing numbers of observations throughout the year, little information was available about how classrooms were selected for the visits or whether individual classrooms were visited at similar intervals or frequencies. Without this information, it is difficult to discern whether the trends in engagement, alignment, and rigor were representative of the whole campus or SLC.

Because the family advocacy portion of the FTF initiative was fundamental in the development of student engagement and academic success, this evaluation examined the outcomes of the family advocacy component in relation to the relationships built between the students, their parents, and their family advocates. Each school established family advocacy classes that met regularly throughout the school year on their respective campuses. Students' reports that they interacted with their advocates outside of the scheduled class period indicate that relationships of varying degrees were built between the students and their advocates. Approximately one-third of students at Travis and LBJ claimed they had met with their family advocate at least one time outside of their advisory period, while 78% of respondents at Regan had met informally with their family advocate.

The outreach to and relationships with parents were variable across the campuses. Large differences were found between the percentages of students at LBJ and Travis, compared with the percentage of students at Reagan, who reported their family advocate had met with their parent or guardian. The factors influencing whether advocates met with parents were not explored. Some teachers reported that parent contact information was not up to date. Additionally, expectations about conducting parent conferences may have differed at the respective campuses. For instance, conferences may have been prioritized according to student need. Furthermore, an advocate's time for parent conferencing may have been limited and prohibited conferencing with all of his or her students' families.

Notably, the evaluation of the family advocacy component of the FTF model was limited in scope within this report and only looked at the relationship between the student, parent, and advocate. Emphasis on parent outreach and on relationship development with parents differentiates this model from the district's student advisory program implemented across the other high schools in the district. A more comprehensive evaluation of the district's

student advisory/family advocacy program is provided in the report entitled *High School Redesign: Student Advisory Evaluation, 2007–2008* (Looby & Garland, 2008).

Standardized test performance improvements were evident at all of the FTF schools. The TAKS test outcomes were particularly impressive at Travis compared to other FTF schools, where the percentage of students who met the math TAKS standard increased significantly in 2007–2008. Facing state intervention due to poor statewide assessment performance, Travis staff sought a wide range of innovative and intensive instructional strategies and other types of reforms to increase student achievement. A math instructional specialist at Travis described the efforts of their PLCs, which met after school to review students' math performance and discuss and design activities or strategies for improved student learning. They received strong district and campus administration support and were compensated for the extra time spent after school. Although these departmental meetings were not explicitly prescribed by FTF, such innovations that emerge organically at campuses can play a vital, complementary role in increasing student preparation and achievement.

Discipline referral rates showed similarly notable improvements in two schools between the 2006–2007 and 2007–2008 school years. The reductions identified at Travis and Reagan were in sharp contrast with the increase in discipline referrals found at LBJ. Because the referral rates at LBJ differed greatly from rates at the other two schools, further investigation of the patterns and types of referrals reported is needed to determine whether certain types of referrals influenced these changes across the three schools. To determine whether these changes in referral rates were detectable to students, student survey responses may yield additional insight into how student perceptions of the behavioral environment on their campus corresponded with changes in referral patterns.

## CONCLUSION AND RECOMMENDATIONS

The FTF initiative includes three major components: SLCs for students; a family advocacy system; and instructional improvement efforts focused on student engagement, curricular alignment, and rigorous instruction. In 2007–2008, the LBJ, Reagan, and Travis high schools engaged in a full-scale implementation and experienced positive results. With ongoing support provided by district and school administrators, implementation can continue to progress. Recommendations for continuing support and improvement are provided for consideration.

1. *Ensure a lasting commitment to conducting EAR classroom observation visits throughout the school year to ensure sustained instructional improvement and to meet student achievement goals.* Across FTF campuses, district and campus staff demonstrated admirable enthusiasm early in the 2007–2008 school year for conducting EAR classroom visits. District- and campus-level administrators may need to articulate expectations regarding the frequency of the observations and the use of the observation data to improve instruction and learning. Both district- and campus-level observers, representing administrative and curriculum offices, should complete the observations to guard against possible bias issues. This will facilitate a more even distribution of responsibility for conducting classroom visits, while also counteracting any systematic data collection errors or biases that may be introduced when observations are disproportionately conducted by only a few campus staff members. Distributing this responsibility will also serve as a hedge against future attrition among campus staff assigned observation duty which may contribute to the sustainability and institutionalization of the EAR protocol tool. Furthermore, administrators need to provide the resources required (e.g., time, training, and technology) to conduct the observations.
2. *Although the self-reported frequency of student, parent/guardian, and family advocate conferences was comparable to those reported in national evaluations after the first year of implementation, campus and district stakeholders must continue to monitor and assist family advocates in performing this essential duty.* Currently, campuses do not have a formal, centralized data collection system to track family advocates' fidelity to this objective. In the absence of this type of collection tool, it is both difficult to

monitor compliance and unfeasible to assess in a systematic manner how these interventions affect individual student outcomes.

3. *Family advocates must continue to cultivate close, durable ties with their advisees.* Students who felt comfortable discussing personal and academic issues with their advocates were more likely to interact frequently with their advisor than were those who were uncomfortable. This is a pillar of the FTF initiative, and the effectiveness and success of the reform effort is greatly enhanced by these relationships.
4. *The collection and utilization of teacher, student, and parent feedback should be expanded to inform the continuous improvement process.* Student response rates for the FTF survey differed noticeably across campuses. Campus staff should strengthen efforts to improve survey response performance in subsequent years. The low response rates, coupled with the small percentage of respondents who provided their unique student identification number, prevented the linking of survey response information to student demographic and academic data. Future survey administrations should explore the feasibility of a web survey with pre-programmed student identification numbers to ensure survey responses can be connected to other district data sources. Additionally, teacher- and parent-specific surveys and focus groups representing all stakeholder groups should be utilized to explore perceptions of these groups to the initiative in relation to outcomes for students.
5. *The program evaluation should explore teacher planning and professional learning activities taking places within the SLC structures.* In the FTF model, teachers are expected to engage in planning and professional learning within their interdisciplinary SLCs and within their departmental groups. At this time, little information is collected regarding how the FTF school staff facilitate these processes for both interdisciplinary and content-focused groups.

**REFERENCES**

- Connell, J. P. (2002). *Getting off the dime: First steps toward implementing First Things First*. Institute for Research and Reform in Education. Retrieved November 25, 2008, from [http://www.irre.org/publications/pdfs/getting\\_off\\_the\\_dime.pdf](http://www.irre.org/publications/pdfs/getting_off_the_dime.pdf)
- Connell, J. P., & Broom, J. (2004). *The toughest nut to crack: First Things First (FTF) approach to improving teaching and learning*. Institute for Research and Reform in Education. Retrieved October 16, 2008, from [http://www.irre.org/publications/pdfs/the\\_toughest\\_nut\\_to\\_crack.pdf](http://www.irre.org/publications/pdfs/the_toughest_nut_to_crack.pdf)
- Ekstrom, R. B., Goertz, M.E., Pollack, J.M., & Rock, D.A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record*, 84 (3), 356-373.
- Institute for Research and Reform in Education (IRRE) (2003). *Kansas City, Kansas School District Trends on the Kansas State Assessment Since District-Wide Implementation of First Things First: 2001 to 2003*. Institute for Research and Reform in Education. Retrieved December 1, from [http://irre.org/publications/pdfs/KCKSD\\_trends.pdf](http://irre.org/publications/pdfs/KCKSD_trends.pdf)
- Klem, A. M., Levin, L., Bloom, S., & Connell, J. P. (2003). *First Things First's family advocate system: Building relationships to support student success in secondary schools*. Institute for Research and Reform in Education. Retrieved October 16, 2008, from [http://www.irre.org/publications/pdfs/building\\_relationships.pdf](http://www.irre.org/publications/pdfs/building_relationships.pdf)
- Looby, K. & Garland, M.W. (2008). *High School Redesign: Student Advisory Evaluation, 2007–2008*. (DPE Publication No. 07.72). Austin, TX. Austin Independent School District Department of Program Evaluation.
- Quint, J., Bloom, H. S., Black, A. R., Stephens, L., & Akey, T. M. (2005). *The challenge of scaling up educational reform: Findings and lessons from First Things First*. New York: MDRC. Retrieved October 21, 2008, from <http://www.mdrc.org/publications/412/full.pdf>
- Quint, J. C., Byndloss, D. C., & Melamud, B. (2003). *Scaling up First Things First: Findings from the first implementation year*. New York: MDRC. Retrieved October 21, 2008, from <http://www.mdrc.org/publications/363/full.pdf>
- Schmitt, L.N.T., Bush-Richards, A., & Cornetto, K. (2008). *Summary of 2005-2006 Through 2007-2008 AISD Staff Climate Survey Results*. (DPE Publication No. 07.23). Austin, TX. Austin Independent School District Department of Program Evaluation.
- Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, 34 (2), 317–342.

### APPENDIX A: FTF STUDENT SURVEY PROFILE AND SAMPLING METHODOLOGY

Although the entire population of students enrolled at LBJ, Travis, and Reagan during the Spring 2008 semester were sampled, the composition of survey respondents at each school was not representative of the respective school's enrollment. Thus, non-response or incomplete survey response was correlated with student-level characteristics, including grade level. For instance, 1,085 students successfully completed the survey, and 939 respondents accurately indicated their grade level, but only 755 provided a valid student identification number that would permit linking the responses to other district data sources. The patterns of non-response resulted in the underrepresentation or overrepresentation of particular student populations. For instance, although 12<sup>th</sup> grade students comprised 21% of the Travis student body when the survey was administered, only 11% of Travis respondents classified themselves as a 12<sup>th</sup> grade student. To correct for this lack of representativeness, poststratification weights by grade-level populations within each school sampled were applied. The poststratification weight was equal to the inverse of the probability of being selected as a result of the sampling procedure. More formally,  $W_{ij} = N_{ij}/n_{ij}$ , where  $W_{ij}$  = the probability weight,  $N_{ij}$  = the population of students within each grade level, by school, and  $n_{ij}$  = the total number of survey respondents within each grade level, by school. Table A. 1 presents the response rates for each FTF campus.

Table A 1. FTF Survey Sample and Response Rates, Spring 2008

	Valid Responses	Enrollment	Response Rate
Reagan	441	829	53.20%
Travis	398	1,300	30.62%
LBJ	246	841	29.25%
Totals	1,085	2,970	36.53%

Source. AISD student records and FTF survey, prepared by the DPE, October 2008

**APPENDIX B: SELECTION OF COMPARISON CAMPUSES**

Several methods were used to select comparison campuses for the FTF evaluation analyses. First, and least rigorously, schools were identified that resembled FTF campuses on an assortment of demographic and achievement indicators, including ethnic composition, high-needs population, and TAKS performance. Schools must have had a student body comprising at least 80% economically disadvantaged students, and an average 2006–2007 math TAKS score below .5 standard deviations from the overall mean. This method identified Lanier and Johnston high schools as the campuses most similar to the FTF schools.

Second, and more rigorously, a variant on propensity score matching (PSM) was used to quantify the multiple school and student-level characteristics associated with selection for FTF intervention. More precisely, a multivariate logistic regression was estimated to determine the conditional probability that a given student at a particular school was enrolled at a FTF campus. These probabilities then were aggregated up to the campus level. Confirming the results from the first, more rudimentary procedure, Lanier and Johnston were found to be most like the FTF campuses, after controlling for student-level demographic and academic characteristics.

**APPENDIX C: TECHNICAL MATERIAL FOR TAKS MULTIVARIATE ANALYSES SECTION**

Because we were investigating the change in student performance that stemmed from the introduction of the FTF reform model, compared with performance at similar schools that did not implement the initiative, the improvement attributable to FTF was captured by the change in students' TAKS performance—or the change in any student outcome measures—between 2006–2007 and 2007–2008. This assumed that all other confounding school-level and student-level factors that could influence student academic performance were included. The impact of FTF implementation on student performance was represented by the inclusion of a school-level dummy variable that was coded as “1” for the 2007–2008 school year and “0” for 2006–2007. If this implementation variable was statistically significant for a FTF school, the improvement resulting from FTF intervention was considered to be meaningful and not generated by sampling or measurement anomalies.

The outcome of interest was whether an individual student satisfied the minimum standard for the TAKS subject area in a given school year. Thus, the dependent variable was binary, assuming values of “1” if a given student met the standard and “0” if the student did not. Limited dependent variables necessitate econometric techniques that adjust for the non-continuous and non-linear structure of the dependent variable. Adopting standard ordinary least squares (OLS) regression techniques with a limited dependent introduces numerous statistical violations and jeopardizes the researchers' ability to extract sound inferences from the statistical results. To avert these dangers, maximum likelihood estimation (MLE) techniques (i.e., logistic regression) were used to produce estimates of the impact of student and school-level characteristics on the likelihood a student met the TAKS standard. These logit estimates then were converted to predicted probabilities to ease the interpretation of the results.

**APPENDIX D: TECHNICAL MATERIAL FOR THE DISCIPLINARY ANALYSIS SECTION**

To investigate whether discipline referrals showed greater improvements at FTF schools than at FTF-comparison campuses, student discipline records were extracted from district data systems (ADIS) and collapsed down to the individual student level. The total number of referrals in a given school year was calculated for each student, ignoring the severity and type of offense prompting the referral. School-level referral rates by school year were calculated by dividing the total number of referrals received by students at a particular campus by the total number of students enrolled. All referrals recorded in the district's student disciplinary system were reported, including those classified as discretionary.

For the multivariate analyses, the student-level structure of the data was retained. Thus, disciplinary referral records were not aggregated up to the campus level using the procedure described in the previous paragraph. Moreover, because the dependent variable was a count of the number of referral incidences each enrolled student received, standard OLS regression was an inappropriate statistical technique. With a severely abnormally distributed dependent variable that was zero-inflated, and an unconditional variance that exceeded the unconditional mean, multivariate count models was prescribed. Moreover, given the overdispersion and zero-inflated nature of the dependent variable, a negative binomial estimation was chosen. Statistical tests to determine the suitability of this model rather than a zero-inflated negative binomial concluded no important differences existed between the estimation results derived from the different techniques.

**AUSTIN INDEPENDENT SCHOOL DISTRICT**

**SUPERINTENDENT OF SCHOOLS**

Pascal D. Forgione, Jr., Ph.D.

**OFFICE OF ACCOUNTABILITY**

Anne Ware, Ph.D.

**DEPARTMENT OF PROGRAM EVALUATION**

Holly Williams, Ph.D., Director

**AUTHORS**

Marshall W. Garland, MA.

Karen Looby, Ph.D.



**BOARD OF TRUSTEES**

Mark Williams, President

Vincent Torres, M.S., Vice President

Lori Moya, Secretary

Cheryl Bradley

Annette LoVoi, M.A.

Christine Brister

Robert Schneider

Karen Dulaney Smith

Sam Guzman