MEMORANDUM September 6, 2018

TO: Mechiel Rozas
Secondary Literacy Director

FROM: Carla Stevens
Assistant Superintendent, Research and Accountability

SUBJECT: **ACHIEVE3000®, 2017–2018**

The 2017–2018 academic year was the fourth year that the Houston Independent School District offered the differentiated reading instruction program, Achieve3000®. While in previous years the program was offered primarily to 9th and 10th grade students, in 2017–2018 it was expanded and made available to all secondary students, or about 100,000 6th through 12th grade students in 98 HISD schools. Only about 17,000 students in grades 6–10 were included in the attached evaluation, which assessed the relationship between Achieve3000 activity completion (i.e., usage) and students' end-of-year STAAR Lexile level.

Key findings include:

- Ignoring potential control factors, schools that had a mean number of activities completed below the district mean number of activities completed had a **HIGHER** end-of-year Lexile level than schools that had a mean number of activities complete at or above the district mean number of activities completed.
- Of the 96,698 students who had access to Achieve3000 in the 2017–2018 academic year, just 107 students completed the recommended 80 activities or more.
- Among middle school grade students (i.e., those in grades 6–8), and controlling for a number of individual-level factors, there was an approximate 1 percent increase in the odds of being on track for college and career readiness as usage of Achieve3000 increased.
- Among high school grade students (i.e., those in grades 9 and 10), and controlling for a number of individual-level factors, there was an approximate 1.5 percent increase in the odds of being on track for college and career readiness as usage of Achieve3000 increased.
- Blacks and males had a lower probability at both school levels (i.e., middle and high school) of being on track for college and career readiness relative to whites and females, respectively.
- There were no statistically significant differences in end-of-year Lexile performance between whites and Hispanics or between whites and Asian/other race students.

Further distribution of this report is at your discretion. Should you have any further questions, please contact me at 713-556-6700.

Carla J. Stevens CJS

Attachment
cc: Noelia Longoria
Annie Wolfe

HOUSTON INDEPENDENT SCHOOL DISTRICT

RESEARCH

Educational Program Report

ACHIEVE3000 EVALUATION 2017–2018

HISD | Research and Accountability
ANALYZING DATA, MEASURING PERFORMANCE.

2018 BOARD OF EDUCATION

Rhonda Skillern-Jones
President

Jolanda Jones
First Vice President

Anne Sung
Second Vice President

Sergio Lira
Secretary

Holly Maria Flynn Vilaseca
Assistant Secretary

Wanda Adams
Diana Dávila
Susan Deigaard
Elizabeth Santos

Grenita Lathan, Ph.D.
Interim Superintendent of Schools

Carla Stevens
Assistant Superintendent
Department of Research and Accountability

D. Diego Torres, Ph.D.
Research Specialists

Jessica Vasan
Research Manager

Houston Independent School District
Hattie Mae White Educational Support Center
4400 West 18th StreetHouston, Texas 77092-8501

www.HoustonISD.org

HISD | Research and Accountability
ANALYZING DATA, MEASURING PERFORMANCE.

EVALUATION REPORT

BUREAU OF PROGRAM EVALUATION

The Relationship between Achieve3000® Usage and HISD Students' College and Career Readiness in Grades 6–10

Prepared by D. Diego Torres, Ph.D.

Abstract

Achieve3000®, a web-based differentiated reading program based on the Lexile® Framework and designed to improve student reading ability and comprehension of increasingly complex texts, was made available to approximately 100,000 secondary students across 98 HISD schools during the 2017–2018 academic year. The present evaluation, using a select subsample of these data, examined the relationship between Achieve3000 usage throughout the year and students' end-of-year State of Texas Assessments of Academic Readiness (STAAR) Lexile level and students' college and career readiness. The research questions were (1) whether schools with mean completion of Achieve3000 activities completed at or above the district average had a higher mean end-of-year Lexile level than schools with mean completion of activities below the district average, and (2) whether, net of the effects of student demographic factors and between-school variance, greater usage of Achieve3000 associated with a higher probability of ending the year on track for being college and career ready as determined by the end-of-year Lexile level, and whether any difference varies by race/ethnicity. Regarding question 1, findings suggest that the mean end-of-year Lexile level was not higher in schools where the mean number of activities was at or above the district average, but was instead higher in schools where the mean number of activities was below the district average. Regarding question 2, findings point to a modest relationship between Achieve3000 usage and students' probability of being on track for college and career readiness. For an increasing number of activities completed, there is an associated rise in the predicted probability of being on track for career and college ready, holding constant other predictors. Black students have a lower odds of being on track for college and career readiness compared to their white, Hispanic, and Asian and other race peers. Implications of the findings are discussed and recommendations for future use are given to help maximize implementation of Achieve3000.

Background

It is estimated that by 2020, 65 percent of all jobs in the U.S. economy will require at least some postsecondary education (Carnevale, Smith, & Strohl, 2013). In the fastest growing occupations–i.e. STEM, healthcare professions and support, and community services–upwards of nearly 90 percent of jobs will require a postsecondary degree, while 63 percent will specifically require a bachelor's degree or higher (Carnevale et al., 2013; Lockard & Wolf, 2012). Unfortunately, the college and career readiness rates of U.S. students are not rising sufficiently to meet the demands of either postsecondary education or the workforce (Almond, 2017; DePaoli, Balfanz, Bridgeland, Atwell, & Ingram, 2017). Data from the class of 2017 reveals that, nationally, only 39 percent of 12th graders taking the ACT test met the minimum score in 3 of 4 subjects (English, reading, math, and science) necessary to achieve a high probability of success in a first-year core college course (ACT, 2017). Forty-six percent of class of 2017 students taking the SAT met the minimum score to ensure a 65 percent probability of earning a B- or higher grade point average during the first year at a four-year university (College Board, 2017).

As non-Asian minority students face greater barriers to being college and career ready due to their lower likelihood of taking rigorous or advanced placement (AP) courses (U.S. Department of Education, 2014), among other factors, this crisis is especially concerning for them. Only 12 percent of 2017 ACT-tested black high school graduates and

23 percent of ACT-tested Hispanic high school graduates met three or more benchmarks necessary to achieve a high probability of success in a first-year college course; percentages for whites and Asians and Pacific Islanders were 50 and 62, respectively (ACT, 2017). Corresponding percentages for college and career readiness on the SAT for the class of 2017 were 20 percent for blacks and 31 percent for Hispanics, compared to 59 percent for whites and 70 percent for Asians and Pacific Islanders (College Board, 2017).

These sobering facts notwithstanding, the majority of high school students, including those with bad grades, low test scores, and poor high school attendance, believe they are going to college (Rosenbaum, 2004). Many, in fact, do. Because the level of work demanded in college tends to be much more rigorous than that taught in high school, however, the least prepared students are often required to complete remedial courses before taking more challenging courses (Deil-Amen & Rosenbaum, 2002). Few students with a C average or lower will go on to complete an associate's degree, much less a bachelor's degree, within eight years of graduating high school (Rosenbaum & Gordon-McKeon, 2003). Indeed, among students with a high school average of C or lower, it is less likely that they'll achieve even one college credit than not (Rosenbaum, 2001).

Reading Ability is Essential to College and Career Readiness

Research by Williamson (2008) has shown that high remediation rates in college and poor college graduation statistics can be attributed in large part to the gap in the reading expectations and text complexity between high school and college. While the reading demands of college and the workplace have held steady or increased over the past half century (Stenner, Koons, & Swartz, 2010), K–12 reading texts have decreased in difficulty (Adams, 2009; Hayes, Wolfer, & Wolfe, 1996). The result has been a 350L (Lexile) gap between the difficulty of 12th grade and college texts, or the equivalent of about two to three years (Williamson, 2006).

Because reading is an essential component of college and workplace success, failure to achieve high levels of reading comprehension, or being unable to read complete texts independently and proficiently, necessarily leads to failure or underperformance in both. In college, specifically, poor readers often struggle in text-heavy courses, preventing them from not only being able to enroll in more challenging courses, but from completing a degree altogether (Au, 2000; Wirt, Choy, Rooney, Provasnik, Sen, & Tobin, 2004). The Common Core State Standards and broader college and career readiness agenda aim to capture the range of text complexity required in the workplace and in college (Wei, Cromwell, & McClarty, 2016). Higher level metacognitive reading abilities than those resulting from high school preparation are required, often in combination with computational skills, to fully comprehend and effectively put to use material printed in many workplace technical manuals (Autor, Levy, & Murnane, 2003; Levy & Murnane, 2004)

The general decline in text complexity at the end of high school relative to the demands of college and the workplace highlights the need for teachers well-trained in literacy instruction. Effective knowledge transmission in the classroom has a direct and positive relationship with student outcomes (McCombes-Tolis & Feinn, 2008). Unfortunately, research has shown that teachers possess insufficient knowledge about reading processes to effectively facilitate readers' development (Allington & McGill-Franzen, 2008; McCombes-Tolis & Feinn, 2008; Washburn, Joshi, & Cantrell, 2011), often as a consequence of poor university preparation (Joshi, Binks, Graham, Ocker-Dean, Smith, & Boulware-Gooden, 2009; Joshi, Binks, Hougen, Dahlgren, Ocker-Dean, & Smith, 2009).

Part of the problem of delivering instruction aimed at helping students achieve higher levels of reading ability centers on the fact that there is a great variability in text complexity at any grade level (Stenner, Koons, & Swartz, 2010). This means that developing and advanced readers–with their broad experience, domain knowledge, and vocabulary–are in a position to be challenged more and see gains when they are. Struggling readers–lacking the requisite understanding of more complicated text structures and strategies to manage comprehension–are not in a position to be challenged too far beyond where they are developmentally, and tend to withdraw from the educational experience when they are.

Struggling readers are most likely to benefit from differentiated reading instruction (i.e., matching students with texts suited to their developmental level, or slightly more challenging) and from strategies that target students' identified weaknesses and strengths (Spear-Swerling, 2011). To the degree that success in learning requires the ability to cope with occasional frustration and failure, some level of challenging instruction in reading is important, and, indeed, research has shown that the use of more complex texts might result in better long-term progress for struggling readers (Shanahan, Fisher, & Frey, 2012).

To be sure, the importance of raising the reading ability of K-12 students cannot be overstated, particularly as it impacts their college and career readiness, and their general preparation for life beyond high school. As noted in a report from the Common Core State Standards Initiative (NGA Center & CCSSO, 2010):

> A turning away from complex texts is likely to lead to a general impoverishment of knowledge, which, because knowledge is intimately linked with reading comprehension ability, will accelerate the decline in the ability to comprehend complex texts and the decline in the richness of text itself. This bodes ill for the ability of Americans to meet the demands placed upon them by citizenship in a democratic republic and the challenges of a highly competitive global marketplace of goods, services, and ideas. (p.4)

Improving Student Achievement through Technology-Based Resources

A policy brief from the International Society for Technology in Education® (ISTE, 2008) maintains that the integration of education technology in traditional instructional delivery gives students an advantage not just in their knowledge of core academic areas, but also in the skills necessary to be productive in higher education and the workplace, namely the abilities to "communicate, collaborate, analyze, create, innovate, and solve problems" (p.4). Such skills are transmitted in various ways, depending on the subject matter. Typically, for reading, helping students achieve readiness beyond the years of compulsory education involves imparting the ability to comprehend difficult prose text. This process begins by taking students where they are developmentally and growing them from there by matching them to texts whose readability is at or slightly above their current individual levels.

Technology-based resources, such as Achieve3000®, now in its fourth year in use by the Houston Independent School District (HISD), help administrators and educators identify students who are behind where they should be educationally, providing information regarding their placement in tiered instruction. Rooted in the Lexile® Framework for Reading, Achieve3000 is a leader in online differentiated literacy instruction, the aim of which is to measure student progress and forecast their development. A strategic curriculum aligned to Texas's grade-level standards for language arts, Achieve3000 also aims to assist students stay on track throughout K–12 for readiness after formal education where readiness is defined as being prepared to meet the rigorous demands of higher education, the workplace, as well as the military.

To help students achieve readiness at each grade, Achieve3000 first administers an assessment that establishes a baseline Lexile level. Students then receive differentiated, grade-appropriate, non-fiction adaptive reading passages that are aligned to that Lexile level. Subsequent passages are determined based on the adjusted Lexile level based on the prior passages end-of-lesson assessments, and, when progress is detected, text rigor is increased (Achieve3000, 2017). Teachers can be instrumental in further accelerating the literacy gains of their students by following students' independent work with direct instruction around Achieve3000's Stretch Articles. Stretch Articles give students the opportunity to boost their growth by engaging a version of the Lexile-aligned lesson written at the grade-appropriate text complexity. The basic idea of this approach is to (1) build students' stamina with independent work at their precise reading level and (2) build students' strength with close reading of grade-appropriate complex texts. It is an approach that has been proven to accelerate student learning.

In randomized controlled trials and studies utilizing quasi-experimental statistical methods to remove selection or omitted variables bias, students who used Achieve3000's differentiated learning tools outperformed their peers who had not used them. Borman, Park, and Min (2015) found that late elementary and middle school aged students who were assigned Achieve3000 outperformed those who did not use the tool by 2.4 points on the California State Test (CST). Shannon and Grant (2015) used a randomized controlled trial and found that students who used Achieved3000 regularly had, during the academic year, larger gains and performed better in both vocabulary and reading comprehension than their control group peers who were limited to a traditional English Language Arts curriculum. Previous evaluations of Achieve3000 have also shown it to be effective at improving the achievement of HISD 9th and 10th grade students (Torres, 2015, 2016).

Research Questions

For the 2017–2018 academic year, HISD's Secondary Curriculum and Development Department invested $2,000,000 to make Achieve3000 available to all of its students in grades 6–12. As outlined in the agreement between the district and Achieve3000, Achieve3000's differentiated literacy solutions included student licenses in both English and Spanish; teacher and

administrator licenses; parent licenses available in 17 languages; LevelSet™ assessments three times yearly; product training; implementation planning; principal awareness sessions; live online workshops; 24/7 self-paced, on-demand resources; reporting for teachers; and alignment to curriculum frameworks for science and social studies. Given this large monetary expenditure in a time of constrained budgets, it is worth assessing whether the extensive catalog of tools provided by that cost actually led to greater student achievement. To that end, the present study seeks to answer the following research questions.

1. Do schools with mean usage (i.e., number of activities completed) of Achieve3000 at or above the district average have a higher mean end-of-year Lexile level than schools with mean usage below the district average?
2. Net of the effects of student demographic factors and between school differences, is greater usage of Achieve3000 associated with a higher probability of ending the year on track for being college and career ready as determined by the end-of-year Lexile level, and does any difference vary by race/ethnicity?

While the high school graduation rate has recently edged up nationally, from 79 percent in 2011 to 84 percent in 2016, education experts have cautioned policymakers and educators about assuming that this positive change means that students are prepared for higher education and work. At the same time the graduation rate rose, students' performance on the National Assessment of Educational Progress (NAEP), often referred to as America's Report Card, remained relatively flat. As is typically the case, of course, the figures were lower, in either case, for disadvantaged and underserved populations compared to their white and Asian/Pacific Islander peers.

This evaluation focuses on whether students are on track for college and career readiness, and not graduation only, based on their end-of-year State of Texas Assessments of Academic Readiness (STAAR) 3–8 reading or End-of-Course (EOC) English Lexile level. Being college and career ready will–particularly with respect to reading ability, the gateway to more difficult, text-heavy courses and workplace technical material–ensure that HISD students are able to take advantage of the growing, dynamic, and knowledge-based Texas economy. A technology-based resource like Achieve3000 may not be able to make up for the general lack of rigorous courses available to, or taken by, non-Asian minorities, but it may help achieve parity for them in at least one domain if engaged on a regular basis. Focusing on Achieve3000's differential impact by race/ethnicity, then, is also important.

Data and Method

Sample

For the 2017–2018 academic year, Achieve3000 was made available to 96,698 students in 98 middle, combined, and high schools. While effective implementation requires a pretest window in the first thirty days of the school year (i.e., September), Houston ISD was impacted by Hurricane Harvey, which delayed school by more than two weeks for many students, thereby reducing a potential analytic sample to only 8,247 students in 63 schools. This analysis, however, allowed for a pretest window that extended into October, overcoming to some degree the negative impacts of the storm on school attendance and increasing the potential sample to 29,676 students across 89 schools.

Achieve3000 is maximized when students are encouraged to complete 80 activities over the course of the academic year, working through them on a regular basis. This study excludes students who had activity Lexile values recorded for fewer than two months, which suggests infidelity to that goal. Also excluded were students who were in 11th and 12th grades since they had no STAAR EOC English Lexile, and students who were absent from the previous year's STAAR 3–8 or STAAR EOC, given that previous year's reading or English performance constituted a baseline control. Only first-time test takers were included. The final analytic sample, then, consisted of 17,010 students (or about 18 percent of the students for whom Achieve3000 was accessible) nested in 73 schools.

Measures

Dependent Variables. This evaluation used students' Lexile level on the STAAR 3–8 reading (for grades 6–8) and STAAR EOC English I (for grade 9) and English II (grade 10) assessments as the main outcome variables. Levels were dichotomized according to the Achieve3000 performance standards such that students were categorized as either Not On Track or On Track for their specific grade (see **Table A1, p. 16**).

Independent Variables. According to Achieve3000, the completion of at least 80 activities in an academic year is correlated with higher Lexile levels compared with lower activity levels. By extension, that means higher activity levels should be related to a higher probability of being on track for college and career readiness by the end of the

academic year. The primary independent variable for the analyses here, then, was the number of Achieve3000 activities completed during the 2017–2018 academic year.

To avoid overestimating the relationship between activity level and students' probability of college and career readiness, the analyses here had as a baseline control students' previous year's performance on the STAAR 3–8 reading (for students in grade 6–9) and STAAR EOC English I (for students in grade 10) assessments.

Additional student-level controls were entered for race/ethnicity, gender, economic disadvantage, English Language Learner (ELL) status, special education status, magnet student status, gifted/talented student status, homeless student status, and displaced student status. There were three race/ethnicity variables, each of which was dummy coded: Black (0 = non-Black, 1 = Black), Hispanic (0 = non-Hispanic, 1 = Hispanic), Asian/Pacific Islander and other (0 = non-Asian/Pacific Islander, 1 = Asian/Pacific Islander). Whites served as the reference group. Gender was coded 0 if a student was female and 1 if a student was male. Economic disadvantage was determined by whether a student did not qualify for reduced or free lunch or receive other family aid or services (coded 0) or did qualify for reduced or free lunch or received other family aid or services (coded 1). English Language Learner (ELL) status was coded 0 for non-ELL students and 1 for ELL students. Dummy variables were also included for whether a student was categorized as a special education student (0 = non-special education student, 1 = special education student), magnet student (0 = non-magnet student, 1 = magnet student), gifted/talented student (0 = non-gifted/talented student, 1 = gifted/talented student), homeless student (0 = non-homeless student, 1 = homeless student), or displaced by Hurricane Harvey (0 = not displaced student, 1 = displaced student).

Analytic Strategy

To answer the first research question, this study compared grade-level means in schools' Lexile level based on whether mean usage of Achieve3000 was below or at or above the district mean usage. A useful starting point for understanding the mere presence of a difference, this comparison, absent any controls, limits what may be extrapolated about the differences uncovered.

To understand more fully how Achieve3000 usage is related to student performance, and to answer the second research question, this evaluation regressed students' status of college and career readiness (0 = Not On Track, 1 = On Track) on Achieve3000 usage and other covariates (**Appendix B, p.17** provides a more complete explanation of the statistical method used here). Given that the data consist of students nested within schools, analyses also account for the amount of variation in college and career readiness that is due to differences between schools. Since the STAAR Lexile levels used as the main outcome come from two different STAAR tests (i.e., STAAR 3–8 and STAAR End-of-Course [EOC]), analyses were conducted separately for students in grades 6, 7, and 8, and students in grades 9 and 10.

Results

Summary Statistics

Table 1 (**p. 6**) presents summary statistics for the middle school-grades sample and the high school-grades sample to provide some context. As might be expected, the mean STAAR Lexile level and STAAR scale score are lower in the lower grades and higher in the higher grades, suggesting the cumulative impact of reading and English instruction from grade to grade. Students in grade 7, building on what they learned in grade 6, are expected to acquire more knowledge and skills such that their standing relative to current 6th graders is higher, and so on. Minimum college and career ready Lexile levels are shown for comparison.

Demographic factors are similar across the two school levels, with a few exceptions. More middle school grade students in the sample were gifted and talented, English language learners, and displaced due to Hurricane Harvey, while more high school grade students in the sample were enrolled in career and technical education (CTE) courses or homeless at some point during the academic year.

Turning to the main predictor in this study, i.e., number of activities completed, the means for both middle school and high school students was far below what is suggested by Achieve3000 to maximize its impact. While it is suggested that students complete at least 40 activities per semester, or at least 80 activities per academic year, the mean number of activities completed by middle school and high school students was, respectively, about 15 and 18 for the academic year, or less than 20 and 25 percent of the recommended usage. Of the 96,698 students who had access to Achieve3000 in the 2017-2018 academic year, just 107 students completed the recommended 80 activities or more.

School-Level Usage

The first research question asked whether schools with mean usage at or above the district

Table 1. Summary Statistics by School Level.

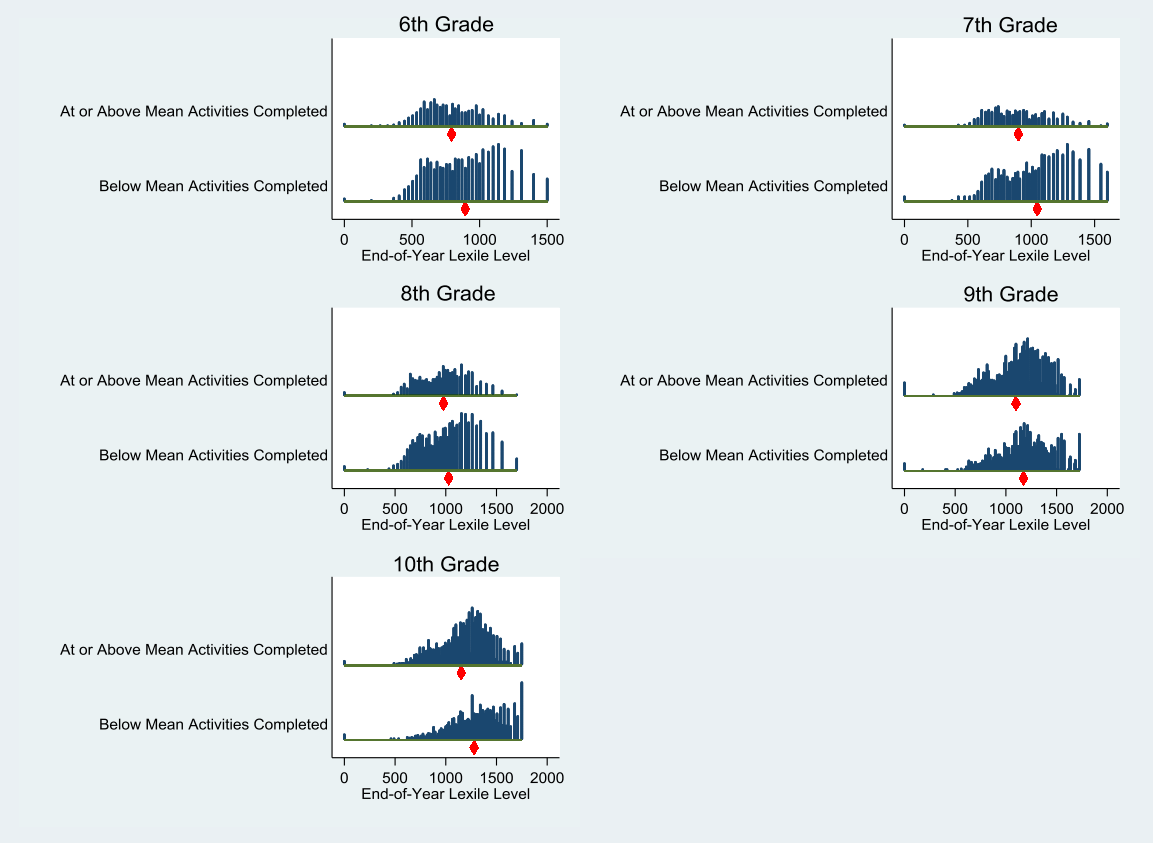
	Middle School Students (N = 7,129)		High School Students (N = 9,881)		College and Career Ready Minimum Lexile Level†
	Mean	SD	Mean	SD	
End-of-Year STAAR Lexile Level					
Grade 6	865.2	248.3			925L
Grade 7	1010.7	274.1			970L
Grade 8	1011.8	249.1			1010L
Grade 9			1133.0	271.9	1050L
Grade 10			1210.4	266.6	1080L
Prior Years' STAAR Scale Score					
Grade 6	1518.6	150.9			
Grade 7	1539.5	156.0			
Grade 8	1575.5	128.9			
Grade 9			1772.4	516.7	
Grade 10			3872.5	524.6	
# of Activities Completed	14.6	15.7	17.6	15.7	
Race/Ethnicity					
%White	5.3	22.5	4.5	20.7	
%Black	22.1	41.5	23.4	42.3	
%Hispanic	68.8	46.4	69.2	46.2	
%Asian/Other Race	3.8	19.2	2.9	16.7	
Sex					
%Female	49.2	50.0	49.2	50.0	
%Male	50.8	50.0	50.8	50.0	
%Economically Disadvantaged	87.2	33.4	82.5	38.0	
%Special Education	7.8	26.8	8.2	27.4	
%Magnet	26.6	44.2	26.2	44.0	
%Gifted/Talented	20.6	40.5	14.4	35.1	
%English Language Learner (ELL)	28.0	44.9	18.4	38.7	
%Career and Technical Education	10.1	30.1	88.8	31.6	
%Homeless During Year	0.3	5.0	2.0	13.9	
%Displaced by Hurricane Harvey	10.0	30.1	7.1	25.7	

†These minimums Lexile levels are included for quick reference and comparison. A full table with performance standards in the Lexile metric is included in Appendix A on page 16.

Note: Previous Years' STAAR Scale Score on the STAAR 3-8 is for Reading and on STAAR End-of-Course is for English I.

average had a higher mean end-of-year Lexile level than schools with mean usage below the district average. Based on the assumption of a positive correlation between the number of activities completed and Lexile performance, the answer should be affirmative. As **Figure 1** (**p. 7**) shows, however, that is not the case. Schools with mean number of activities completed below the district mean number of activities completed had a higher end-of-year Lexile level than schools with mean number of activities completed at or above the district mean number of activities completed. This held across each of the five grades examined in this study, grades 6–10. The differences in means were not only statistically significant, as **Table C1** (**p. 18**) shows, but so were the effect sizes.

Figure 1. Mean End-of-Year Lexile by Status on Mean Number of Achieve3000 Activities Completed (At or Above District Mean OR Below District Mean) by Grade.



It is important to keep in mind that this finding does not take into account mitigating factors that might yield a different result. Such mitigating factors could be attributable to mean differences in the characteristics that motivate students from dissimilar student bodies to utilize a district program that is not captured in simple summary data. Teacher and school effects, particularly with respect to implementation, are also likely correlated with Achieve3000 usage, such that these could also serve to attenuate the findings in Table C1. It is no doubt clear from this that analyses must go beyond mere descriptive statistics to fully understand the association between Achieve3000 usage and students' end-of-year STAAR Lexile level. Net of a host of demographic and other factors, is Achieve3000 usage associated with higher end-of-year Lexile level performance?

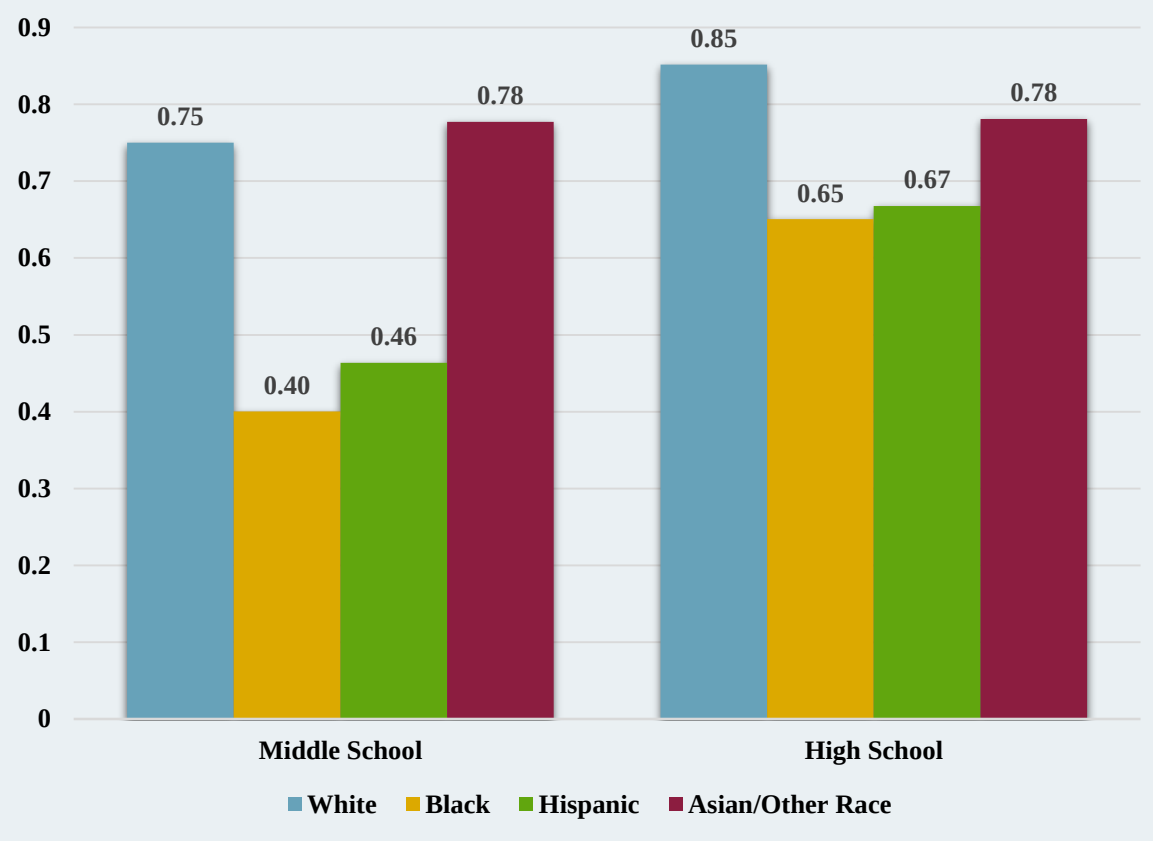
The Association between Achieve3000 Usage and College and Career Readiness

Before turning to the logistic regression models examining the relationship between Achieve3000 usage and college and career readiness, it is important to first highlight racial/ethnic group differences in end-of-year college and career readiness by school level. **Figure 2** (**p. 8**) reveals that, among students in the middle school grades (i.e., grades 6–8), the proportion of white and Asian and other race students who were college and career ready based on their Lexile level was about 3 out of 4, while the corresponding proportions for blacks and Hispanics were 2 out of 5 and approximately 1 out of 2, respectively. High school-age students appeared to be better prepared with nearly 9 in 10 whites and 8 in 10 Asians and other race students ready for college or career. Two thirds of both black and Hispanic 9th and 10th graders were college and career ready based on their Lexile level.

Given the well-known gap in academic achievement, as well as the concomitant lower likelihood of non-Asian minorities to have had rigorous educational opportunities, the racial/ethnic differences in college and career readiness proportion shown in Figure 2 were not unexpected. They do, however, spotlight potential deficiencies that can be targeted for amelioration using differentiated reading instruction like that offered in Achieve3000.

Table D1 (**p. 19**) and **Table D2** (**p. 20**) present, for middle school grades (i.e., grades 6–8) and high school grades (i.e., grades 9 and 10), respectively,

Figure 2. Proportion of College and Career Ready Students by School Level and Race/Ethnicity as Measured by Lexile Level.



results from the regression models. Panel A under Model 1 in Table D1 shows that the overall proportion of middle school-grade students who were on track for being college ready at the end of the 2017–2018 academic year was 0.477, or approximately 48 percent (see Appendix B for how to calculate the prediction of the proportion). Panel B under Model 1 in Table D1 shows both the mean proportion of college and career ready students across schools and the amount of variation in college and career readiness that is due to differences between schools. From the estimates in this model, the marginal probability of being on track for college and career readiness was 0.575, or roughly 58 percent, among Achieve3000 users in grades 6–8 (see Appendix B for how to calculate the marginal probability from the random effects model). A comparison of these two models suggest significant between-school differences in college and career readiness. Based on the estimate of ρ, students in most schools are more like their same-school peers than they are unlike them, and there is a decent amount of difference between schools.

Panel A under Model 2 in Table D1 added covariates for student-level demographic and other factors, while Panel B under Model 2 added the same, but allowed for random intercepts. Both models provided similar estimates of the covariate effects. Panel B in Model 2, however, showed a small 12 percent reduction in the variation in school-level unobserved heterogeneity over Panel B in Model 1. This suggests that there could be unobserved school-level measures that are associated with the school-level random intercept that might account for the relationship this study is examining. A comparison of all models suggested the random intercept model with covariates was the better fit.

The number of Achieve3000 activities completed, previous years' STAAR scale score, magnet status, gifted and talented status, and career and technical education course taking were all positively associated with being on track for college and career readiness. Being black, relative to being white, and being male, relative to being female, were both negatively associated with being on track for college and career readiness, as were being identified as special education or English language learner. The relationships between being on track for college and readiness and being Hispanic and

Asian and other race, economically disadvantaged, homeless during the year, or displaced by hurricane Harvey were not statistically significant.

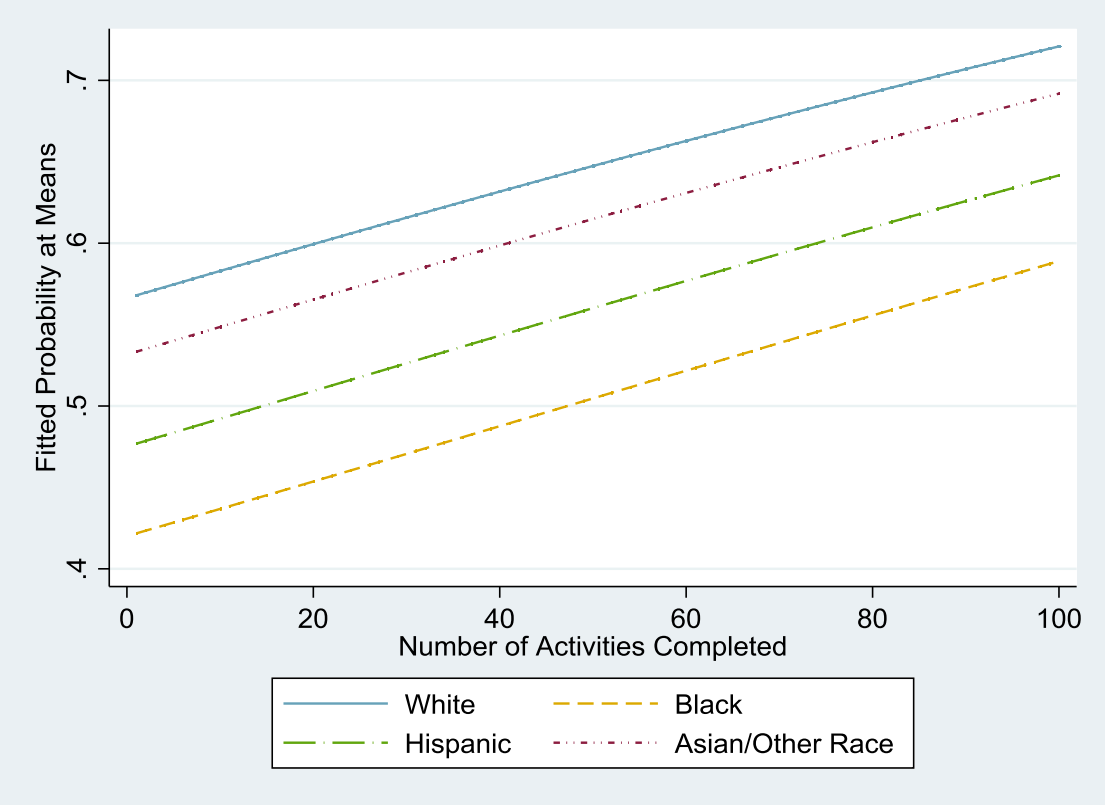
For each Achieve3000 activity completed, there was a 1 – exp(0.00682) = 0.0068, or about 1 percent increase in the odds of being on track for college and career readiness. Blacks had a 1 – exp(-0.589) = 0.4452, or 45 percent lower odds of being on track for college and career readiness than did their white peers. Whites, Hispanics, and Asian and other race students had about the same odds of being college and career ready.

Figure 3 presents, for middle school-grade students, a graphical representation of the predicted probabilities of being on track for college and career readiness by race/ethnicity based on the random intercept model with covariates. The positive relationship between Achieve3000 usage is evident here, with greater number of activities completed being associated with a higher probability of being on track for college and career readiness. The white-black gap in end-of-year readiness is large at both the low and high ends of usage; that is, there appears to be very little or no convergence in white and black college and career readiness as more activities are completed during the academic year.

Ignoring student-level demographic factors, the overall proportion of 9th and 10th graders who were on track for being college and career ready at the end of the 2017–2018 academic year was 0.675, or approximately 68 percent (Panel A under Model 1 in Table D2). Accounting for the amount of variation in college and career readiness due to school differences, which, according to ρ, was about 19 percent, the marginal probability for being college and career ready for these students was 0.616, or about 62 percent (Panel B under Model 1 in Table D2).

The standard full and random intercept models in Panels A and B in Model 2 revealed similar results for the covariate effects. The latter model, however, though it added largely individual-level measures, reduced the variation in school-level unobserved heterogeneity by just under 19 percent over Panel B in Model 1. Panel B in Model 2 provided a better fit over all other models in Table D2.

Figure 3. Predicted Probabilities of College and Career Readiness by Number of Achieve3000 Activities Completed and Race/Ethnicity from Panel B under Model 2 in Table D1, Grades 6, 7, and 8.

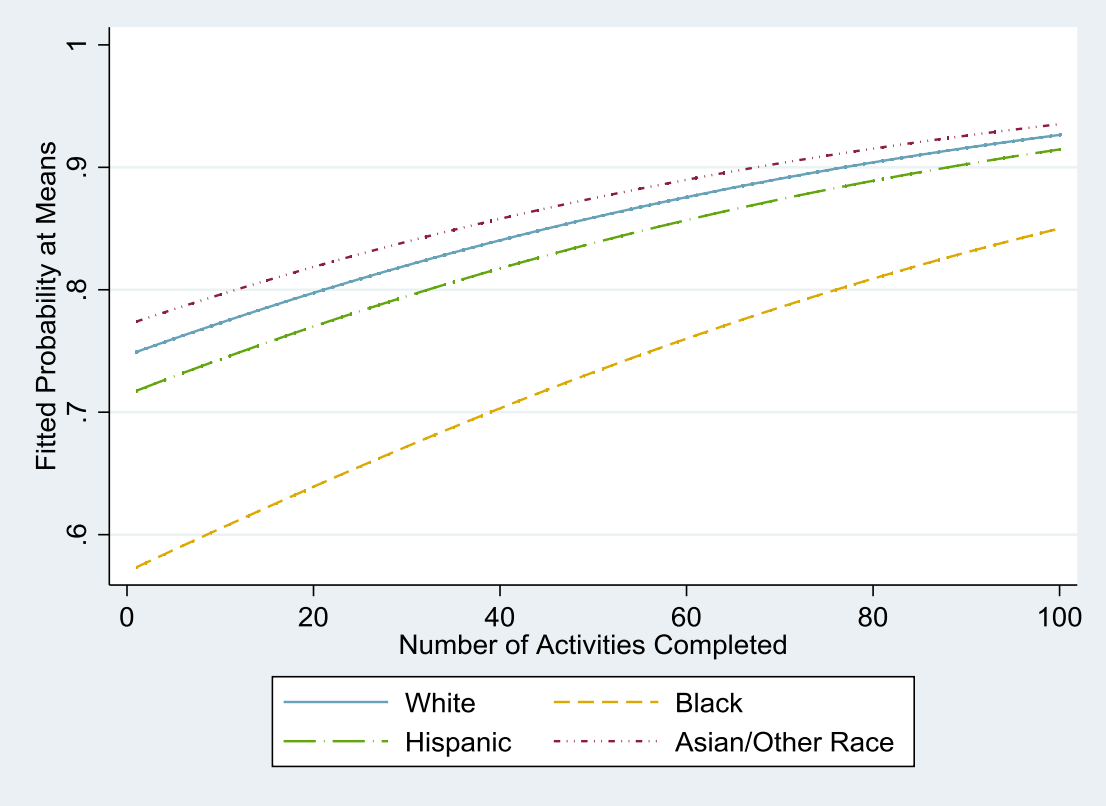


As was the case among the middle-school grade sample, the number of Achieve3000 activities completed, previous years' STAAR scale score, magnet status, gifted and talented status, and career and technical education course taking were positively associated with being on track on for college and career readiness. Blacks and males had a lower probability of being on track for college and career readiness relative to white and females, respectively. Special education students and English language learners were also less likely than their counterparts to be on track for college and career readiness. There were no statistically significant associations between being on track for college and career readiness and being Hispanic and Asian and other race, economically disadvantaged, homeless during the year, or displaced by Hurricane Harvey.

Focusing on the research question of whether greater Achieve3000 usage is associated with a higher probability of ending the year on track for being college and career ready, as determined by end-of-year Lexile performance, and whether any association differs by race/ethnicity, the results of this analysis indicate that the answer to both questions is yes. For each Achieve3000 activity completed, there was a 1 – exp(0.0145) = 0.0146, or about 1.5 percent increase in the odds of being on track for college and career readiness. While the differences in the odds of being on track for college and career readiness between Hispanic and white students and between Asian and other race and white students were statistically nonsignificant, black high school-grade students had a 1 – exp(-0.799) = 0.5503, or 55 percent lower odds of being on track for college and career readiness than did their white peers.

Figure 4 presents a graphical representation of the predicted probabilities of being on track for college and career readiness by race/ethnicity among 9th and 10th grade students. As Achieve3000 usage increases, there is an associated rise in the predicted probability, holding covariates at their means, of being on track for college and career readiness. It is notable that there also appears to be a modest convergence in the wide black-white probability gap at the low end of Achieve3000 usage as students complete more activities.

Figure 4. Predicted Probabilities of College and Career Readiness by Number of Achieve3000 Activities Completed from Panel B under Model 2 in Table D2, Grades 9 and 10.



Discussion

The aim of this study was to answer two research questions: (1) whether schools with mean completion of Achieve3000 activities completed at or above the district average had a higher mean end-of-year Lexile level than schools with mean completion of activities below the district average, and (2) whether, net of the effects of student demographic factors and between-school variance, greater usage of Achieve3000 was associated with a higher probability of ending the year on track for being college and career ready as determined by the end-of-year Lexile level, and whether any difference varies by race/ethnicity.

While the first research question was answerable in the negative, this finding should be qualified as representing qualitative differences between student bodies. It is interesting to note that the students from schools with less usage of Achieve3000 had higher STAAR scale scores at the end of the previous year than did their peers in schools with higher Achieve3000 usage. That the same students go on to achieve higher end-of-year Lexile levels despite lower Achieve3000 usage may simply suggest that these students do not need the added benefit from Achieve3000 usage as they are already higher-level achievers. On the other hand, the mean Achieve3000 usage among students in schools with mean usage at or above the district mean is so suboptimal that its impact is negligible. These potentially mitigating factors, when accounted for, might yield a different result than was revealed in basic descriptive statistics.

Both parts of the second research question were answerable in the affirmative. There was a positive relationship between Achieve3000 usage and students' college and career readiness and this relationship did indeed vary by race/ethnicity. Among students in the middle school grades (i.e., grades 6–8), each activity completed was associated with a 1 percent increase in the odds of being on track for college and career readiness. Blacks, however, experienced lower likelihood of readiness relative to whites, whose likelihood of readiness over Hispanic and Asian and other race students was statistically nonsignificant. While all groups saw increased odds of college and career readiness when more activities were completed, there did not appear to be any added advantage for blacks that the readiness gap decreased in size. Among students in the high school grades (i.e., grades 9 and 10), the same pattern was present: for each Achieve3000 activity completed, there was a 1.5 percent increase in the odds of being on track for career and college readiness. Black 9th and 10th graders had lower likelihood of readiness than did white, Hispanic, and Asian, and other race students, but there was greater convergence in the likelihood of readiness with the completion of more activities.

If it is the case that non-Asian minorities are less likely than whites and Asians to be prepared for and take rigorous courses that would benefit their reading comprehension skills, then black students' lower odds of being on track for career and college readiness is easy to understand. It is interesting, however, that Hispanic students' odds of readiness was not different from white students' odds of readiness. That this non-Asian minority group did as well as whites and Asians, while blacks did not, requires explanation. One possible explanation for the different outcomes between Hispanic and black students might be the fact that, as a majority Hispanic district, HISD does a good job of matching more of its Hispanic students to the rigorous courses traditionally unavailable to this group in more traditional white-majority districts. Blacks, on the other hand, even though they constitute a significant minority of the district's student population, still have the disadvantage of being unevenly distributed across those school catchment areas that provide the courses that would likely be more advantageous to their educational achievement and attainment. The result is that enough Hispanics, those in schools with advanced and rigorous courses, lift the achievement of the group, distinguishing them from their black counterparts.

Limitations

Though promising, the results reported here are not entirely indisputable. There are a few limitations that should be cause for reflection. First is the fact that the data analyzed here excluded a large portion of the overall population of students who had access to Achieve3000. To be included in the analytic sample, a student had to first engage Achieve3000 in either September or October to allow time for the completion of Achieve3000 activities before the end-of-year STAAR test. A student also had to be in grade 6, 7, 8, 9, or 10 and had to have a STAAR scale score in the previous year, to be used as a baseline control. The resulting sample included approximately 17,000 students, or about 17 percent of the overall secondary population. Implications for analyses are whether the distribution in Achieve3000 usage for the sample used was what it would have been expected under conditions where Achieve3000 was implemented with fidelity. It is not unreasonable to assume that the distribution for usage does not follow a bell-shaped curve and that there might be a great deal of self-selecting into usage among

students less likely to need the intervention. That the proportion of college and career ready middle and high school students is so high, given what we know from other assessments measuring college readiness, seems to support this assumption. Depending on the strength of the Achieve3000 tool to improve reading ability among those students most in need of the intervention, the relationship between usage and college and career readiness could be smaller or greater based on a sample that accurately reflects the distribution of Achieve3000 usage.

A second limitation centers on the question of whether the range of completed activities among most students was adequate to the task of truly improving reading ability so that students were college and career ready. Maximizing the impact of Achieve3000 implementation requires progressing toward its 40-activity goal per semester, or 80 activities per year. The analytic samples used here showed a mean number of activities completed far below that; the mean among the middle school-grade and high school-grade samples were both below 20 activities completed. The question, of course, is whether the relationship between the number of activities completed and being on track for college and career readiness is causal in any way, or whether the increase in readiness is a consequence of traditional curriculum delivery. Absent a randomized controlled trial or the use of quasi-experimental statistical techniques, it is difficult to know for sure. A higher mean number of activities completed, though, would give greater confidence that at least some of the relationship is causal, even if causal methods are not used.

Finally, the quality of the Achieve3000 data was poor to low. While effective implementation of the intervention requires exact testing windows in the first 30 days, mid-year, and end-of-year, there is little indication this happened among HISD students (see **Table 2**). Given that many students and schools were adversely impacted by Hurricane Harvey during the start of the 2017–2018 academic year, one might expect negative effects on the implementation of program tools and resources that are considered supplemental. That said, even if one extends the initial testing window into October, only 30,770 students out of nearly 100,000 took the Achieve3000 pre-test. Another 24,723 students took the initial test from November of 2017 through June of 2018. Fewer than 1,300 students completed a mid-year test in either December of 2017 or January of 2018, while another 245 completed a mid-year test from February of 2018 through May of 2018. Out of nearly 100,000 students, fewer than 5,000 completed an end-of-year Achieve3000 assessment.

Table 2. LevelSet Assessments Administered by Month and Testing Window.

	Testing Window		
Month-Year	Pre-Test	Interim Test	Post-Test
Sep-2017	8,565	0	0
Oct-2017	22,205	0	0
Nov-2017	7,769	0	0
Dec-2017	3,671	848	0
Jan-2018	3,727	383	0
Feb-2018	4,204	126	0
Mar-2018	2,637	93	0
Apr-2018	2,201	19	20
May-2018	512	7	4,905
Jun-2018	2	0	26
Total	**55,493**	**1,476**	**4,951**

The implication of low fidelity of implementation is that it was difficult to assess Achieve3000 using Achieve3000 data. That is, while it was possible to cull an analytic sample from the Achieve3000 data by filtering for those who provided pre-test scores in September and October of 2017, as this study did, the outcome, the college and career ready indicator, had to be constructed from the end-of-year Lexile level provided by STAAR rather than that provided by Achieve3000. Though Lexile levels from different tests can generally be compared, the same student can receive different scores depending on whether the test is summative, formative, low-stakes, or high-stakes. STAAR, which is an end-of-grade high-stakes test, could very well reveal Lexile levels that are not in line with the progress-monitoring aims of the Achieve3000 tool. It would have been preferable to have had, for the majority of students with access to the intervention, complete Achieve3000 data that included a pre-test score, interim score, and post-test score, along with the cumulative number of Achieve3000 activities completed. Complete data allow for a truer picture of progress over the course of the year, net of the effects of a student's initial

performance, as well as demographic and other factors.

Conclusion and Recommendations

With the foregoing limitations in mind, the following course of action is recommended for consideration to improve both data quality and statistical analysis for the 2018–2019 evaluation. (For a comprehensive document that sets out aims to support effective Achieve3000 implementation, see the document titled, "Maximizing the Impact of Your Achieve3000® Implementation.")

1. Schedule testing windows for the first 30 days, middle of the year, and end of the year, during which students will take Achieve3000's LevelSet assessment.
2. Ensure fidelity to the testing windows by communicating implementation expectations to relevant teachers or other school staff. Using HISD's Academic Services Memo system is one such approach that may be used to meet this recommendation.
3. Run regular reports during each testing window to guarantee all teachers have administered the LevelSet assessment within the testing window.
4. Schedule teacher workshops to communicate the importance of regularly encouraging students to strive for the 80-activity goal for the full academic year. Students can reach Achieve3000's 80-activity goal by completing 2 to 3 activities or more per week.
5. Train and regularly remind teachers and literacy specialists to better integrate Achieve3000 lessons into the core curriculum through literacy professional development sessions.
6. Educate and enroll parents in Achieve3000's Home Edition so that they, too, can encourage their children to strive for the 80-activity goal for the academic year.
7. Have central office staff and relevant teachers run regular reports with the express purpose of making sure students are progressing toward Achieve3000's 80-activitiy goal for the academic year.

To assure that students achieve the maximum literacy gains, make significant academic progress, and are college and career ready will warrant integrating the Achieve3000 differentiated reading tool into the district's comprehensive secondary reading initiative through actions that include the recommendations above. The more integrated the approach, and the more that fidelity to program implementation is adhered to, the greater the benefit to HISD students.

References

ACT, Inc. (2017). *The condition of college & career readiness, 2017*. ACT, Incorporated.

Achieve3000®. (2012). *College and career readiness: Are my students on track for college and career readiness?* Retrieved from http://doc.achieve3000.com/marketing/A3K12-0007_CCR_TechnicalPaper-MidYear.pdf

Achieve3000®. (2017). *How it works*. Retrieved from https://www.achieve3000.com/how-it-works/

Adams. M. (2009) The challenge of advanced texts: The interdependence of reading and learning. In E. H. Hiebert (Ed.), *Reading more, reading better: Are American students reading enough of the right stuff?* (pp. 163–189). New York, NY: Guilford.

Allington, R.L., & McGill-Franzen, A. (2008). Comprehension difficulties among struggling readers. In S.E. Israel, & G.G. Duffy (Eds.), *Handbook of research on comprehension* (pp. 551–568). New York, NY: Routledge.

Almond, M.. (2017). *Paper thin? Why all high school diplomas are not created equal.* Washington, DC: The Alliance for Excellent Education.

Au, K.H. (2000). A multicultural perspective on policies for improving literacy achievement: Equity and excellence. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III) (pp. 835–851). Mahway, NJ: Lawrence Erlbaum Associates.

Autor, D.H., Levy, F., & Murnane, R.J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, *118*(4), 1279–1333.

Borman, G.D., Park, S.J., & Min, S. (2015). *The district-wide effectiveness of the Achieve3000 Program: A quasi-experimental study*. Retrieved from https://files.eric.ed.gov/fulltext/ED558845.pdf

Carnevale, A.P., Smith, N., & Strohl, J. (2013). *Recovery: Job growth and education requirements through 2020*. Washington, DC: Georgetown University Center on Education and the Workforce.

College Board. (2017). *2017 SAT Suite of Assessments Annual Report*. Retrieved from https://reports.collegeboard.org/sat-suite-program-results/detailed-2017-reports

Deil-Amen, R., & Rosenbaum, J.E. (2002). The unintended consequences of stigma-free remediation. *Sociology of Education*, 249-268.

DePaoli, J.L., Balfanz, R., Bridgeland, J., Atwell, M., & Ingram, E.S. (2017). Building a grad nation: Progress and challenge in raising high school graduation rates. *America's Promise Alliance.* Retrieved from http://gradnation.americaspromise.org/report/2017-building-grad-nation-report

Hayes, D.P., Wolfer, L.T., & Wolfe, M.F. (1996). Schoolbook simplification and its relation to the decline in SAT-verbal scores. *American Educational Research Journal*, *33*(2), 489–508.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Hoboken, New Jersey: John Wiley & Sons.

ISTE Policy Brief. (2008). *Technology and student achievement – the indelible link*. International Society for Technology in Education.

Joshi, R.M., Binks, E., Graham, L., Ocker-Dean, E., Smith, D. L., & Boulware-Gooden, R. (2009). Do textbooks used in university reading education courses conform to the instructional recommendations of the National Reading Panel?. *Journal of Learning Disabilities*, *42*(5), 458-463.

Joshi, R.M., Binks, E., Hougen, M., Dahlgren, M.E., Ocker-Dean, E., & Smith, D.L. (2009). Why elementary teachers might be inadequately prepared to teach reading. *Journal of Learning Disabilities*, *42*(5), 392–402.

Levy, F., & Murnane, R.J. (2004). *The new division of labor: How computers are creating the next job market*. Princeton, NJ: Princeton University Press.

Lockard, C.B., & Wolf, M. (2012). Occupational employment projections to 2020. *Monthly Lab. Rev.*, *135*, 84.

MacGibbon, B., & Tomberlin, T.J. (1989). Small area estimates of proportions via empirical Bayes techniques. *Survey Methodology 15*, 237–252.

McCombes-Tolis, J., & Feinn, R. (2008). Comparing teachers' literacy-related knowledge to their state's standards for reading. *Reading Psychology*, *29*(3), 236–265.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects: Appendix A. Washington, DC : Authors.

Rosenbaum, J.E. (2001). *Beyond college for all*. New York, NY: Russell Sage.

Rosenbaum, J.E. (2004). It's time to tell the kids: If you don't do well in high school, you won't do well in college (or on the job). *American Educator*, *28*, 8–42.

Rosenbaum, J.E., & Gordon-McKeon, B. (2003). *College for all: How has it changed?* Unpublished paper, Institute for Policy Research, Northwestern University, Chicago, IL.

Shanahan, T., Fisher, D., & Frey, N. (2012). The challenge of challenging text. *Educational Leaderships, 69*(6), 58–62.

Shannon, L., & Grant, B. (2015). *A final report for the evaluation of the Achieve3000 programs.* Charlottesville, VA: Magnolia Consulting, LLC.

Spear-Swerling, L. (2011). Patterns of reading disabilities across development. In A. McGill-Franzen, & R.L. Allington (Eds.), *Handbook of reading disability research* (pp. 149–161). New York, NY: Routledge.

Stenner, A.J., Koons, H., & Swartz, C.W. (2010). *Text complexity and developing expertise in reading*. Durham, NC: MetaMetrics, Inc..

Torres, D. D. (2015). *Achieve3000® impacts on student reading and STAAR EOC English I, English II, and Biology exams for the 2014–2015 academic year*. Houston, TX: Houston Independent School District.

Torres, D. D. (2016). *Achieve3000® impacts on students' STAAR EOC English I and English II performance for the 2015–2016 academic year.* Houston, TX: Houston Independent School District.

US Department of Education Office for Civil Rights. (2014, March). *Civil rights data collection data snapshot: College and career readiness* (Issue Brief No. 3). Washington, DC: Author.

Washburn, E.K., Joshi, R.M., & Cantrell, E.B. (2011). Are preservice teachers prepared to teach struggling readers?. *Annals of Dyslexia*, *61*(1), 21–43.

Wei, H., Cromwell, A.M., & McClarty, K.L. (2016). Career readiness: An analysis of text complexity for occupational reading materials. *The Journal of Educational Research*, *109*(3), 266–274.

Williamson, G.L. (2006). *Aligning the Journey with a Destination: A Model for K–16 Reading Standards*. Durham, NC: MetaMetrics, Inc.

Williamson, G.L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, *19*(4), 602–632.

Wirt, J., Choy, S., Rooney, P., Provasnik, S., Sen, A., & Tobin, R. (2004). The condition of education 2004 (NCES 2004-077). Washington, DC: US Department of Education. *National Center for*

Education Statistics.

For additional information contact the HISD Department of Research and Accountability at 713-556-6700 or e-mail Research@Houstonisd.org.

Appendix A

Table A1. College and Career Readiness Performance Standards in the Lexile Metric.

	College and Career Readiness			
	Not On Track		On Track	
Grade	Falls Far Below	Approaches	Meets	Exceeds
1	-111 and Below	-110 to 185L	190L to 530L	535L and Above
2	150L and Below	155L to 415L	420L to 650L	655L and Above
3	265L and Below	270L to 515L	520L to 820L	825L and Above
4	385L and Below	390L to 735L	740L to 940L	945L and Above
5	500L and Below	505L to 825L	830L to 1010L	1015L and Above
6	555L and Below	560L to 920L	925L to 1070L	1075L and Above
7	625L and Below	630L to 965L	970L to 1120L	1125L and Above
8	660L and Below	665L to 1005L	1010L to 1185L	1190L and Above
9	775L and Below	780L to 1045L	1050L to 1260L	1265L and Above
10	830L and Below	835L to 1075L	1080L to 1335L	1340L and Above
11/12	950L and Below	955L to 1180L	1185L to 1385L	1390L and Above

Source: Adapted from table printed in a technical paper published by Achieve3000 (2012).

Appendix B

Random Intercept Logistic Regression Model

This study assumes, following the work of Hedeker and Gibbons (2006), that there are N level-2 units with n_i level-1 units in the ith cluster producing a total sample size of $n = \sum_{i=1}^{N} n_i$. It supposes that the response, y_{ij}, is binary, and follows the generalized linear mixed model (GLMM) originally proposed by MacGibbon and Tomberlin (1989), $\Pr(y_{ij} = 1|p_{ij}) = p_{ij}$. Modeling the response pattern for the binary dependent variable using the logit model yields the following specification:

$$\text{logit}\left(p_{ij}\right) = \mathbf{x}'_{ij}\beta + u_i; \quad u_i \sim N(0, \sigma_u^2). \tag{1}$$

Here p_{ij} denotes the response probability $Pr(y_{ij} = 1)$ for the jth student from the ith school. Per the conventional model, **x** is included to account for observed sources of variation in the response. The level-two residual u_i is assumed to be normally distributed with mean 0 and variance σ_u^2, and independent of **x**. Conditional on **x**, the random effect u_i increases an expected response for a student in school i when it is positive and decreases an expected response for a student in school i when it is negative.

Predicting the Overall Proportion of College and Career Ready Students

Predicting the overall proportion of students who are college and career ready can be conducted in the usual way from the null, or unconditional, models under Panel A in Model 1 (i.e., the models without covariates) as

$$\hat{p} = \Lambda\left(\hat{\beta}_0\right) = \frac{\exp(\hat{\beta}_0)}{1 + \exp(\hat{\beta}_0)} \tag{2}$$

Estimating the Marginal Probability from the Random Effects

An estimate of the marginal expectation for college and career readiness when random effects are added under Panel B in Model 1 can be quantified as

$$\hat{p} = \int_u \frac{\exp(\hat{\beta}_0 + u_i)}{1 + \exp(\hat{\beta}_0 + u_i)} g(u)\mathrm{d}u\,, \tag{3}$$

where $g(u)$ is the normal probability function with mean 0 and variance σ_u^2, and u_i is the subject-specific random effect.

Appendix C

Table C1. Student's *t* and Hedge's *g* (effect size) Comparing Lexile Level between Schools Below the District Mean Number of Activities Completed and Schools At or Above the District Mean Number of Activities Completed.

	School Below District Mean Activities Completed			School At or Above District Mean Activities Completed						
Grade	Observations	Mean	SD	Observations	Mean	SD	Difference	*t*-value	*p*-value	Effect Size (Hedge's *g*)
Grade 6	1,565	**893.5**	253.0	624	**793.1**	220.3	**-100.4**	8.7	<.001	**.41**
Grade 7	1,828	**1048.3**	278.1	629	**898.9**	229.5	**-149.3**	12.1	<.001	**.56**
Grade 8	1,677	**1029.0**	255.1	806	**976.2**	232.2	**-52.8**	5.0	<.001	**.21**
Grade 9	2,377	**1173.8**	266.1	2,897	**1099.7**	272.0	**-74.1**	9.9	<.001	**.27**
Grade 10	2,094	**1280.0**	261.8	2,513	**1152.9**	255.6	**-127.1**	16.6	<.001	**.49**

Appendix D

Table D1. Logit Models for College and Career Readiness, Middle School Grades (6, 7, and 8).

	Model 1		Model 2	
	A	B	A	B
	Estimate (std. error)	Estimate (std. error)	Estimate (std. error)	Estimate (std. error)
Intercept	-0.0916***	0.348***	0.952***	0.654**
	(0.0237)	(0.0321)	(0.216)	(0.227)
# of Activities Completed			0.00151	0.00682**
			(0.00232)	(0.00241)
Previous Years' STAAR Scale Score (centered)			0.150***	0.150***
			(0.00452)	(0.00461)
Race/Ethnicity (Ref. = White)				
Black			-0.819***	-0.589**
			(0.218)	(0.225)
Hispanic			-0.512*	-0.366
			(0.210)	(0.216)
Asian/Other Race			-0.206	-0.140
			(0.304)	(0.311)
Sex (Ref. = Female)				
Male			-0.354***	-0.337***
			(0.0690)	(0.0703)
Economically Disadvantaged			-0.321**	-0.185
			(0.122)	(0.128)
Special Education			-0.964***	-1.037***
			(0.179)	(0.181)
Magnet			0.348***	0.432***
			(0.0793)	(0.0864)
Gifted/Talented			0.667***	0.590***
			(0.106)	(0.109)
English Language Learner (ELL)			-0.838***	-0.866***
			(0.0900)	(0.0914)
Career and Technical Education			0.407***	0.454***
			(0.115)	(0.121)
Homeless During Year			0.438	0.623
			(0.734)	(0.732)
Displaced by Hurricane Harvey			-0.185	-0.155
			(0.114)	(0.117)
σ_u^2		0.783***		0.686***
		(0.0270)		(0.0745)
ρ		0.192		0.173
-2 log *L*	9865.14	8768.94	5327.94	5251.14
df	7128	7127	7113	7112

Note: *p < .05, **p < .01, ***p < .001; two-tailed tests.

Table D2. Logit Models for College and Career Readiness, High School Grades (9 and 10).

	Model 1		Model 2	
	A	B	A	B
	Estimate (std. error)	Estimate (std. error)	Estimate (std. error)	Estimate (std. error)
Intercept	0.732***	0.837***	2.075***	1.472***
	(0.0215)	(0.0263)	(0.204)	(0.214)
# of Activities Completed			0.0100***	0.0145***
			(0.00182)	(0.00195)
Previous Years' STAAR Scale Score (centered)			0.00780***	0.00689***
			(0.000599)	(0.000609)
Race/Ethnicity (Ref. = White)				
Black			-1.151***	-0.799***
			(0.187)	(0.194)
Hispanic			-0.508**	-0.162
			(0.182)	(0.189)
Asian/Other Race			-0.000900	0.137
			(0.270)	(0.280)
Sex (Ref. = Female)				
Male			-0.471***	-0.453***
			(0.0567)	(0.0580)
Economically Disadvantaged			-0.240**	-0.121
			(0.0823)	(0.0856)
Special Education			-2.334***	-2.349***
			(0.102)	(0.105)
Magnet			1.276***	1.399***
			(0.0852)	(0.111)
Gifted/Talented			1.893***	1.889***
			(0.175)	(0.179)
English Language Learner (ELL)			-2.649***	-2.718***
			(0.0764)	(0.0790)
Career and Technical Education			0.0584	0.00334
			(0.0848)	(0.0926)
Homeless During Year			0.0296	-0.163
			(0.194)	(0.197)
Displaced by Hurricane Harvey			-0.191	-0.149
			(0.116)	(0.120)
σ_u^2		0.747***		0.609***
		(0.0261)		(0.0396)
ρ		0.185		0.156
-2 log *L*	12462.07	11271.26	8044.15	7803.38
df	9880	9879	9866	9865

Note: *p < .05, **p < .01, ***p < .001; two-tailed tests.