MEMORANDUM   June 14, 2018

TO: Lance Menster
           Officer, Elementary Curriculum and Development

FROM: Carla Stevens
           Assistant Superintendent, Research and Accountability

SUBJECT: **LONGITUDINAL COHORT STUDY ON THE SHORT-TERM, INTERMEDIATE,
           AND LONG-TERM EFFECTS OF HISD PREKINDERGARTEN ON ACADEMIC
           PERFORMANCE, BEHAVIOR, COLLEGE PREPAREDNESS, AND SCHOOL
           PERSISTENCE, 2017–2018**

Attached is a copy of the prekindergarten longitudinal study for the 2017–2018 academic year.
The study tracked the performance of three cohorts of students who were eligible or attended
HISD prekindergarten during the 2001–2002 (Cohort I), 2002–2003 (Cohort II), and 2003–2004
(Cohort III) academic years.

Key findings include:
- Kindergarten students in Cohorts I, II, and III outperformed students who did not attend
  HISD prekindergarten on nationally-normed Stanford and Aprenda reading and mathematics
  tests, indicating a statistically significant positive effect of HISD prekindergarten in the short-
  term.
- HISD prekindergarten had extended effects for Cohort II students, given that they
  consistently outperformed non-HISD prekindergarten students at targeted grades third, fifth,
  and seventh.
- Long-term effects related to college preparedness on the PSAT, SAT, and ACT reading and
  mathematics tests were mostly in favor of non-HISD prekindergarten students; however,
  these results were not consistently significant.
- A higher proportion of HISD prekindergarten students demonstrated school persistence,
  having graduated on-time at 12 years, with significantly lower dropout rates compared to
  non-HISD prekindergarten students.
- Attendance rates at each level of measurement were significantly in favor of HISD
  prekindergarten 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
    Gabrielle Coleman

HOUSTON INDEPENDENT SCHOOL DISTRICT

RESEARCH
Educational Program Report

LONGITUDINAL COHORT STUDY ON THE SHORT-TERM, INTERMEDIATE, AND LONG-TERM EFFECTS OF HISD PREKINDERGARTEN ON ACADEMIC PERFORMANCE, BEHAVIOR, COLLEGE PREPAREDNESS, AND SCHOOL PERSISTENCE, 2017-2018

HISD | Research and Accountability
ANALYZING DATA, MEASURING PERFORMANCE.

2018 BOARD OF EDUCATION

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

www.HoustonISD.org

It is the policy of the Houston Independent School
District not to discriminate on the basis of age, color,
handicap or disability, ancestry, national origin,
marital status, race, religion, sex, veteran status,
political affiliation, sexual orientation, gender identity
and/or gender expression in its educational or
employment programs and activities.

Longitudinal Cohort Study on the Short-term, Intermediate, and Long-term Effects of HISD Prekindergarten on Academic Performance, Behavior, College Preparedness, and School Persistence, 2017–2018

Prepared by Venita R. Holmes, Dr.P.H.

Abstract
This study hypothesized that HISD prekindergarten participation consistently correlated with better educational outcomes for students. Taking into account students' age differences through kindergarten enrollment, propensity score matching controlled for students' background characteristics. Three cohorts of students in this study were eligible or attended HISD prekindergarten during the 2001–2002 (Cohort I), 2002–2003 (Cohort II), and 2003–2004 (Cohort III) academic years. The study found that kindergarten students in Cohorts I, II, and III outperformed students who did not attend HISD prekindergarten on nationally-normed Stanford and Aprenda reading and mathematics tests, indicating a statistically significant positive effect of HISD prekindergarten in the short-term. HISD prekindergarten had extended effects for Cohort II students, given that they consistently outperformed non-HISD prekindergarten students at targeted grades third, fifth, and seventh. Long-term effects related to college preparedness on the PSAT, SAT, and ACT reading and mathematics tests were mostly in favor of non-HISD prekindergarten students; however, these results were not consistently significant. A higher proportion of HISD prekindergarten students demonstrated school persistence, having graduated on-time at 12 years, with significantly lower dropout rates compared to non-HISD prekindergarten students. Attendance rates at each level of measurement were significantly in favor of HISD prekindergarten students. The findings demonstrated that educating at-risk students during their early years has benefits. Evidence that continued support is needed for these students was also shown as they progress through school. Future research is needed to determine whether full-day HISD prekindergarten provides more lasting educational effects for students and to track their progress to ensure that the instructional model continues to build students' knowledge and skills.

Introduction

Research has shown that participation in early education programs may have a significant impact on academic progress, employment, and behavioral outcomes of economically-disadvantaged children later in life (Almond & Currie, 2011; Barnett, 1995; Heckman, 2008; Knudsen et al., 2006; Schweinhart, Montie, Xiang, et al., 2005). A review of relevant studies found that "early childhood programs can produce large short-term benefits for children on intelligence quotient (IQ) and sizable long-term effects on school achievement, grade retention, placement in special education, and social adjustment" (Barnett, 1995, p. 25).

Ansari and Winsler (2013) noted that children who attend prekindergarten programs in public schools are more likely to be exposed to higher quality teachers and a more academically-challenging curriculum (**Figure 1**) compared to children who attend center-based programs, family care, or who do not attend prekindergarten programs.

This program evaluation seems timely considering the growing interest in building public awareness and

Figure 1: HISD prekindergarten students engaged in problem-solving

Figure 2a: Prekindergarten skill development area - Technology

support for prekindergarten services (Burchinal, Mashburn, Pianta, & Vandergrift, 2009; Texas Education Agency, 2017).

To that end, this evaluation measured the short, intermediate, and long-term effects of HISD prekindergarten programs on students' academic performance, behavior, college preparedness, and school persistence using attendance, discipline, and nationally-normed reading and mathematics assessment data. Three cohorts of students were identified based on the year that they participated or were eligible to participate in HISD prekindergarten. The cohort years for the study were 2001–2002, 2002–2003, and 2003–2004. A multi-cohort year approach was conducted to determine whether there was consistency in results over time. The years for prekindergarten eligibility were selected to ensure that all outcomes of interest, particularly graduation, could be measured in the program evaluation.

Background

In compliance with Texas Education Code §29.153, school districts in Texas offer free prekindergarten for eligible children (Texas Education Agency, 2008). The Houston Independent School District (HISD) initiated its prekindergarten program during the 1985–1986 academic year. Consistent with common practices, HISD initially operated a state-funded, half-day prekindergarten program for three hours daily on elementary campuses. For HISD's two early childhood centers (Martin Luther King and Ninfa

Laurenzo), full-day prekindergarten programs were implemented in 2004. HISD expanded the prekindergarten program to full-day, during the 2006–2007 academic year, which consisted of seven hours of programming. The half-day program was implemented with state funds, while the full-day program was implemented with local and federal funds (Mellon, 2006). HISD provided half-day prekindergarten to students in each cohort for this study.

HISD prekindergarten was offered to tuition and non-tuition students (Houston Independent School District, n.d.). To be eligible for participation in the non-tuition prekindergarten program, students were required to be (a) four years old on or before September 1 of the school year, (b) live in the HISD attendance boundary, (c) meet immunization requirements, and (d) meet at least one of the following conditions: (1) unable to speak and comprehend the English language, (2) homeless (3) economically disadvantaged, (4) a child of an active duty member of the armed forces, (5) in or ever been in the conservatorship of the Department of Family and Protective Services, and (6) eligible for Head Start. Prekindergarten participants could be the child of a person eligible for the Star of Texas Award as a peace officer as defined in Texas Education Code, Section 3106.002, a firefighter as defined in Section 3106.003, or an emergency medical first responder as defined in Section 3106.004 (Texas Education Agency, 2008). The prekindergarten eligibility criteria related to the Star of Texas Award was added in 2017 by the Texas Legislature. Thus, students with Star of Texas Award designations were not included in this study.

TEA Curriculum Model

The Texas Education Agency (2011) developed guidelines to help teachers define and implement a comprehensive curriculum, focused on specific skill areas and outcomes. The TEA Prekindergarten Curriculum Guidelines implemented when the prekindergarten cohorts used in this study was conducted were established in 1999. The guidelines were obtained through a correspondence with TEA staff (TEA correspondence, 2017). **Figures 2a** through **2e** depict HISD prekindergarten students engaging in cognitive, social, and intellectual development activities that are aligned to the guidelines. The guidelines emphasized the following areas:

• Language and Literacy (listening comprehension, speech production and speech discrimination, vocabulary, verbal

Figure 2b: Prekindergarten skill development area - Fine Arts

Figure 2c: Prekindergarten skill development area - Health and Safety

Figure 2d: Prekindergarten skill development area - Science

Figure 2e: Prekindergarten skill development area - Physical Development

expression, phonological awareness, print and book awareness, letter knowledge and early word recognition, motivation to read, developing knowledge of literacy forms, written expression);

- Personal and Social Development (a sense of self, interpersonal and social skills for communicating with others);
- Health and Safety (hygiene and nutrition education, everyday routines and procedures to remain safe and avoid injury);
- Mathematics (numbers and operations, patterns, geometry and spatial sense, measurement, classification and data collection)
- Science (science processes, science concepts);
- Social Studies (individual, culture, and community; history; economics);
- Fine Arts (arts, music, dramatic play);
- Physical Development (physical movement, gross-motor development, fine-motor development), and
- Technology Applications (basic functions of the computer and related technologies).

During the study period, HISD relied on the Texas prekindergarten curriculum model to build its prekindergarten program and classroom instructional strategies. These strategies can be summarized as follows: (1) Children begin foundation instruction for literacy, numeracy, and integrated learning; (2) Children are actively involved in experiences, making meaningful choices, making decisions, and solving problems; (3) Classrooms contain materials and activities for a wide range of interests and abilities; (4) Teacher-directed and child-initiated activities are balanced throughout the day; (5) Children develop reading and communication skills by using language to express insights and solve problems through interaction with adults and peers; and (6) Children are provided time for working individually or in small groups for intentional and purposeful learning (Houston Independent School District, 2017). Given the instructional model and strategies implemented by HISD, this evaluation addressed the following research questions.

Research Questions:
1. What were the short-term and intermediate effects of HISD prekindergarten on students' academic performance, attendance, and behavior?

2. What were the long-term effects of HISD prekindergarten on college preparedness?
3. To what extent did HISD prekindergarten affect school persistence relative to graduation and school dropout?
4. What were the best predictors of school persistence for Cohort I, II, and III student groups, considering their demographic characteristics?

For this study, students were eligible or attended HISD prekindergarten during the 2001–2002 (Cohort I), 2002–2003 (Cohort II), and 2003–2004 (Cohort III) academic years.

Review of the Literature

Research regarding the impact of prekindergarten programs on student outcomes is mixed (Andrews, Jargowsky, & Kuhne, 2012; Lipsey, Farran, & Hofer, 2015). Specifically, Barnett (1995) found positive short and long-term effects of prekindergarten on school achievement, grade retention, placement in special education, and social adjustment. Research from a longitudinal study of students who attended the Abbott Preschool Program demonstrated that children continued to outperform their peers in reading and mathematics through early elementary school and were less likely than their peers to be retained or require remediation (Frede, Jung, Barnett, & Figueras, 2009). A follow-up of the Perry Preschool study, with randomly-assigned three and four-year old children, found that at age 40, children who participated in the High Scope participatory learning program had higher earnings, were more likely to hold a job, had committed fewer crimes, and were more likely to have graduated from high school than adults who did not participate in the program (Schweinhart et al., 2005).

A meta-analyses of 22 experimental and quasi-experimental studies conducted between 1960 and 2016 found that, on average, participation in early childhood education led to "statistically significant reductions in special education placement ($d = 0.33$ SD, 8.1 percentage points) and grade retention ($d = 0.26$ SD, 8.3 percentage points) and increases in high school graduation rates ($d = 0.24$ SD, 11.4 percentage points)" (McCoy et al., 2017). A cohort study conducted by the Texas Education Agency (2017), from 1999 through 2002, to examine the state's long-term investment in public half-day prekindergarten observed small, but statistically significant differences between eligible student attendees and eligible students who did not attend. Prekindergarten attendance

	N	HISD Prek	Non-HISD Prek		n	HISD Prek	Non-HISD Prek
			Before Matching				After Matching
Cohort I (2001–2002)	14,219	8,123	6,096		12,046	8,123	3,923
Cohort II (2002-2003)	13,837	8,092	5,745		11,746	8,092	3,654
Cohort III (2003-2004)	13,420	8,307	5,113		11,652	8,307	3,345
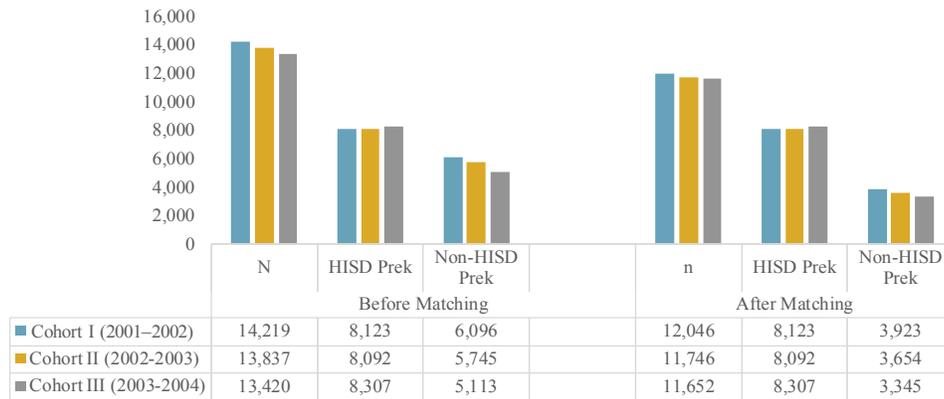
Figure 3: Cohorts I, II, and III counts before and after propensity matching of HISD prekindergarten students ("treated") and non-HISD prekindergarten students ("control") - Note: Cohort year represents the academic year that students attended HISD prekindergarten or were eligible to attend HISD prekindergarten.

was associated with a negligible, but statistically significant difference in performance for students on the state-mandated reading test. The study also noted that "prekindergarten attendance was associated with a lower likelihood of dropping out of school (7% vs. 9%), a higher likelihood of graduating high school on time (72% vs. 66%), and enrolling and persisting in college (38% vs. 31% and 28% vs. 22%, respectively)" (Texas Education Agency, 2017, p.2).

In contrast, an examination of data from the national Early Childhood Longitudinal Study (2010-11) revealed academic readiness gaps continued to exist between low-income and high-income children who did and did not participate in prekindergarten programs, on average, in mathematics and reading (U.S. Department of Education, 2011; Nores & Barnett, 2014). Lipsey, Farran and Hofer's (2015) randomized-controlled study at 58 Tennessee schools found statistically significant positive prekindergarten effects on students' Woodcock Johnson achievement measures and on teacher ratings obtained at kindergarten. A follow-up study found that all positive achievement differences favoring state prekindergarten participants at the end of the prekindergarten year were no longer statistically significant by the end of kindergarten. Moreover, by the end of third grade, non-prekindergarten control group students outperformed, on average, the prekindergarten participants on some achievement measures.

Methods
Study Population
This longitudinal observational study compared three cohorts of students who participated in HISD prekindergarten (treatment) with students who did not participate in HISD prekindergarten (control), adjusting for pretreatment effects (race/ethnicity, gender, economic status, at risk, special education, gifted/talented, and limited English proficiency (LEP). The identification of the study population occurred in multiple phases to control for age differences in the groups and took into account that kindergarten attendance is not mandatory in the state of Texas (Texas Education Agency, 2008, 2009; TEC §25.085). Moreover, the use of multiple cohorts allowed for data triangulation and validation of study findings to determine whether outcomes were consistently observed over time. The selected cohort years for HISD prekindergarten or prekindergarten eligibility (Cohort I: 2001–2002, Cohort II: 2002–2003, and Cohort III: 2003–2004) allowed sufficient time for students to experience outcomes of interest, including academic performance and graduation during elementary, middle, and high school. The Public Education Information Management System

(PEIMS) was used to identify and gather data on background characteristics of the study population, which was used in the propensity score matching procedures prior to data collection on student outcomes. Differentiation was not made regarding whether HISD prekindergarten students were tuition or non-tuition students. Half-day prekindergarten was being implemented in HISD during the years that the cohorts were developed.

To isolate HISD prekindergarten participation, it was necessary to determine which students were enrolled in HISD kindergarten programs during the designated cohort years. To control for kindergarten participation and age differences, the study population was comprised of students who successively attended both kindergarten and first grade, to offset the non-mandatory prekindergarten attendance requirement in Texas. **Figure 3** depicts the total number of students in each cohort before and after statistical adjustments were made to generate treatment and comparison groups. More details are, subsequently, provided in this evaluation regarding sample selection.

Study Sample Selection
Propensity score matching (PSM) using "nearest-neighbor matching" techniques was conducted to identify treatment (HISD prekindergarten) and comparison groups (non-HISD prekindergarten), controlling for race/ethnicity, gender, economic status, at risk, special education, gifted/talented, and limited English proficiency (LEP). PSM is considered a viable method to estimate causal treatment effects, selection bias, and control for observed bias (Rosenbaum & Rubin, 1985; Rubin, 1997; Joffe & Rosenbaum, 1999; Murnane & Willett, 2011). SPSS software applied a PSM procedure (Ho, Imai, King, & Stuart, 2007; Ho, Imai, King, & Stuart, 2011; Thoemmes & Kim, 2011; Thoemmes, 2012) to create the groups. When more than one good match existed for the treatment group, PSM randomly selected one of the duplicates with replacement to become the actual neighbor (Murnane & Willett, 2011). Thus, PSM identified some students as matches for multiple treatment group students. Control variables were used as fixed effects and random effects to adjust for the variation within and between groups. The demographic, pretreatment variables were selected because they were considered to be unrelated to assignment to treatment (Rosenbaum & Rubin, 1985).

The quality of matches can be affected by the order in which subjects are selected for matching and the maximum permitted difference between matched subjects (the "caliper") (Lunt, 2013).

Table 1: Standardized Mean Differences on Covariates based on Overall Balance Test After Propensity Score Matching

Baseline Measures	Cohort I (2001–2002*)			Cohort II (2002–2003*)			Cohort III (2003–2004*)		
	Means Treated	Means Control	Std. Mean Diff.	Means Treated	Means Control	Std. Mean Diff.	Means Treated	Means Control	Std. Mean Diff.
Race/Ethnicity									
Asian	.021	.014	.044	.020	.020	.001	.021	.020	.007
Black	.259	.258	.004	.238	.245	-.018	.234	.250	-.038
Hispanic	.707	.716	-.019	.725	.724	.003	.732	.722	.023
White	.013	.012	.004	.017	.011	.046	.013	.008	.042
Gender (F=0, M=1)	.503	.512	-.020	.508	.506	.003	.501	.517	-.032
Eco Disadv.	1.000	1.000	n/a	.951	.953	-.007	.942	.947	-.020
At Risk	.796	.798	-.005	.802	.815	-.033	.843	.842	.001
Special Ed	.030	.033	-.016	.032	.043	-.060	.041	.040	.003
G/T	.068	.054	.054	.076	.047	.112	.093	.071	.074
LEP	.582	.576	.011	.578	.568	.019	.568	.544	.049

Note: Hansen and Bowers (2008) Overall Balance Test conducted in SPSS using MatchIt, R1tools and Coarsened Exact Matching (CEM) to balance covariates
*Cohort year represents the academic year that students attended HISD prekindergarten or were eligible to attend HISD prekindergarten.; Std. Mean Diff < .25 considered balanced
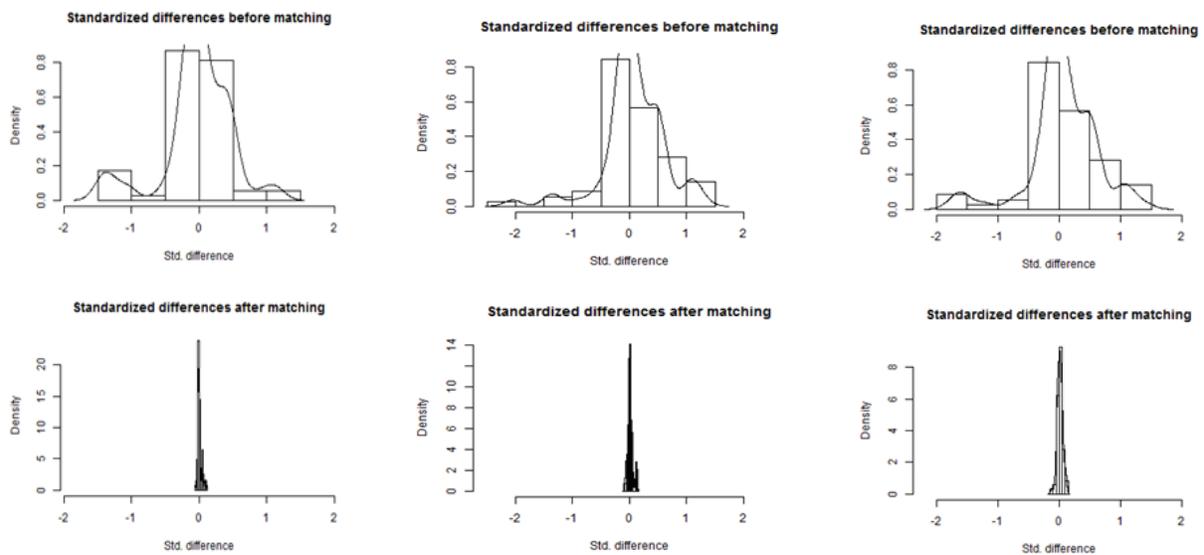
Figure 4: Histograms for Cohorts I, II, and III, respectively, with overlaid kernel density estimates of standardized differences before and after propensity matching of HISD prekindergarten students ("treated") and non-HISD prekindergarten students ("control") using MatchIt package (Thoemmes & Kim, 2011; Thoemmes, 2012.)
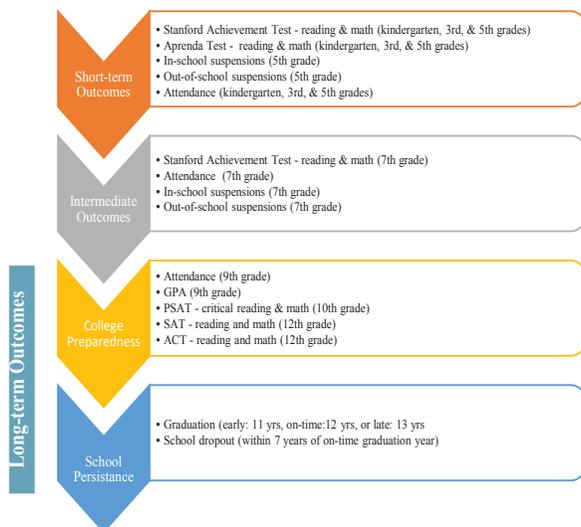
Figure 5: Short-term, intermediate, and long-term outcomes study measures

The caliper for this evaluation was .2 standard deviations, which is the default statistic for SPSS PS Matching. A caliper of .2 standard deviations is more rigid than the .25 standard deviations applied in Rosenbaum and Rubin's (1985) and Cochran and Rubin's (1973) logistic regression model to predict exposure to treatment.

After matching, the balance of all observed covariates, interactions among all covariates, and quadratic terms of all covariates were examined. No imbalances remained as assessed through univariate and multivariate tests. **Appendix A** (pp. 12–14) presents pre-intervention characteristics for Cohorts I, II, and III baseline and analytic samples. **Table 1** shows cohort balance after propensity score matching. Histograms depicted in **Figure 4** reveals the covariate balance was greatly improved in the matched sample compared to the non-matched sample for Cohorts I, II, and III, respectively.

Data Collection and Analyses

Short-term, intermediate, and long-term outcome measures used in this study are presented in **Figure 5**. Short-term and intermediate academic measures included students' performance on

the nationally-normed Stanford (English language) and Aprenda (Spanish language) tests. Long-term academic outcome measures related to college preparedness consisted of ninth grade GPA along with tenth grade PSAT, twelfth grade SAT, and twelfth grade ACT reading and mathematics scores. Long-term school persistence was comprised of graduation (early in 11 years, on-time in 12 years, and late in 13 years) from school entry in first grade, and school dropout. School dropout was based on whether students dropped out of school within seven years of expected graduation in 12 years. Additional short-term, intermediate, and long-term outcome measures to determine whether behavioral factors were affected by HISD prekindergarten participation were kindergarten, third, fifth, seventh, and ninth grade attendance along with in-school and out-of-school suspensions at fifth and seventh grades.

Independent t-tests were conducted using Stanford and Aprenda normal curve equivalents to determine whether there were statistically significant differences in the means between the groups. Similar analyses were conducted between the groups relative to attendance rates, GPA, PSAT, SAT, and ACT. Benchmark data for PSAT, SAT, and ACT were extracted from College Board and ACT, Inc. reports to facilitate interpretation of results (ACT, 2013; Wyatt, Smith, & Proestler, 2014). The level of statistical significance was p<.05. Effect sizes were also computed to measure the magnitude of the program's impact. Hedge's g is a standard deviation-based measure used to compute the effect size for groups with different sample sizes. Hedge's g follows similar criteria to Cohen's d for determining the strength of an intervention with an effect size of 0.2 = small effect, 0.5 = moderate effect, and 0.8 = large effect. The What Work's Clearinghouse notes that an effect size of 0.25 standard deviations or larger is considered to be substantively important (What Works Clearinghouse, n.d.; **Appendix B**, p. 15).

Odds ratios were calculated for discrete outcomes, including in and out-of-school suspensions (1 = yes and 0 = no), graduation (3 = early, 2 = on-time, and 1= late), and dropout at specific points in time. The level of statistical significance was p<.05. Confidence Intervals (CI) were computed. If a 95% CI included the null value of 1, then there was insufficient evidence to conclude that the groups were statistically significantly different.

Multinomial logistic regression analyses were conducted to determine the best predictors of on-time (at twelve years), early (at 11 years), and late (at 13 years) graduation for Cohort I, II, and III student groups. The model controlled for gender, economic status, gifted/talented, special education identification, at risk status, limited English proficient (LEP), and whether or not students participated in HISD prekindergarten.

Study Limitations

There were several limitations of the study which may affect confidence in the analyses. First, it was unknown whether students who did not attend HISD prekindergarten programs attended other early education programs prior to enrollment in HISD at kindergarten. Measurement of outcomes was limited to the availability of the data in HISD data systems; therefore, students who lacked data on variables of interest were excluded from the study. Graduation status was limited to a review of student records one year before and one year after expected graduation. Therefore, it was unknown whether students whose data were not available graduated in other school districts or even dropped out of school prior to the seven-year dropout criteria observed in this study. Other factors that may have greatly influenced students' educational outcomes, in-

cluding the quality of instruction, parental involvement (Camilli et al., 2010), and family practices (Fuller et al., 2017) were not measured in this study. However, propensity score matching was conducted to establish baseline equivalence between the groups by controlling for background characteristics that may have influenced prekindergarten participation. In addition, multiple cohorts were used to understand the effect of prekindergarten and important contributors to student success based on prekindergarten research (Texas Education Agency, 2017).

Results

What were the short-term and intermediate effects of HISD prekindergarten on students' academic performance, attendance, and behavior?

The nationally-normed Stanford and Aprenda reading and mathematics tests were used to determine the short-term and intermediate effects of HISD prekindergarten on students' academic achievement. Students' attendance and disciplinary actions were also used as measures of these effects at these points in time in the study. Short-term measures were observed at kindergarten, third grade, and fifth grade, while intermediate effects were observed at seventh grade. The results for Cohorts I, II, and III can be found in **Tables 2a**, **2b**, and **2c**, respectively in **Appendix C**.

The most notable results in favor of HISD prekindergarten were observed at the kindergarten level relative to Stanford and Aprenda reading and mathematics performance. Specifically, there were statistically significant positive effects of HISD prekindergarten on students' kindergarten Stanford and Aprenda reading and mathematics performance in all cohorts (**Figures 6a**, **6b**, and **6c**). The largest effect was found on the Aprenda reading and mathematics tests for Cohort I (d=.326, d=.323); Cohort II (d=.355, d=.354); and Cohort III (d=.368, d=.291) in the respective student groups (Appendix C).

By third grade, slight differences in Stanford reading and mathematics performance were observed between the groups in favor of non-HISD prekindergarten students in Cohort I; however, these differences were neither statistically significant nor substantially important (Appendix C, p. 16). At third grade, Cohort II students outperformed non-HISD prekindergarten students on Stanford reading and mathematics tests, and the results
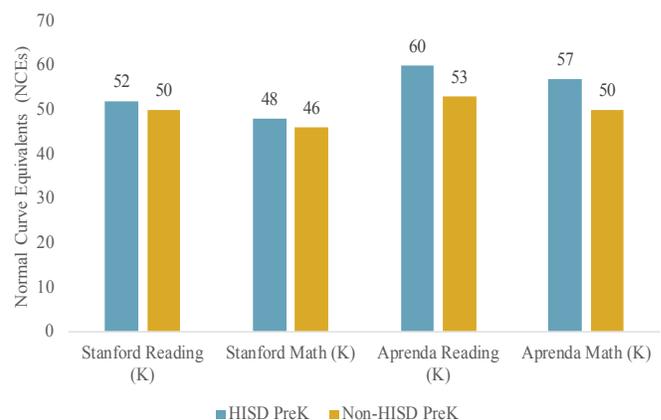
Figure 6a: Cohort I - Short-term kindergarten Stanford and Aprenda reading and math results (NCEs rounded to nearest whole number)
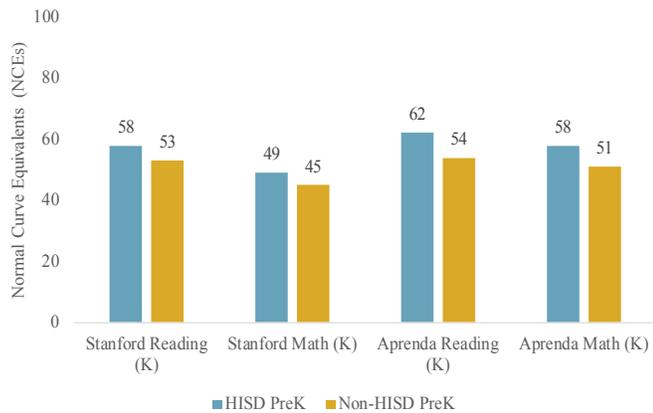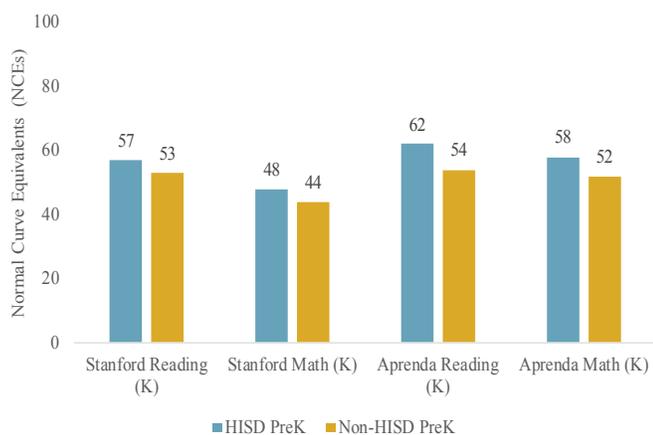
Figure 6b: Cohort II - Short-term kindergarten Stanford and Aprenda reading and math results (NCEs rounded to nearest whole number)

were statistically significant in mathematics (Appendix C, p. 17; p=.015). Group differences in Cohort III at third grade (Appendix C, p. 18) were not statistically significant, although HISD prekindergarten students achieved a slightly higher mean on the Stanford mathematics test and non-HISD prekindergarten students achieved a slightly higher mean than HISD prekindergarten students on the Stanford reading test.

At fifth grade, on the Stanford test, HISD prekindergarten students in Cohort I had a slightly higher mean in mathematics, while non-HISD prekindergarten students had a slightly higher mean in reading (Appendix C, p. 16). The findings were not statistically significant. However, a statistically significant positive effect of HISD prekindergarten was observed for Cohort II students at fifth grade in mathematics (Appendix C, p. 17). Cohort III comparison-group students outperformed HISD prekindergarten students at fifth grade in reading, with the results reflecting statistical significance (Appendix C, p. 18).

The intermediate effects at seventh grade on the Stanford reading and mathematics tests of HISD prekindergarten were also measured. Aprenda is typically not administered at this grade level in HISD. The findings varied among the groups. Cohorts I and III HISD prekindergarten students had a statistically-significant lower mean Stanford reading score than non-HISD prekindergarten students, while Cohort II attained a higher mean reading score that was statistically insignificant (Appendix C, pp. 16–18). In each case, the effect sizes were small for Cohorts I (d=-.082),

Figure 6c: Cohort III - Short-term kindergarten Stanford and Aprenda reading and math results (NCEs rounded to nearest whole number)

Cohort II (.012), and Cohort III (d=-.072) (Appendix C). On the Stanford mathematics test at seventh grade, HISD prekindergarten students in Cohorts I and III attained lower mean scores and HISD prekindergarten students in Cohort II attained a higher mean score that were statistically insignificant (Appendix C, pp. 16–18).

The short-term and intermediate effects of HISD prekindergarten on attendance were consistently in favor of HISD prekindergarten students at kindergarten, third, and fifth grades. The largest effects were observed at third grade in Cohort I (d=.230), Cohort II (d=.219), and Cohort III (d=.241); and at fifth grade in Cohort II (d=..220) (Appendix C, pp. 16–18). The estimated effects were positive and statistically significant.

The odds of in-school suspensions at fifth grade (short-term) were greater for HISD prekindergarten students than non-HISD prekindergarten students in Cohorts I and II, and greater for out-of-school suspensions in Cohort I (Appendix C). However, there was insufficient evidence to conclude that the groups were statistically signficanly different. At the intermediate level in seventh grade, the odds were greater for HISD prekindergarten Cohort II and III students to have a higher rate of in-school suspensions and Cohort I students to have a higher rate of out-of-school suspensions than non-HISD prekindergarten students. However, Cohort II and III HISD prekindergarten students had a slightly lower rate of out-of-school suspensions compared to non-HISD prekindergarten students at seventh grade. This finding for Cohort III reflected a statistically significant positive effect.

What were the long-term effects of HISD prekindergarten on college preparedness?

Multiple indicators were used to measure the long-term impact of HISD prekindergarten on college preparedness, including tenth-grade PSAT, twelfth-grade SAT, and twelfth-grade ACT reading and mathematics scores. In addition, students' ninth-grade attendance and ninth-grade GPAs were used to determine long-term impact in this area. Means were calculated for the measures in each cohort. Benchmark scores were obtained from College Board and ACT, Inc. reports to compare minimum reading and mathematics scores indicative of college readiness (ACT, 2013; Marini, Beard, & Shaw, 2018; Wyatt, Smith, & Proestler, 2014). The percentage of students who met benchmarks were compared with cohort group means. The findings for Cohorts I, II, and III are depicted in **Figures 7a**, **7b**, and **7c** (p. 8).

It is evident that the mean reading and mathematics scores for HISD prekindergarten and non-HISD prekindergarten students fell below PSAT, SAT, and ACT reading and mathematics benchmarks in each cohort (Figures 7a, 7b, and 7c). In addition, non-HISD prekindergarten students in each cohort attained a higher mean score than HISD prekindergarten students on all tests, except for Cohort II on the PSAT mathematics test. This difference was not statistically significant (Appendix C). The differences between the groups that yielded statistical significance were observed on all tests for Cohort I students and on the SAT and ACT mathematics tests for Cohort II students. For Cohort III students, statistically significant differences were observed in favor of non-HISD prekindergarten students on the PSAT reading, the SAT reading and mathematics, and the ACT mathematics tests.

Relative to attendance, mean ninth-grade attendance rates
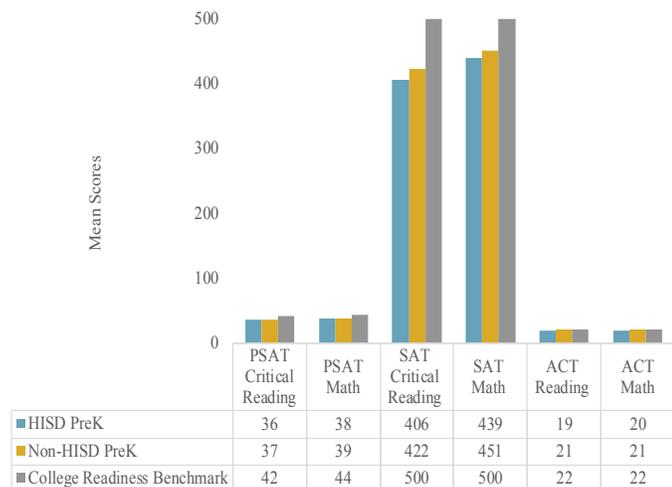
Figure 7a: Cohort I PSAT (10th gr), SAT (12th gr), and ACT (12th gr) reading and math performance (Scores rounded to nearest whole number)

	PSAT Critical Reading	PSAT Math	SAT Critical Reading	SAT Math	ACT Reading	ACT Math
HISD PreK	36	38	406	439	19	20
Non-HISD PreK	37	39	422	451	21	21
College Readiness Benchmark	42	44	500	500	22	22

and effect sizes were calculated to determine the differences between the groups. The study found that there was a statistically significant positive effect of HISD prekindergarten on students' attendance in Cohort I (94.34 vs. 93.83), Cohort II (94.55 vs. 93.55), and Cohort III (94.43 vs. 93.28) (Appendix C).

The ninth-grade mean GPAs for students in Cohorts I and III were lower for HISD prekindergarten students compared to non-HISD prekindergarten students. In contrast, HISD prekindergarten students in Cohort II had a higher GPA. The differences observed in all cohorts between the groups were not statistically significant.

To what extent did HISD prekindergarten affect school persistence relative to graduation and school dropout?

Students' persistence in school was measured using the graduation and dropout rates among the total students remaining in the cohort. The sample sizes are provided in Tables 2a, 2b, and 2c in Appendix C (pp. 16–18) for the respective cohorts. For Cohort I, early graduation was indicated at 11 years (2013–2014),
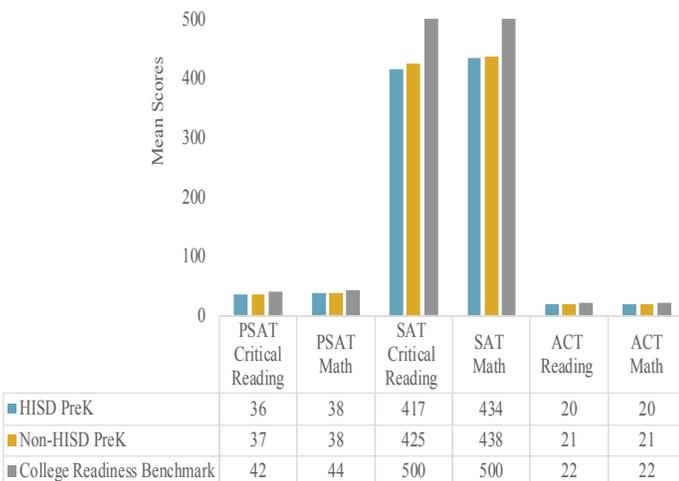
Figure 7b: Cohort II - PSAT (10th gr), SAT (12th gr), and ACT (12th gr) reading and math performance (Scores were rounded to nearest whole number.)

	PSAT Critical Reading	PSAT Math	SAT Critical Reading	SAT Math	ACT Reading	ACT Math
HISD PreK	36	38	417	434	20	20
Non-HISD PreK	37	38	425	438	21	21
College Readiness Benchmark	42	44	500	500	22	22

on-time graduation at 12 years (2014–2015), and late graduation was at 13 years (2015–2016), considering the 12-year expected graduation benchmark starting with first grade. A similar method was used to determine graduation status for Cohorts II and III. For Cohort II, early graduation was observed in 2014–2015, on-time graduation in 2015–2016, and late graduation in 2016–2017. Finally, for Cohort III, early graduation was reflected in 2015–2016, on-time graduation in 2016–2017, and late graduation in 2017–2018. A limitation is that students could have graduated from other school districts in Texas or in other states.
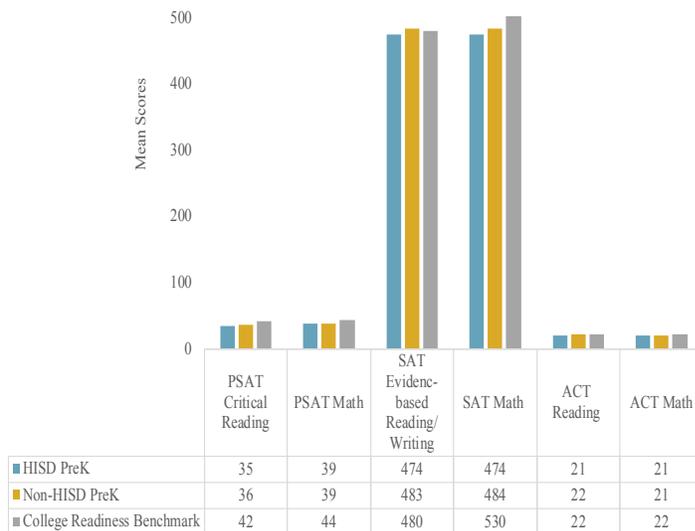
Figure 7c: Cohort III - PSAT (10th gr), SAT (12th gr), and ACT (12th gr) reading and math performance (Scores were rounded to nearest whole number.)

	PSAT Critical Reading	PSAT Math	SAT Evidence-based Reading/Writing	SAT Math	ACT Reading	ACT Math
HISD PreK	35	39	474	474	21	21
Non-HISD PreK	36	39	483	484	22	21
College Readiness Benchmark	42	44	480	530	22	22

Overall, a higher percentage of HISD prekindergarten students compared to non-HISD prekindergarten students graduated on time in Cohort I (81.0% vs. 76.3%,), Cohort II (99.2% vs. 94.4%), and Cohort III (98.8% vs. 98.5%) (Tables 2a, 2b, and 2c, respectively, Appendix C).

In addition, HISD prekindergarten Cohort I students were 1.32 times more likely than non-HISD prekindergarten students to graduate on time (CI=0.127, 0.432), and equally likely to graduate early (OR=.998, CI=0.413, 2.411) compared to non-HISD prekindergarten students. Moreover, HISD prekindergarten had a statistically positive effect on Cohort I's on-time graduation rate.

Cohort II HISD prekindergarten students were equally likely to graduate on time (OR=.998, (CI=0.901, 1.106), but more likely to graduate early than non-HISD prekindergarten students (OR=1.306, CI=0.528, 3.231). For Cohort III students, the odds of graduating on time was slightly higher for HISD prekindergarten students compared to non-prekindergarten students (OR=1.003, CI=0.904, 1.112). The results for Cohorts II and III were not statistically significant based on the data.

Tables 2a, 2b, and 2c in Appendix C depict the percent of Cohort I, II, and III students in both groups who dropped out of school within 7 years of expected graduation at 12 years. It is evident that the mean dropout rate was lower for HISD prekindergarten students compared to non-HISD prekindergarten students. However, there was insufficient evidence to conclude that the groups were statistically significantly different.

What were the best predictors of school persistence for Cohort I, II, and III student groups, considering their demographic characteristics?

Multinomial logistic regression analyses were conducted to determine the best predictors of on-time (at 12 years), early (11 years), and late (13 years) graduation for Cohort I, II, and III student groups. The findings are depicted for the respective cohorts in Tables 3a, 3b, and 3c. The model controlled for economic status, gender, limited English proficient (LEP), gifted/talented, special education status, and whether students participated in HISD prekindergarten.

In Cohort I (**Table 3a**), it was more likely that students who attended HISD prekindergarten graduated on-time compared to students who did not attend HISD prekindergarten (B=.294, p<.0001). In addition, HISD prekindergarten, students who were not at risk (B=-.633, p<.0001), gifted/talented (B=.613), LEP (B=.592, p<.0001), and female (B=-.240, p<.0001) made significant contributions to on-time graduation. A similar pattern was observed among cohort students who graduated late, except students who were at risk were more likely to graduate late than students who were not at risk. No significant findings were found at early graduation.

Cohort II HISD prekindergarten students (**Table 3b**) were more likely than non-HISD prekindergarten to graduate on-time (at 12 years) (B=.314, p<.0001). In addition, HISD prekindergarten, students who were not at risk (B=-.786, p<.0001), LEP (B=.599, p<.0001), gifted/talented (B=.535, p<.0001), and female (B=-.217, p<.0001), made significant contributions to on-time graduation. No significant findings were observed for early graduation, and no students were observed in the data to indicate late graduation

The findings for Cohort III students are depicted in **Table 3c**, revealing that these students were more likely than non-HISD prekindergarten students to graduate on-time (at 12 years). HISD prekindergarten, students who were gifted/talented (B=.655, p<.0001), not at risk (B=-.573, p<.0001), LEP (B=.407, p<.0001), and female (B=-.188) made statistically significant contributions to on-time graduation. Being a

male made a statistically significant contribution to early graduation (B=.834), and no students were observed in the data for late graduation as this information is not available at the time of this reporting.

Table 3b: Cohort II - Mutinomial Logistic Regression Predicting Graduation of Cohort Participants

	On-Time (12 yrs: 2015–2016) B	Early (11 yrs: 2014–2015) B	Late (13 yrs: 2016–2017)† B
Variable			
Constant	.420	-20.569	-
HISD Prekindergarten Status	.314***	-.296	-
Gender (1=M, 0=F)	-.217***	.750	-
Economic Status	-.268	-.375	-
Gifted/Talented	.535***	-.259	-
Special Education	..222	15.904	-
At Risk	-.786***	-.147	-
LEP	.599***	.242	-

*** p < .0001, ** p < .001; R2=.035
Data Source: PEIMS graduation file
†Information limited at the time of reporting

Table 3c: Cohort III - Mutinomial Logistic Regression Predicting Graduation of Cohort Participants

	On-Time (12 yrs: 2016–2017) B	Early (11 yrs: 2015–2016) B	Late (13 yrs: 2017–2018)† B
Variable			
Constant	.137	-5.906	-
HISD Prekindergarten Status	.352***	.240	-
Gender (1=M, 0=F)	-.188***	.834**	-
Economic Status	-.144	-.170	-
Gifted/Talented	.655***	-.259	-
Special Education	.139	.498	-
At Risk	-.573***	-.415	-
LEP	.407***	.183	-

*** p < .0001, ** p < .001; R2=.031
Data Source: PEIMS graduation file
†Information not available at the time of reporting

Table 3a: Cohort I - Mutinomial Logistic Regression Predicting Graduation of Cohort Participants

	On-Time (12 yrs: 2014–2015) B	Early (11 yrs 2013–2014) B	Late (13 yrs 2015–2016) B
Variable			
Constant	0.27	-21.101	-2.862
HISD Prekindergarten Status	.294***	.094	.269**
Gender (1=M, 0=F)	-.240***	.158	-.358***
Economic Status	-	-	-
Gifted/Talented	.613***	.139	1.970***
Special Education	.018	15.904	-.349
At Risk	-.633***	.-.742	.579***
LEP	.592***	.263	.579***

*** p < .0001, ** p < .001; R2=0.36
Data Source: PEIMS graduation file

Discussion

The prekindergarten program is a complex subsystem of early childhood education that has the responsibility of promoting the equitable development, learning, and school readiness of all children. Each child, regardless of their abilities, should be respected and carefully included to the fullest extent and with the highest expectations in prekindergarten programs (NAEYC, 2009; NAEYC & NAECS/SDE, 2003). To determine the effects of prekindergarten in HISD, this study hypothesized that participation will consistently correlate with better short-term, intermediate,

and long-term academic, behavioral, college preparedness, and school persistence outcomes. HISD students who were included in the study were exposed to half-day prekindergarten based on a Texas instructional model of language and literacy, mathematics, science, social studies, fine arts, physical education, and technology.

This study controlled for preexisting factors, including kindergarten enrollment, ethnicity, gender, special education, gifted/talented, at risk, and English proficiency, which had the potential to affect student outcomes. The three cohorts of propensity score matched students identified in this study attended HISD prekindergarten or were eligible to attend HISD prekindergarten in the district during the 2001–2002 (Cohort I), 2002–2003 (Cohort II), and 2003–2004 (Cohort III) academic years.

The study found that, at the kindergarten level, students in Cohorts I, II, and III outperformed students who did not attend HISD prekindergarten on nationally-normed Stanford and Aprenda reading and mathematics tests. Thus, HISD prekindergarten had a statistically significant positive effect on students' academic performance at kindergarten. Moreover, the positive effects of HISD prekindergarten was extended for Cohort II students who consistently outperformed non-HISD prekindergarten students at third, fifth, and seventh grades on these tests.

Long-term effects of HISD prekindergarten on the PSAT, SAT, and ACT reading and mathematics tests were mostly in favor of non-HISD prekindergarten students in all cohorts. Specifically, the study found that college preparedness, based on PSAT, SAT, and ACT performance, were below College Board benchmarks. However, students scoring below the benchmark may still do well and succeed in high school and college, considering that college success requires a wide range of cognitive and non-cognitive knowledge and skills, including motivation, self-regulation, and perseverance (Wyatt, Smith, & Proestler, 2014).

While intermediate and long-term academic effects of HISD prekindergarten were typically insignificant between the groups, attendance rates of HISD prekindergarten students were significantly higher compared to non-HISD prekindergarten students at short-term, intermediate, and long-term measurement levels. Moreover, a higher proportion of HISD prekindergarten students graduated at 12 years, the expected time of graduation, compared to non-HISD prekindergarten students in all cohorts.

Consistent with a recent study conducted by the Texas Education Agency (2017), this study found that, over time, the positive effects of prekindergarten in all areas may not be sustained. Nevertheless, this study demonstrated that educating at-risk students during their early years has educational benefits, especially on school persistence. Evidence indicated that students identified as at risk and who qualify for prekindergarten services require continued support as they progress through school. The study highlights the importance of maintaining academically-stimulating environments at all levels of education. Future research is needed to determine whether full-day HISD prekindergarten provides more lasting effects for students. The research should consider tracking the progress of HISD prekindergarten students to ensure that the instructional model continues to build students' knowledge and skills acquired during prekindergarten.

References

ACT. (2013). What Are the ACT College Readiness Benchmarks? Retrieved from http://www.act.org/content/dam/act/unsecured/documents/benchmarks.pdf

Almond, D., & Currie, J. (2011). Handbook of Labor Economics. Vol. 4B, Chapter Human Capital Development Before Age 5, 1315-1486. Amsterdam: Elsevier.

Andrews, R. J., Jargowsky, P., & Kuhne, K. (2012). The Effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance. National Bureau of Economic Research.

Ansari, A., & Winsler, A. (2013). Stability and sequence of center-based and family childcare: Links with low-income children's school readiness. Children and Youth Services Review, 35(2), 358-366.

Barnett, W. (1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. The Future of Children. Retrieved from https://www.princeton.edu/futureof-children/publications/docs/05_03_01.pdf

Burchinal, M., Mashburn, A., Pianta, R., & Vandergrift, N. (2009). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. Early Childhood Research Quarterly, 25, 166-167.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. (2010). Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development. Teachers College Record. 112. Retrieved from http://gregorycamilli.info/papers/early%20education%20interventions.pdf

Cochran, W.G., & Rubin, D.B. (1973). Controlling bias in observational studies: a review. Sankhya: Indian J Stat, Ser A. 1973; 35(4):417–446.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. Routledge.

Fuller, B., Bein, E., Bridges, M., Kim, Y., & Rabe-Hesketh, S. (2017). Do academic preschools yield stronger benefits? Cognitive emphasis, dosage, and early learning, Journal of Applied Developmental Psychology, Volume 52, p. 1-11, ISSN 0193-3973, https://doi.org/10.1016/j.appdev.2017.05.001.

Frede, E., Jung, K., Barnett, W. S., & Figueras, A. (2009). The APPLES blossom: Abbott preschool program longitudinal effects study (APPLES) preliminary results through 2nd grade (Interim Report). New Brunswick, NJ: Rutgers University, National Institute for Early Education Research. Retrieved from http://ieer.org/pdf/apples_second_grade_results.pdf

Hansen, B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. Statistical Science, 2, 219-236.

Heckman, J. J. (2008). Schools, Skills and Synapses. Economic Inquiry, 46(3): 289-324.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Analysis, 15, 199–236. Google Scholar, Crossref

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. Journal of Statistical Software, 42, 1–28. Retrieved from http://www.jstatsoft.org/v42/i08/ Google Scholar

Houston Independent School District. (n.d.). Early Childhood. Enrollment Process. Retrieved from http://www.houstonisd.org/Page/126445

Houston Independent School District. (2017). HISD Curriculum, Instruction, and Development. Prekindergarten Guidelines. Retrieved from http://www.houstonisd.org/PreK

Joffe, M. M., & Rosenbaum, P. R. (1999). Propensity Scores.

American Journal of Epidemiology, 150, 327–333.

Knudsen, E. I., Heckman, J. J., Cameron, J., Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proceedings of the National Academy of Sciences 103 (27), 10155–10162.

Lipsey, M. W., Farran, D., & Hofer, K. (2015). A Randomized Control Trial of a StatewideVoluntary Prekindergarten Program on Children's Skills and Behaviors through Third Grade. Nashville, TN: Vanderbilt University, Peabody Research Institute. Retrieved from https://peabody.vanderbilt.edu/research/pri/VPK-through3rd_final_withcover.pdf

Lunt, M. (2013). Selecting an Appropriate Caliper Can Be Essential for Achieving Good Balance With Propensity Score Matching. American Journal of Epidemiology. Vol. 179, No. 2. DOI: 10.1093/aje/kwt212

Marini, J., Beard, J., & Shaw, E. (2018). Student Ranking Differences within Institutions Using Old and New SAT Scores. ERIC. Retrieved from https://eric.ed.gov/?q=%22%22&ff1=-souCollege+Board&ff2=subCareer+Readiness&ff3=subCollege+Readiness&id=ED581513

Mellon, E. (2006). Pre-K HISD students putting in a 7-hour day. Houston Chronicle. Retrieved from https://www.chron.com/news/houston-texas/article/Pre-K-HISD-students-putting-in-a-7-hour-day-1865794.php

McCoy, D., Yoshikawa, H., Ziol-Guest, K., Duncan, G., Schindler, H., Magnuson, K.,Yang, R., Koepp, A., & Shonkoff, J. (2017). Impacts of Early Childhood Education on Medium- and Long-Term Educational Outcomes. Educational Researcher, Vol 46, Issue 8, pp. 474 - 487.

Murnane, R. J., & Willett, J. B. (2011). Methods matter: Improving causal inference in educational and social science research. New York, NY: Oxford University Press.

National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education. (2003). Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. Position Statement. Washington DC: NAEYC.

National Association for the Education of Young Children. (2009). Developmentally appropriate practice in early childhood programs serving children from birth through age 8. Position statement. Washington, DC: NAEYC.

Nores, M., & Barnett, W.S. (2014). Access to High Quality Early Care and Education: Readiness and Opportunity Gaps in America (CEELO Policy Report). New Brunswick, NJ: Center on Enhancing Early Learning Outcomes.

Rosenbaum, P.R., & Rubin, D.B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat. 1985;39(1):33–38.

Rubin, D. B. (1997). Estimating causal effects from large datasets using propensity scores. Annals of Internal Medicine, 127(8), 757–763.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). Lifetime effects: The HighScope Perry Preschool study through age 40. (Monographs of the HighScope Educational Research Foundation, 14). Ypsilanti, MI: HighScope Press.

Texas Education Agency (1999). Prekindergarten Curriculum Guidelines.

Texas Education Agency. (2008). Revised Texas prekindergarten guidelines. Retrieved from http://www.tea.state.tx.us/ed_init/pkguidelines/PKG_Final_100808.pdf

Texas Education Agency. (2009). Kindergarten Frequently Asked Questions and Answers. https://tea.texas.gov/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=5921&libID=5933

Texas Education Agency. (2011). Early Childhood Outcomes and Prekindergarten Guidelines Alignment. Retrieved from http://ontrack-media.net/AF/TEA/PDF/L3_0-SE_File%201_EC_Guidelines_Alignment.pdf

Texas Education Agency. (2017). Prekindergarten Outcomes for Texas Public School Students. Retrieved from file:///C:/Users/vholmes/Downloads/Prekindergarten_Longitudinal_Study.pdf

Thoemmes, F. (2012). Propensity Score Matching in SPSS. Retrieved from https://www.human.cornell.edu/sites/default/files/HD/qml/Thoemmes_2012.pdf

Thoemmes, F., & Kim, E. S. (2011). A Systematic Review of Propensity Score Methods in the Social Sciences. Multivariate Behavioral Research, 46, 90-118.

U.S. Department of Education, National Center for Education Statistics (2011). Early childhood longitudinal study, Kindergarten cohort (ECLS-K) [2011 restricted-use Kindergarten data file and electronic codebook]. Institute of Education Sciences.

What Works Clearinghouse. Standards Handbook, Version 4.0. (n.d.). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Wyatt, J., Smith, K., & Proestler, N. (2014). Process: Implications for Practitioners. College Board. Retrieved from https://files.eric.ed.gov/fulltext/ED556463.pdf

Appendix A

Cohort I Pre-Intervention Population Sample Sizes and Characteristics for the Baseline Sample (Cohort Year: 2001-2002*)				
Sample Characteristics (N= 14,219)	Sample Size	% within group	Sample Size	% within group
Baseline Measures	Treatment (n = 8,123)		Controls (n=6,096)	
Race/Ethnicity				
Asian	167	2.1	251	4.1
Black	2,107	25.9	1,569	25.7
Hispanic	5,745	70.7	3,250	53.3
White	104	1.3	1,026	16.8
Gender				
Male	4,083	50.3	3,121	51.2
Female	4,042	49.7	2,977	48.8
Eco Disadv.	8,125	100.0	6,098	100.0
At Risk	6,466	79.6	3,967	65.1
Special Ed	247	3.0	386	6.3
G/T	550	6.8	782	12.8
LEP	4,725	58.2	2,191	35.9

Note: Due to the small sample size (n = 4), American Indian and Alaskan students were not included in the analyses.
*Cohort year indicates the year that students were eligible or attended HISD prekindergarten

Cohort I Pre-Intervention Population Sample Sizes and Characteristics for the Analytic Sample based on Propensity Score Matching (Cohort Year: 2001-2002*)				
Sample Characteristics (N = 12,046)	Sample Size	% within group	Sample Size	% within group
Baseline Measures	Treatment (n = 8,123)		Control (n = 3,923)	
Race/Ethnicity				
Asian	167	2.1	81	2.1
Black	2,107	25.9	1,140	29.1
Hispanic	5,745	70.7	2,608	66.5
White	104	1.3	94	2.4
Gender				
Male	4,083	50.3	2,002	51.0
Female	4,042	49.7	1,921	49.0
Eco Disadv.	8,125	100.0	3,923	100.0
At Risk	6,466	79.6	2,942	75.0
Special Ed	247	3.0	178	4.5
G/T	790	6.5	240	6.1
LEP	4,725	58.2	1,902	48.5

Note: Due to the small sample size (n = 4), American Indian and Alaskan students were not included in the analyses.
*Cohort year indicates the year that students were eligible or attended HISD prekindergarten

Cohort II Pre-Intervention Population Sample Sizes and Characteristics for the Baseline Sample (Cohort Year: 2002-2003*)				
Sample Characteristics (N= 13,837)	**Sample Size**	**% within group**	**Sample Size**	**% within group**
Baseline Measures	Treatment (n = 8,092)		Controls (n=5,745)	
Race/Ethnicity				
Asian	160	2.0	236	4.1
Black	1,929	23.8	1,453	25.3
Hispanic	5,869	72.5	2,964	51.6
White	134	1.7	1,092	19.0
Gender				
Male	4,109	50.8	2,954	51.4
Female	3,987	49.2	2,794	48.6
Eco Disadv.	7,700	95.1	4,117	71.6
At Risk	6,491	80.2	3,666	63.8
Special Ed	263	3.2	370	6.4
G/T	618	7.6	820	14.3
LEP	4,677	57.8	1,867	32.5

Note: Due to the small sample size (n = 5), American Indian or Alaskan students were not included in the analyses.
*Cohort year indicates the year that students were eligible or attended HISD prekindergarten

Cohort II Pre-Intervention Population Sample Sizes and Characteristics for the Analytic Sample based on Propensity Score Matching (Cohort Year: 2002-2003*)				
Sample Characteristics (N = 11,746)	**Sample Size**	**% within group**	**Sample Size**	**% within group**
Baseline Measures	Treatment (8,092)		Control (3,654)	
Race/Ethnicity				
Asian	160	2.0	107	2.9
Black	1,929	23.8	1,043	28.5
Hispanic	5,869	72.5	2,424	66.3
White	134	1.7	80	2.2
Gender				
Male	4,109	50.8	1,857	50.8
Female	3,987	49.2	1,799	49.2
Eco Disadv.	7,699	95.1	3,350	91.6
At Risk	6,490	80.2	2,772	75.8
Special Ed	263	3.2	190	5.2
G/T	618	7.6	221	6.0
LEP	4,676	57.8	1,674	45.8

Note: Due to the small sample size (n = 5), American Indian or Alaskan students were not included in the analyses.
*Cohort year indicates the year that students were eligible or attended HISD prekindergarten

Cohort III Pre-Intervention Population Sample Sizes and Characteristics for the Baseline Sample (Cohort Year: 2003-2004*)				
Sample Characteristics (N= 13,420)	**Sample Size**	**% within group**	**Sample Size**	**% within group**
Baseline Measures	Treatment (n = 8,307)		Controls (n=5,113)	
Race/Ethnicity				
Asian	171	2.1	230	4.5
Black	1941	23.4	1255	24.5
Hispanic	6086	73.2	2644	51.7
White	109	1.3	984	19.2
Gender				
Male	4161	50.1	2621	51.2
Female	4149	49.9	2497	48.8
Eco Disadv.	7832	94.2	478	5.8
At Risk	7003	84.3	3552	69.4
Special Ed	339	4.1	336	6.6
G/T	772	9.3	898	17.5
LEP	4724	56.8	1598	31.2

Note: Due to the small sample size (n = 8), American Indian or Alaskan students were not included in the analyses.
*Cohort year indicates the year that students were eligible or attended HISD prekindergarten

Cohort III Pre-Intervention Population Sample Sizes and Characteristics for the Analytic Sample based on Propensity Score Matching (Cohort Year: 2003-2004*)				
Sample Characteristics (N = 11,652)	**Sample Size**	**% within group**	**Sample Size**	**% within group**
Baseline Measures	Treatment (8,307)		Control (3,345)	
Race/Ethnicity				
Asian	171	2.1	103	3.1
Black	1,941	23.4	975	29.1
Hispanic	6,086	73.2	2,204	65.9
White	109	1.3	63	1.9
Gender				
Male	4,161	50.1	1,744	52.1
Female	4,149	49.9	1,602	47.9
Eco Disadv.	7,832	94.2	3022	90.3
At Risk	7,003	84.3	2,665	79.6
Special Ed	339	4.1	180	5.4
G/T	772	9.3	292	8.7
LEP	4,724	56.8	1,440	43.0

Note: Due to the small sample size (n = 8), American Indian or Alaskan students were not included in the analyses.
*Cohort year indicates the year that students were eligible or attended HISD prekindergarten

Appendix B

WWC Characterization of Findings of an Effect Based on a Single Outcome Measure

Statistically significant positive effect	The estimated effect is positive and statistically significant (correcting for clustering when not properly aligned).
Substantively important positive effect	The estimated effect is positive and not statistically significant but is substantively important.
Indeterminate effect	The estimated effect is neither statistically significant nor substantively important.
Substantively important negative effect	The estimated effect is negative and not statistically significant but is substantively important.
Statistically significant negative effect	The estimated effect is negative and statistically significant (correcting for clustering when not properly aligned).

Note: A statistically significant estimate of an effect is one for which the probability of observing such a result by chance is less than one in 20 (using a two-tailed t-test with $p = 0.05$). A properly aligned analysis is one for which the unit of assignment and unit of analysis are the same. An effect size of 0.25 standard deviations or larger is considered to be substantively important.

Source: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

Table 2a: Cohort I (2001–2002*) Post-Intervention Outcomes for the Analytic Sample and Estimated Effects								
	Intervention Group (N=8,123)			Comparison Group (N=3,923)			Estimated Effects		
Outcome Measures (grade level)	n	Mean (% within group)	Standard Devia.	n	Mean (% within group)	Standard Devia.	Mean Diff.	Effect Size (Odds Ratio)	p-value (95% CI)
Short-term									
Stanford Reading (K)	4161	52.54	18.311	2218	50.39	19.446	2.15	.115	.000***
Stanford Math (K)	4161	47.70	20.245	2218	45.57	20.343	2.13	.105	.000***
Attendance (K)	4304	95.62	5.063	2357	94.62	5.310	1.00	.194	.000***
Aprenda Reading (K)	3794	60.12	22.025	1565	52.90	22.418	7.22	.326	.000***
Aprenda Math (K)	3794	56.84	19.995	1565	50.35	20.397	6.49	.323	.000***
Stanford Reading (3rd)	3122	51.36	17.429	1486	52.22	19.395	-.872	-.047	.135
Stanford Math (3rd)	3122	56.12	19.648	1486	56.38	20.699	-.677	-.013	.677
Stanford Reading (5th)	4314	45.70	15.765	1754	46.50	17.344	-.80	-.049	.080
Stanford Math (5th)	4314	55.19	16.663	1754	54.65	17.615	.54	.032	.257
Attendance (5th)	5855	97.68	3.124	2614	96.98	4.210	.70	.200	.000***
In-school Suspensions (5th)	855	(10.5%)	-	375	(9.6%)	-	-	(OR=1.007)	(CI=0.828, 1.225)
Out-of-school Suspensions (5th)	541	(6.7%)	-	234	(6.0%)	-	-	(OR=1.027)	(CI=0.845, 1.249)
Intermediate Outcomes									
Stanford Reading (7th)	3469	46.07	15.309	1397	47.38	17.144	-1.30	-.082	.010*
Stanford Math (7th)	3469	55.17	16.748	1397	55.67	17.614	-.50	-.029	.367
Attendance (7th)	4871	95.94	4.989	2176	95.32	5.641	.62	.119	.000***
In-school Suspensions (7th)	501	(6.2%)	-	213	(5.4%)	-	-	(OR=.993)	(CI=0.792, 1.243)
Out-of-school Suspensions (7th)	342	(4.2%)	-	134	(3.4%)	-	-	(OR=1.122)	(CI=0.881, 1.429)
Long-term Outcomes									
College Preparedness									
Attendance (9th)	4604	94.34	8.339	2025	93.83	8.369	.50	.061	.024*
GPA (9th)	852	2.70	.942	347	2.76	.986	-.06	-.062	.301
PSAT (10th)									
Critical Reading	3770	36.40	7.07	1585	37.204	7.920	-.8076	.109	.000***
Math	3770	38.18	8.38	1584	38.61	8.990	-.4291	.050	.022*
SAT Reading (12th)	2636	405.56	94.052	1020	422.07	106.493	-16.51	-.163	.000***
SAT Math (12th)	2636	439.11	93.768	1020	450.637	101.045	-11.53	-.120	.001**
ACT Reading (12th)	501	18.74	6.144	212	20.62	6.886	-1.88	-.294	.000***
ACT Math (12th)	501	19.90	5.027	212	21.18	5.558	-1.28	-.246	.003**
School Persistence									
Graduation: Total N Graduates (Data Source: PEIMS Graduation file, 2013–2014 through 2015–2016)	3296			1354					
Graduation (Early: 11 yrs, 2013–2014)	17	(.5%)	-	7	(.5%)	-	-	(OR=.998)	(CI=0.413, 2.411)
Graduation (On-time: 12 yrs, 2014–2015)	2669	(81.0%)	-	1033	(76.3%)	-	-	(OR=1.323)	(CI=0.127, 0.432)
Graduation (Late: 13 yrs, 2015–2016)	610	(18.5%)	-	314	(23.2%)	-	-	(OR=.752)	(CI=0.645, 0.877)
Dropped Out (within 7 years of on-time graduation year: between 2008–2009 and 2014–2015	276	(3.4%)	-	148	(3.8%)	-	-	(OR=.897)	(CI=0.732, 1.1)

Table 2b: Cohort II (2002–2003*) Post-Intervention Outcomes for the Analytic Sample and Estimated Effects								
	Intervention Group (8,092)			Comparison Group (3,654)			Estimated Effects		
Outcome Measures (grade level)	n	Mean	Standard Devia.	n	Mean	Standard Devia.	Mean Diff.	Effect Size (Odds Ratio)	p-value (95% CI)
Short-term									
Stanford Reading (K)	4290	57.92	21.919	2144	53.01	22.445	4.10	.222	.000***
Stanford Math (K)	4290	48.73	22.02	2144	44.97	22.288	3.76	.170	.000***
Attendance (K)	4403	95.72	4.628	2259	94.89	5.031	.828	.046	.000***
Aprenda Reading (K)	3678	61.79	22.099	1400	53.87	22.778	7.92	.355	.000***
Aprenda Math (K)	3678	58.14	19.718	1400	51.06	20.810	7.08	.354	.000***
Stanford Reading (3rd)	3194	52.58	17.535	1387	52.40	19.371	.177	.010	.761
Stanford Math (3rd)	3194	57.52	19.347	1387	55.87	20.995	1.55	.083	.015*
Attendance (3rd)	6719	97.32	3.491	2872	96.49	4.426	.83	.2186	.000***
Stanford Reading (5th)	4421	49.13	16.517	1643	48.35	18.423	.777	.046	.115
Stanford Math (5th)	4421	55.97	16.517	1643	54.13	18.499	1.833	.108	.000***
Attendance (5th)	5899	97.68	3.303	2404	96.89	4.20	.790	.220	.000***
In-school Suspensions (5th)	721	(8.9%)	-	287	(7.9%)	-		(OR=1.084)	(CI=0.885, 1.329)
Out-of-school Suspensions (5th)	465	(5.7%)	-	217	(5.9%)	-		(OR=0.819)	(CI=0.666, 1.007)
Intermediate Outcomes									
Stanford Reading (7th)	3531	47.65	17.357	1305	47.44	19.368	.210	.012	.718
Stanford Math (7th)	3531	59.35	18.629	1305	58.23	20.446	1.12	.059	.070
Attendance (7th)	4872	96.41	4.517	1974	95.83	5.082	.5825	.124	.000***
In-school Suspensions (7th)	462	(5.7%)	-	187	(5.1%)	-		(OR=1.091)	(CI=0.855, 1.392)
Out-of-school Suspensions (7th)	261	(3.2%)	-	127	(3.5%)	-		(OR=0.811)	(CI=0.626, 1.05)
Long-term Outcomes									
College Preparedness									
Attendance (9th)	4527	94.55	8.230	1822	93.55	8.890	.996	.119	.000***
GPA (9th grade)	1062	3.03	.941	348	3.02	.941	.005	.010	.928
PSAT (10th)									
Critical Reading	3875	36.25	8.50	1476	36.52	8.940	-.2662	.031	.313
Math	3869	38.27	8.36	1474	38.16	8.900	.1040	.013	.690
SAT Reading (12th)	2716	417.02	99.039	956	424.74	104.141	-7.721	-.076	.213
SAT Math (12th)	2716	433.52	97.696	956	438.21	107.446	-4.695	-.046	.041*
ACT Reading (12th)	557	20.052	6.277	237	20.692	6.195	-.639	-.102	.187
ACT Math (12th)	557	20.016	5.361	237	21.025	5.797	-1.009	-.183	.018*
School Persistence									
Graduation: Total N Graduates (Data Source: PEIMS Graduation file, 2014–2015 through 2016–2017. The 2016–2017 data may have been limited at the time of this reporting.)	2782			991					
Graduation (Early: 11 yrs, 2014–2015)	22	(.8%)	-	6	(.6%)	-	-	(OR=1.306)	(CI=0.528, 3.231)
Graduation (On-time: 12 yrs, 2015–2016)	2760	(99.2%)	-	985	(94.4%)	-	-	(OR=.998)	(CI=0.901, 1.106)
Graduation (Late: 13 yrs, 2016–2017)	0	-	-	0	-	-	-	-	-
Dropped Out (within 7 years of on-time graduation year: between 2009–2010 and 2015–2016	8	(.09%)	-	9	(,25%)	-	-	(OR=.401)	(CI=0.155, 1.041)

Table 2c: Cohort III (2003–2004*) Post-Intervention Outcomes for the Analytic Sample and Estimated Effects								
	Intervention Group (N=8,307)			Comparison Group (N=3,345)			Estimated Effects		
Outcome Measures (grade level)	n	Mean	Standard Devia.	n	Mean	Standard Devia.	Mean Diff.	Effect Size	p-value (95% CI)
Short-term									
Stanford Reading (K)	4210	57.05	21.801	747	52.71	22.942	4.34	.197	.000**
Stanford Math (K)	4210	47.69	21.162	747	44.35	21.756	3.34	.157	.000**
Attendance (K)	4472	95.63	4.687	2888	95.13	4.783	0.50	.106	.000**
Aprenda Reading (K)	3891	62.39	21.741	1218	54.53	22.432	7.86	.368	.000**
Aprenda Math (K)	3891	57.98	19.371	1218	52.28	20.329	5.70	.291	.000**
Stanford Reading (3rd)	3274	46.32	16.994	1307	46.46	18.848	-.138	-.007	.809
Stanford Math (3rd)	3274	52.07	19.333	1307	51.97	20.449	.103	.005	.873
Attendance (3rd)	7041	97.59	3.20	2602	96.74	4.295	.841	.241	.000***
Stanford Reading (5th)	4761	46.73	15.441	1606	47.98	17.673	-1.25	-.077	.007**
Stanford Math (5th)	4761	55.93	16.634	1606	55.95	17.634	-.014	-.001	.977
Attendance (5th)	6257	97.46	3.106	2279	96.73	3.999	.723	.217	.000***
In-school Suspensions (5th)	803	(9.7%)	-	311	(9.3%)	-	-	(OR=0.877)	(0.707, 1.089)
Out-of-school Suspensions (5th)	493	(5.9%)	-	188	(5.9%)	-	-	(OR=.947)	(0.765, 1.173)
Intermediate Outcomes									
Stanford Reading (7th)	3765	49.68	17.550	1249	50.99	19.705	-1.317	-.072	.026*
Stanford Math (7th)	3765	57.59	19.316	1249	57.84	20.320	-.250	-.012	.695
Attendance (7th)	5088	96.16	5.345	1810	95.63	5.678	.529	.098	.000***
In-school Suspensions (7th)	558	(6.7%)	-	174	(5.2%)	-	-	(OR=1.219)	(0.948, 1.569)
Out-of-school Suspensions (7th)	287	(3.4%)	-	118	(3.5%)	-	-	(OR=.756)	(0.58, 0.984)
Long-term Outcomes									
College Preparedness									
Attendance (9th)	4775	94.43	8.139	1698	93.28	9.920	1.15	.133	.000***
GPA (9th grade)	2144	2.89	.993	705	2.97	1.029	-.080	-.079	.087
PSAT									
Critical Reading	4063	35.43	8.560	1343	36.24	9.742	-.8143	.091	.004**
Math	4062	38.64	9.210	1342	39.05	9.940	-.4034	.044	.173
SAT Reading	2881	473.87	90.724	899	483.24	99.328	-9.370	-.100	.008**
SAT Math	2881	474.14	90.870	899	484.23	99.890	-10.086	-.108	.005**
ACT Reading	503	20.87	5.824	182	21.84	6.027	-.9758	-.165	.055
ACT Math	503	20.53	4.898	182	21.41	5.214	-.8793	-.172	.042*
School Persistence									
Graduation: Total N Graduates (Data Source: PEIMS Graduation file, 2015–2016 through 2016–2017. The 2017–2018 data were not available at the time of this reporting.)	3027			949					
Graduation (Early: 11 yrs, 2015–2016)	36	(1.2%)	-	14	(1.5%)	-	-	(OR=.806)	(CI=0.433, 1.501)
Graduation (On-time: 12 yrs, 2016–2017)	2991	(98.8%)	-	935	(98.5%)	-	-	(OR=1.003)	(CI=0.904, 1.112)
Graduation (Late: 13 yrs, 2017–2018)	0	-	-	0	-	-	-	-	-
Dropped Out (within 7 years of on-time graduation year: between 2010–2011 and 2016–2017	294	(3.7%)	-	175	(5.2%)			(OR=.401)	(CI=0.155, 1.041)