VOYAGER SOPRIS
LEARNING®

Evaluation of W. T. White
Passport Reading Journeys™ **II Students,**
2006–2007

Katy Denson, Ph.D.
Independent Consultant

Passport Reading Journeys™ is a reading intervention system for adolescent students in the middle grades. The Curriculum Overview states, "*Passport Reading Journeys* provides targeted instruction informed by benchmark assessments to help accelerate struggling students toward reading proficiency" (p. 1). Research-based assessments including the Lexile Framework® for Reading, Reading Benchmarks and Vital Indicators of Progress® (VIP®) Reading Connected Text (RCT) are provided to assist the teacher to group for instruction and to monitor the progress of each student.

In cooperation with the Dallas Independent School District (DISD), Voyager pilot tested the *Passport Reading Journeys* II program with two Reading Improvement teachers at W.T. White High School. Students are placed in Reading Improvement if they (a) are in Grades 7–9, (b) have a scale score below 2200 on TAKS reading, and (c) are proficient in English or classified as bilingual program level "Advanced," "Transitional," or "Post-Transitional." For students without a current TAKS reading score, a current norm-referenced standardized reading test score is honored in determining the Reading Improvement eligibility of students transferring from other districts. Students with a reading comprehension score below the 40th percentile are eligible for Reading Improvement. The Reading Improvement course for Grade 9 students is Reading I.

The purpose of this report is (a) to describe the students participating in the 2006 – 07 Reading Improvement program at W. T. White High School, (b) to describe students' performance on the *Passport Reading Journeys* assessments, (c) to assess the relationship of *Passport Reading Journeys* measures to the standardized achievement tests given by the DISD, including the Iowa Tests of Basic Skills (ITBS), Gates-MacGinite Reading Tests (Gates) and the Texas Assessment of Knowledge and Skills (TAKS), and (d) to assess the growth of Reading Improvement students participating in the *Passport Reading Journeys* program.

METHODOLOGY

DATA MANAGEMENT

A data file of students enrolled in Reading I was obtained from the school. Gates-MacGinite Reading Comprehension scores were supplied by the teachers. *Passport Reading Journeys* assessments, including Lexile, Reading Connected Text (RCT), and Reading Status, were provided. Current year (2007) and previous years (2005 and 2006) TAKS and ITBS scores for participants were obtained from DISD data files.

RELIABILITY AND VALIDITY OF ACHIEVEMENT MEASURES

Iowa Tests of Basic Skills (ITBS)

Reliability is the extent to which a test yields consistent results. Estimates of reliability for Forms K and L of the ITBS were computed with the Kuder-Richardson Formula 20 (KR–20) alternate form reliability coefficient. Reliability for the reading comprehension subtest used in this report was .925. Validity is the extent to which a test measures what it says it is measuring. Content validity is maintained through an on-going research process. Item fairness and differential item functioning were conducted for various subgroups of students[1].

Gates-MacGinitie Reading Test

Estimates of reliability of the Gates were computed with the Kuder-Richardson Formula 20 (KR–20) alternate form reliability coefficient. For the comparison subtests used in this report, reliabilities ranged from .87 to .92 The Gates has an informal degree of content or curricular validity. Construct validity has not been addressed by the test publishers[2].

Texas Assessment of Knowledge and Skills (TAKS)

The KR–20 for TAKS reading assessments ranged from .81 to .93 Both content and criterion-related validity have been established.

Lexile Measures

Reliability for the semantic component was computed by "analyzing more than 50 semantic variables in order to identify those elements that contributed to the difficulty of the 350 vocabulary items on Forms L and M of the *Peabody Picture Vocabulary Test*. The mean log word frequency provided the highest correlation with item rank order (r= –0.779) for the items on the combined form" (p. 4–5)[4]. The reliability of text difficulty was established through a regression equation, yielding the correlation between the observed login difficulties and the theoretical calibrations across the 9 tests of 0.93 after correction for range restriction and measurement error.

Construct validity was well documented through correlations of Lexile measures and 12 separate reading achievement measures including the Iowa Tests of Basic Skills (r= .88), the Stanford Achievement Tests (r= .92), the Metropolitan Achievement Test (r= .93) and the Gates-MacGinite Reading Tests (r= .92). Correlations between theory-based calibrations produced by the Lexile equation and the rank order of unit in basal readers (i.e., grade four is harder than grade 3, the beginning of grade 4 is easier than the end of grade 4) were also high (r= .995).

[1] Impara, J.C. & Plake, B.S. (1998). The Iowa Tests of Basic Skills, Forms K and L. *The Thirteenth Mental Measurements Yearbook*. Buros Institute of Mental Measurements. University of Nebraska-Lincoln. Lincoln, Nebraska.

[2] Cooter, R.B. (1989). Test review: Gates-MacGinite Reading Test, Third Edition. *Journal of Reading*, April 1989, 656–658.

[3] Texas Education Agency (2006). *Technical Digest for the Academic Year 2004–05*. Texas Student Assessment Program, Texas Education Agency.

[4] MetaMetrics, Inc. (2006). Passport Reading Journeys *benchmark assessments: Development and technical guide*.

STATISTICAL ANALYSES

Change in Lexile and RCT score

Repeated measures analyses of variance (ANOVA) were used to measure the change in Lexile and Reading Connected Text (RCT) scores across the three benchmark periods. Students had to have all three measures for each assessment to be included in the analyses. For the change in RCT, students also were coded according to their spring 2006 ITBS reading comprehension percentile score to look for between-subjects effects. Students scoring between the 1st and 25th percentile and the 26th and 39th percentile were included. Differences for both Lexile and RCT were also assessed by ethnicity, English proficiency, and economic status.

Correlations between Passport Reading Journeys *and standardized assessments*

Correlations between the Benchmark One Lexile and the 2006 ITBS Normal Curve Equivalent (NCE) units were computed. Mean NCEs and standard deviations are reported for each disaggregated group, as well as the correlation. For the Gates-MacGinite, correlations were computed between the Benchmark One and Three Lexile and the Gates fall and spring extended scale score. Correlations between TAKS Lexile and *Passport Reading Journeys* Lexile measures were computed for (a) spring 2006 TAKS and Benchmark One and (b) spring 2007 TAKS and Benchmark Three.

Growth on standardized test measures

A repeated measures ANOVA was used to assess growth in reading comprehension using the Gates Extended Scale Score. The spring 2006 ITBS percentile bands, 1st to 25th percentile and 26th to 39th percentile, were used as a between-subjects variable. Also, a repeated measures ANOVA was used to analyze growth on TAKS Reading. Three years of scale scores (2005, 2006 and 2007) were used as the measure. Analyses were also conducted for both measures using ethnicity, English proficiency and economic status as between-subjects variables.

RESULTS

Student Demographic Characteristics

There were 168 students enrolled in nine sections of Reading I (Table 1). Five sections (N = 94) were taught by a teacher who used the program in 2005-06, while the other four (N = 74) were taught by a teacher using *Passport Reading Journeys* for the first time in 2006-07.

There were similar percentages of male (54%) and female (46%) students. Slightly more than half of the students were limited English proficient (LEP) (28%) or exited LEP students (37%). LEP students have a home language other then English reported as the primary language spoken at home and a *Woodcock Muños Language Survey* (WMLS) score of less than 3 (on a scale of 1 – 5). Exited students were previously served in a bilingual education (BE) or English as a Second Language (ESL) program, but exited the BE/ESL program with a WMLS score of 3, 4 or 5 and (a) have passed the *Texas Assessment of Knowledge and Skills* (TAKS) or (b) earned above the 39th percentile on a norm-referenced test of reading comprehension and language.

The majority of students were Hispanic (79%) or African American (14%). Seventy-one percent were classified with low economic status based on free- or reduced-lunch qualification. Thirteen (8%) students served in special education were spread relatively evenly across the nine sections.

TABLE 1: DEMOGRAPHIC CHARACTERISTICS OF *PASSPORT READING JOURNEYS* II STUDENTS AT W.T. WHITE HIGH SCHOOL, 2006–07

Characteristic	Enrolled in 2006–07 N	%	Have Complete 2006–07 Data[a] N	%	Have Complete Longitudinal Data[b] N	%
Gender						
Male	91	54.2	32	47.1	60	50.8
Female	77	45.8	36	52.9	58	49.2
Total	168		68		118	
Ethnicity						
African American	23	13.7	8	11.8	13	11.0
Asian	1	0.6	0	–	1	0.8
Hispanic	132	78.6	57	83.8	100	84.7
White	12	7.1	3	4.4	3	3.4
English Proficiency						
English Proficient	47	28.0	18	30.5	26	22.0
Limited English	59	35.1	11	18.6	34	28.8
Exited	62	36.9	30	50.8	58	49.2
Other Characteristics						
Low Economic	120	71.4	48	70.6	82	69.5
Special	13	7.7	1	1.5	4	3.4

[a]Data include (a) *Passport Reading Journeys* Benchmark One and Three Lexile score,
 (b) Gates-MacGinite fall and spring scores, and (c) 2006 and 2007 TAKS Reading scores.
[b]Data include 2005, 2006 and 2007 TAKS Reading scores.

For some analyses, complete data included (a) *Passport Reading Journeys* Benchmark One and Three Lexile scores, (b) Gates fall and spring scores, and (c) 2006 and 2007 TAKS scores. While most students had a 2006 (88%) and 2007 TAKS score (95%) and *Passport Reading Journeys* Lexile scores (98%), only 43% had Gates fall and spring scores. It appears that one teacher did not give the Gates in the fall and only 77% of the other teacher's students had both a fall and spring Gates assessment. Therefore, only 72% (N = 68) of the total students have complete 2006-07 data (Table 1).

A longitudinal analysis looked at growth on TAKS for students that participated in *Passport Reading Journeys* II in 2006-07. Seventy percent of current students (N = 118) had complete data for this analysis (Table 1).

Standardized Reading Achievement Pretest Scores

The data described below can be thought of as pretest scores because all were collected before students began using *Passport Reading Journeys*. The ITBS and TAKS assessments were conducted in the spring of grade 8 for these students, while the Gates was administered by the teacher in the Reading I classroom.

ITBS READING COMPREHENSION PERCENTILE SCORE

As required for placement in Reading Improvement, all but three students scored below the 40th percentile on the reading comprehension subtest (Table 2). The majority of students (72%) scored below the 26th percentile, and the remaining 20% scored between the 26th and 39th percentiles.

Characteristic	Number	Percent
2006 ITBS Reading Comprehension		
No Score	21	17.8
Below the 26th percentile	70	72.2
26th–39th percentile	24	20.3
40th percentile and above	3	2.5
Fall 2006 Gates-MacGinite Reading Comprehension		
No Score	96	57.1
Below Grade 4	7	4.2
Grade 5 to Grade 6	29	17.3
Grade 7 to Grade 8	14	8.3
Grade 9 and above	5	2.9
2006 TAKS Reading		
Had scorable test	126	88.1
Met minimum expectations	66	44.6

GATES-MACGINITIE READING COMPREHENSION GRADE EQUIVALENT SCORES

The Reading I teachers administered the Gates to students in their classes in fall 2006. Fifty-seven percent of the students had no fall score (Table 2). Of those that had scores, 29 (17% of total students) had a grade equivalent in the grade 5 to grade 6 range. Another 14 (8%) were only slightly below grade level.

TAKS READING SUBTEST

Of the 168 students, 126 had scorable April 2006 TAKS tests (Table 2). Eighteen students were LEP exempt, meaning that they had been in the US less than 30 months, 6 students were absent, and 18 scores were unaccounted for. Of the 126 students, 45% met the minimum requirements for passing the reading subtest.

Passport Reading Journeys Assessments

There are three teacher-administered assessments contained within the *Passport Reading Journeys* program: (a) the Reading Benchmark, including the Lexile measure, (b) the VIP Reading Connected Text (RCT), and (c) Comprehension and Vocabulary Assessments, which are not discussed in this report.

LEXILE FRAMEWORK

MetaMetrics, Inc. developed the *Passport Reading Journeys* Benchmark assessments. Results for the benchmarks are calculated using the Lexile scale and the Lexile Framework for Reading, also developed by MetaMetrics. Materials from the MetaMetrics website (www.Lexile.com) indicate that there is not a direct translation from a specific grade level to the corresponding Lexile measurement. Approximate mid-year reader measures show where students should be reading if they comprehend about 75% of the material. A higher Lexile score for a grade level indicates increased comprehension rather than the ability to read more difficult material. There is considerable overlap of Lexile levels between grades (Table 3). For Grade 9, the recommended reader measure is 855L to 1165L at mid-year.

TABLE 3:
TYPICAL READER AND TEXT
MEASURES BY GRADE

Grade	Reader Measures	Text Measures
1	Up to 300L	200L to 400L
2	140L to 500L	300L to 500L
3	330L to 700L	500L to 700L
4	445L to 810L	650L to 850L
5	565L to 910L	750L to 950L
6	665L to 1000L	850L to 1050L
7	735L to 1065L	950L to 1075L
8	805L to 1100L	1000L to 1100L
9	855L to 1165L	1050L to 1150L
10	905L to 1195L	1100L to 1200L
11–12	940L to 1210L	1100L to 1300L

Note. Data were obtained from the Lexile website.

READING BENCHMARK

The Reading Benchmark is administered three times per year during benchmark periods. The test is administered to the whole class and is a multiple-choice test that takes one class period to complete. Items on the benchmarks are calibrated for reading level using the same equation used to measure books or other text. This allows students to be matched to reading materials on the appropriate Lexile level.

Because of the great overlap in Lexile score bands, it is difficult to assign specific grade levels to students' scores. In Table 4, the Lexile score ranges end at the highest Lexile for each grade. Refer to Table 3 for a more specific grade level range. Regardless of grade level, students clearly made progress from Benchmark One to Three. At Benchmark One, there was only the possibility that 29% of the students fell in the Lexile range for Grade 9 (855L to 1165L), while by Benchmark Three, 72.5% of the students scored in the Grade 9 range.

Lexile Score	First Benchmark		Second Benchmark		Third Benchmark	
	Number	Percent	Number	Percent	Number	Percent
0–300	1	>.1	0	–	0	–
301–500	14	10.1	7	4.2	4	2.4
501–700	50	36.0	28	17.0	9	5.4
701–810	33	23.7	44	26.7	33	19.8
811–910	23	16.5	31	18.8	63	37.7
911–1000	18	12.9	31	18.8	26	15.6
1001–1065	0	–	24	14.5	12	7.2
1066–1100	0	–	0	–	20	12.0
1101–1165	0	–	0	–	0	–
> 1165	0	–	0	–	0	–
Total	139		165		167	

Growth by ITBS pretest percentile band

There were 138 students with Lexile measures for each benchmark assessment. A repeated measures ANOVA found the difference in scores to be significant [$F (2, 136) = 82.837$, $p < .001$, $\eta^2 = .549$]. Students in the 1st to 25th percentile band made steady growth across the three benchmarks, while those in the 26th to 39th band stayed relatively the same from Benchmark Two to Three (Figure 1).
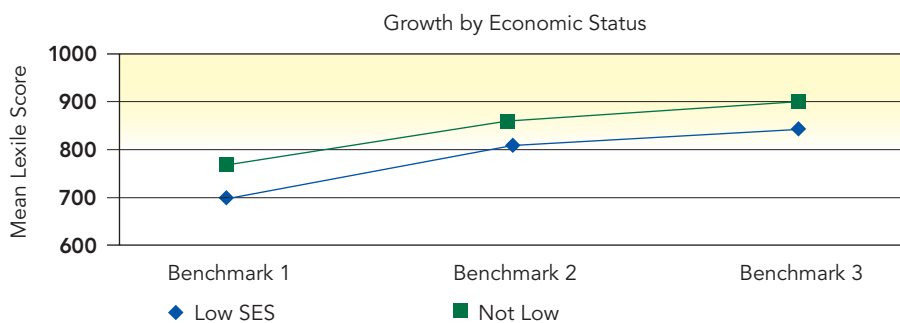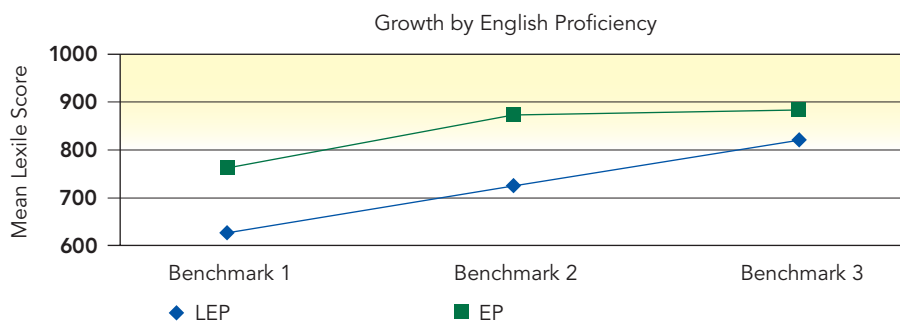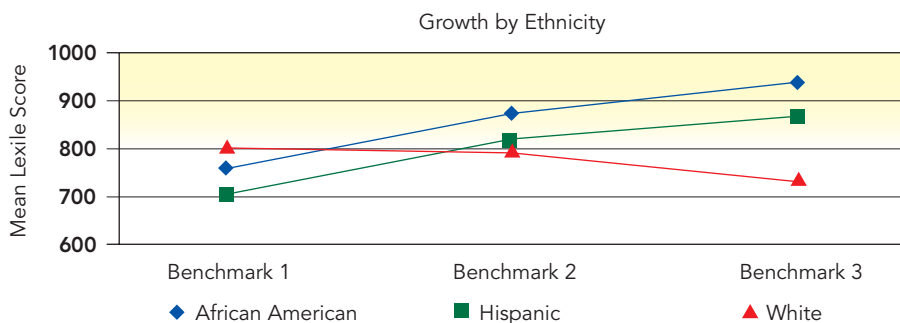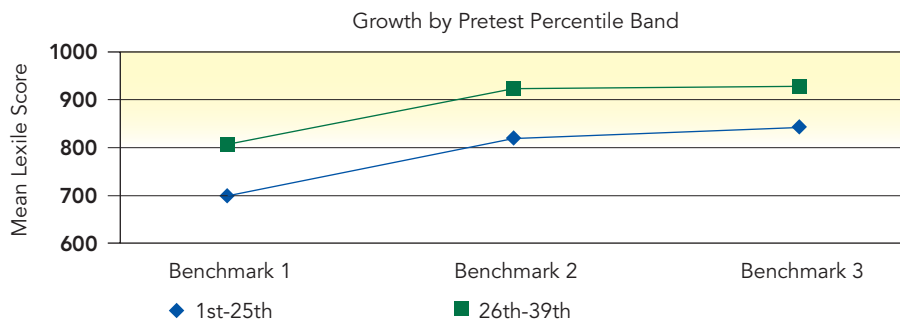
Differences by Ethnicity

There were 138 students with Lexile measures for each benchmark assessment. A repeated measures ANOVA found the difference in scores to be significant [$F (2, 136) = 82.837$, $p < .001$, $\eta^2 = .549$]. Students in the 1st to 25th percentile band made steady growth across the three benchmarks, while those in the 26th to 39th band stayed relatively the same from Benchmark Two to Three (Figure 1).

Differences by LEP status

There was a significant within-subjects interaction in rate of growth on the three Lexile measures [$F (2, 210) = 84.068$, $p < .001$, $\eta^2 = .445$]. Between-subjects effects also was significant [$F (1, 105) = 22.697$, $p = < .000$, $\eta^2 = .178$]. LEP students had a continuous growth from Benchmark One to Three, while English proficient (EP) students had a greater rate of growth from Benchmark One to Two than from Benchmark Two to Three (Figure 1).

VOYAGER SOPRIS
LEARNING®

Growth by Pretest Percentile Band

◆ 1st-25th   ■ 26th-39th

Growth by Ethnicity

◆ African American   ■ Hispanic   ▲ White

Growth by English Proficiency

◆ LEP   ■ EP

Growth by Economic Status

◆ Low SES   ■ Not Low

Differences of economic status

There was a significant within-subjects difference in rate of growth on the three Lexile measures [F (2, 270) = 62.098, p < .001, η^2 = .315], but the interaction was not significant. Between-subjects effects also was significant [F (1, 135) = 4.468, p = .036, η^2 = .032]. Both groups of students had continuous growth from Benchmark One to Three, although low economic status students always had a lower mean Lexile score (Figure 1).

READING CONNECTED TEXT (RCT)

RCT determines a student's fluency rate when reading a grade-level passage. It is administered three times a year at the same time as the Benchmark assessment. It is a one-minute timed measure that is given individually. The *Passport Reading Journeys* materials indicate that RCT is predictive of a student's ability to read on grade level. *Passport Reading Journeys* has goal scores for each benchmark period: (1) 109 words per minute (wpm), (2) 120 wpm, and (3) 125 wpm. Students reading 60 or less wpm are seen as struggling readers.

From the first RCT assessment, 39% (N = 54) of the students read below the goal for Benchmark One (109 wpm) (Table 5). By Benchmark Two, 25% (N = 18) were below the goal of 120 wpm. However, by Benchmark Three, 49% (N = 81) had not reached the goal of 125 wpm, yet only 3 students were below 60 wpm and seen as struggling readers.

Reading Rate	First Benchmark Number	First Benchmark Percent	Second Benchmark Number	Second Benchmark Percent[a]	Third Benchmark Number	Third Benchmark Percent
< 71	7	5.0	3	4.2	10	6.0
71–80	3	2.2	4	5.6	2	1.2
81–90	12	8.6	0	–	8	4.8
91–100	14	10.1	0	–	12	7.2
101–110	18	12.9	5	7.0	28	16.8
111–120	10	7.2	6	8.4	21	12.6
121–130	15	10.8	6	8.4	26	15.6
131–140	22	15.8	14	19.7	17	10.2
141–150	19	13.7	20	28.2	21	12.6
151–160	6	4.3	6	8.4	12	7.2
161–170	5	3.6	4	5.6	5	3.0
171–180	5	3.6	1	1.4	4	2.4
> 180	2	1.4	2	2.8	0	–
Total	139		71		167	

Note. Percents may not add to 100 due to rounding.
[a]Only 71 of the 165 students had an RCT. Percents are based on a total N of 71.

Differences of economic status

There were 41 students with 2006 ITBS scores and RCT scores for three benchmark assessments; 31 in the 1st to 25th percentile and 10 in the 26th to 39th percentile. A repeated measures ANOVA found a significant within-subjects interaction in scores across the three benchmarks [F (2, 78) = 3.702, p = .030, η^2 = .087]. There was a significant quadratic trend in the interaction [F (1, 39) = 6.563, p = .014, η^2 = .144] because there was a drop in mean RCT scores from Benchmark Two to Three for the 1st to 25th percentile band and a rise for the 26th to 39th percentile band. Between-subjects effects were not significant [F (1, 39) = 1.614, p = .211] (Figure 2).

Differences by ethnicity

There was significant within-subjects differences for rate of growth on the three RCT measures [F (2, 110) = 3.312, p = .010, η^2 = .057]. Between-subjects effects was not significant [F (2, 55) = 0.525, p = .595]. Hispanic and White students had growth from Benchmark One to Two, but scores dropped at Benchmark Three. African American students maintained a minimal growth from Benchmark One to Three (Figure 2). Seventy-eight percent of the students with all three scores were Hispanic; therefore, results should be interpreted with caution.
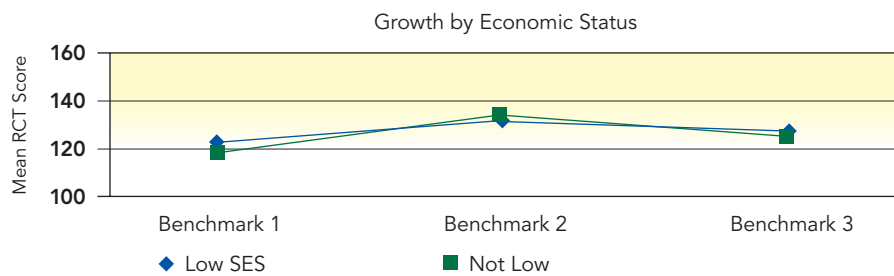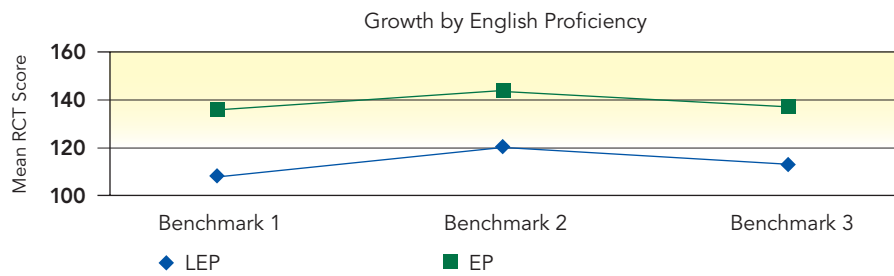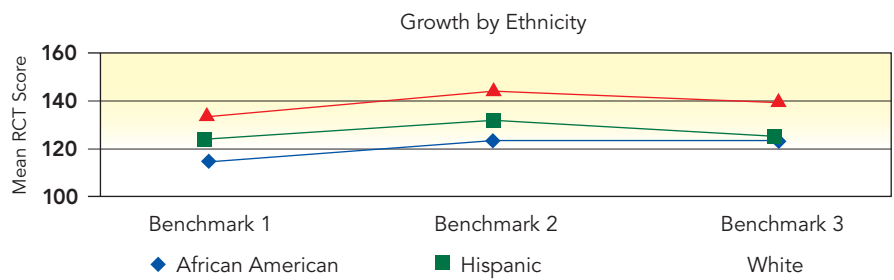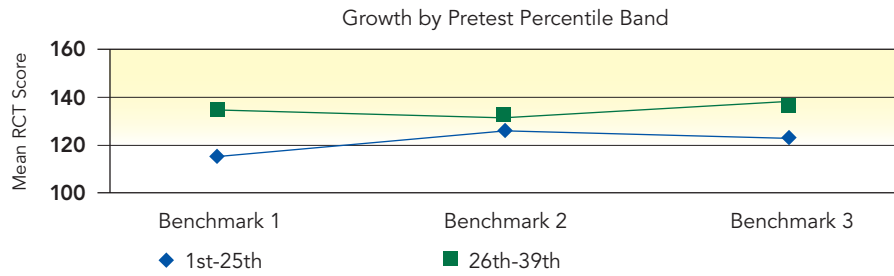
Differences by LEP status

There was a significant within-subjects difference in rate of growth on the three RCT measures [F (2, 86) = 6.936, p = .002, η^2 = .139]. Between-subjects effects also was significant [F (1, 43) = 14.578, p = < .000, η^2 = .253]. Both groups had similar growth patterns, increasing from Benchmark One to Two, but declining from Two to Three (Figure 2). English proficient students maintained higher scores than LEP students at each Benchmark.

Differences by economic status

There was a significant within-subjects difference in rate of growth on the three RCT measures [F (2, 114) = 9.312, p < .001, η^2 = .140], but the interaction was not significant. Between-subjects effects was not significant [F (1, 57) = .050, p = .823]. Scores were almost the same at each Benchmark, regardless of economic status (Figure 2).

FIGURE 2:

GROWTH IN RCT MEASURES
BY PRETEST PERCENTILE
BAND, ETHNICITY, LEP
STATUS AND ECONOMIC
STATUS.

Growth by Pretest Percentile Band

Mean RCT Score

◆ 1st-25th   ■ 26th-39th

Benchmark 1   Benchmark 2   Benchmark 3

Growth by Ethnicity

Mean RCT Score

◆ African American   ■ Hispanic   White

Benchmark 1   Benchmark 2   Benchmark 3

Growth by English Proficiency

Mean RCT Score

◆ LEP   ■ EP

Benchmark 1   Benchmark 2   Benchmark 3

Growth by Economic Status

Mean RCT Score

◆ Low SES   ■ Not Low

Benchmark 1   Benchmark 2   Benchmark 3

Relationship of Standardized Assessments to *Passport Reading Journeys* Assessments

ITBS READING COMPREHENSION

The correlation between the Lexile level from Benchmark One and the 2006 ITBS Reading Comprehension NCE scores was statistically significant [r (100) = .257, p = .01]. No correlations were computed for the Benchmark Three Lexile and ITBS because Grade 9 students did not take ITBS in 2006 – 07.

GATES-MACGINITE READING COMPREHENSION

The correlation between the Benchmark One Lexile and the fall Gates was statistically significant [r (60) = .625, p < .001]. In the spring, there also was a significant correlation between the Benchmark Three Lexile and the Gates scale [r (163) = .475, p < .001].

TAKS READING

There was a significant correlation between the spring 2006 TAKS Lexile and the fall *Passport Reading Journeys* Benchmark One Lexile [r (124) = .391, p < .001]. The Lexile assigned by the TAKS was higher than the Lexile from the *Passport Reading Journeys* Benchmark. There was a significant correlation between the 2007 TAKS Lexile and the *Passport Reading Journeys* Benchmark Three Lexile [r (159) = .442, p < .001].

Growth on Standardized Assessments

GATES-MACGINITE READING COMPREHENSION

Differences by ITBS pretest percentile band

A repeated measures ANOVA found a significant growth in scale score on the Gates [F (1, 59) = 65.511, p < .001, η^2 = .526]. Rate of growth was similar for students scoring between the 1st to 25th and 26th to 39th percentile on the spring 2006 ITBS (Figure 3), although students pretesting in the 26th to 39th percentile band had higher mean scores for both assessments.
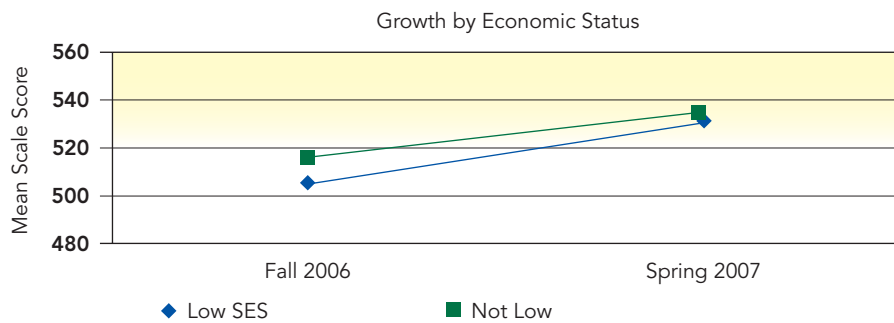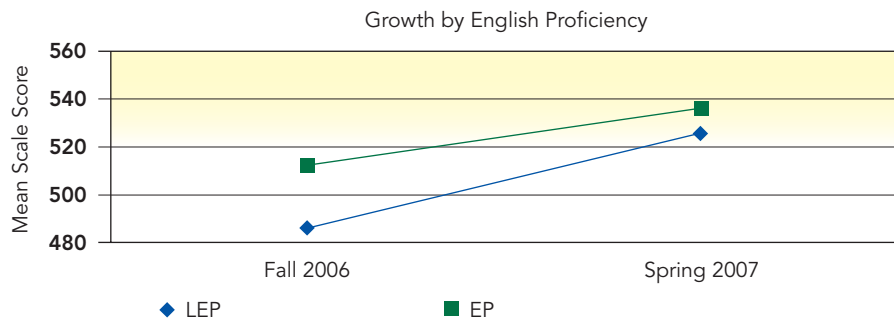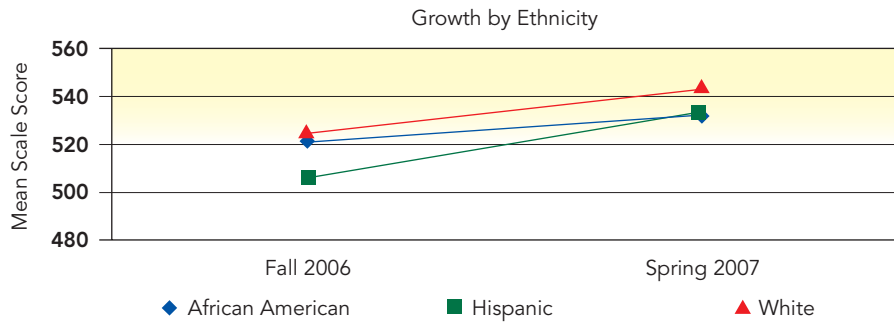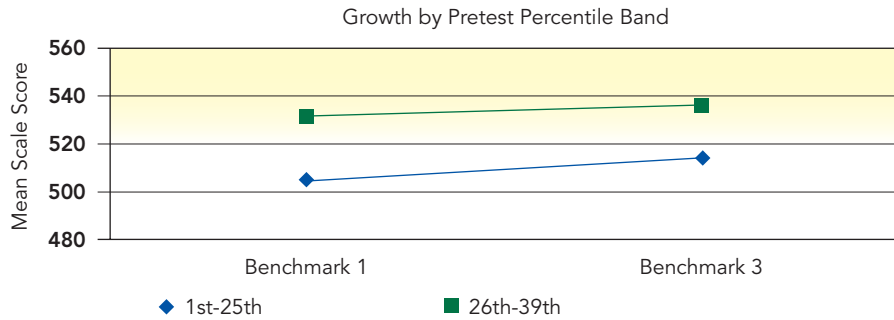
Differences by ethnicity

There was a significant within-subjects difference in rate of growth from fall to spring [F (1, 69) = 14.442, p < .001, η^2 = .173]. Between-subjects effects was not significant [F (2, 69) = 1.577, p = .214]. The mean scale score for Hispanic students was 20 points lower than African American students in the fall, but was the same in the spring (Figure 3). There were only four White students in this analysis, so sample size may have played a role in the results.

Differences by LEP status

There was a significant within-subjects interaction in rate of growth on the Gates [F (1, 50) = 4.396, p = .041, η^2 = .081]. Between-subjects effects also was significant [F (1, 50) = 10.722, p = .002, η^2 = .178]. Both EP and LEP students had growth from fall to spring. However, the rate of growth was much higher for LEP students, gaining 41 points (Figure 3).

VOYAGER SOPRIS
L E A R N I N G®

FIGURE 3:

GROWTH ON THE GATES-
MACGINITE BY PRETEST
PERCENTILE BAND,
ETHNICITY, LEP STATUS AND
ECONOMIC STATUS

Growth by Pretest Percentile Band

Mean Scale Score

	Benchmark 1	Benchmark 3

◆ 1st-25th   ■ 26th-39th

Growth by Ethnicity

Mean Scale Score

	Fall 2006	Spring 2007

◆ African American   ■ Hispanic   ▲ White

Growth by English Proficiency

Mean Scale Score

	Fall 2006	Spring 2007

◆ LEP   ■ EP

Growth by Economic Status

Mean Scale Score

	Fall 2006	Spring 2007

◆ Low SES   ■ Not Low

TAKS READING

A repeated measures ANOVA found a significantly different rate of growth in TAKS scale score from 2005 to 2006 and 2006 to 2007 [F (2, 196) = 26.589, p < .001, η^2 = .213]. There was a significant quadratic trend [F (1, 98) = 9.779, p < .001, η^2 = .091] because mean growth from Grade 7 to Grade 8 (2005 to 2006) was only 6.5 points. However, from Grade 8 to Grade 9, after participating in a Reading Improvement course using *Passport Reading Journeys*, mean growth was 93 points (Figure 4).

FIGURE 4:
MEAN TAKS READING SCALE SCORES FROM 2005 TO 2007
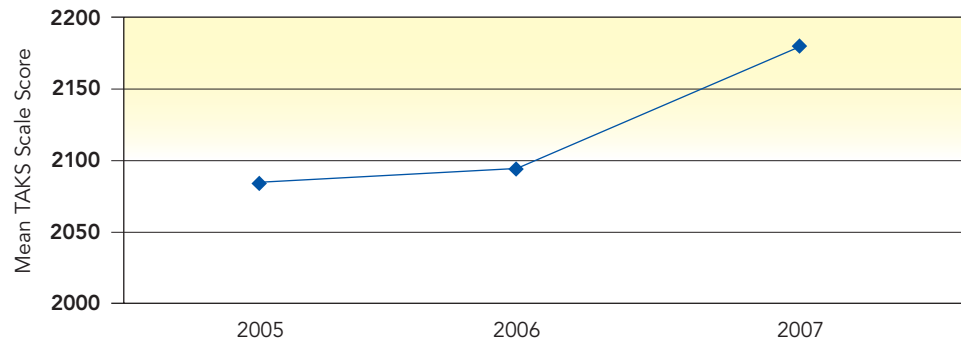
Differences by ITBS pretest percentile band

Another repeated measures ANOVA found a significant difference in the rates of growth for students of differing spring ITBS percentile bands [F (1, 85) = 5.625, p = .020, η^2 = .062]. Although students in the 26th to 39th percentile band had higher TAKS scale scores all three years, they did not make the pronounced growth in Grade 9 of those that scored in the 1st to 25th percentile band in spring 2006 (Figure 5).

Differences by ethnicity

There was a significant within-subjects difference in rate of growth from 2005 to 2007 [F (2, 192) = 3.67, p = .027, η^2 = .037]. Between-subjects effects were not significant [F (2, 96) = 1.256, p = .290]. African American students had the highest mean scale scores each year (Figure 5). Hispanic students made little progress from Grade 7 to Grade 8, but made the greatest amount of growth in Grade 9. There were only four White students in this analysis, so sample size is an issue.
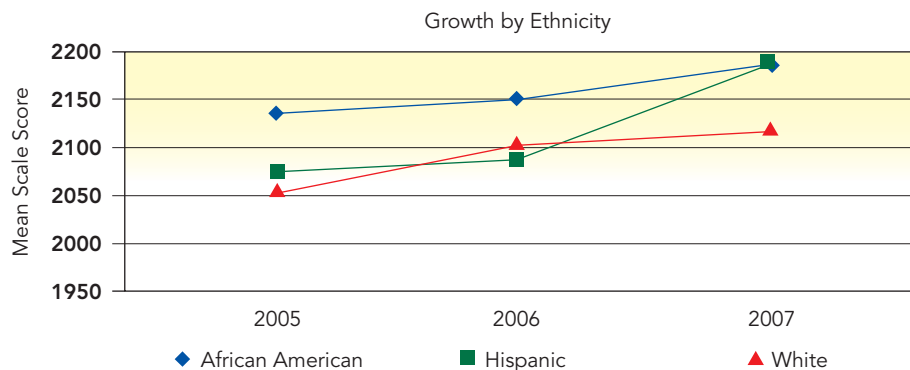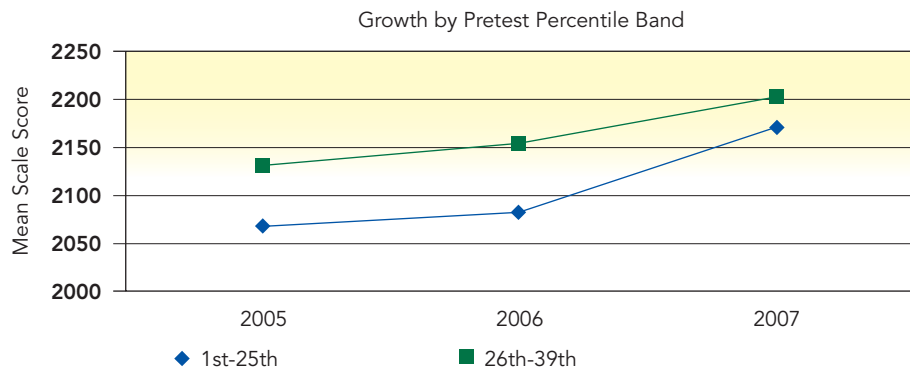
Differences by LEP status

There were statistically significant within-subjects [F (2, 186) = 17.15, p < .001, η^2 = .156] and between-subjects differences for rate of growth on TAKS Reading [F (1, 93) = 82.05, p < .001, η^2 = .469]. Both EP and LEP students had growth from 2005 to 2007. However, the rate of growth was much higher for LEP students, although they remained behind EP students by 237 points (Figure 5).
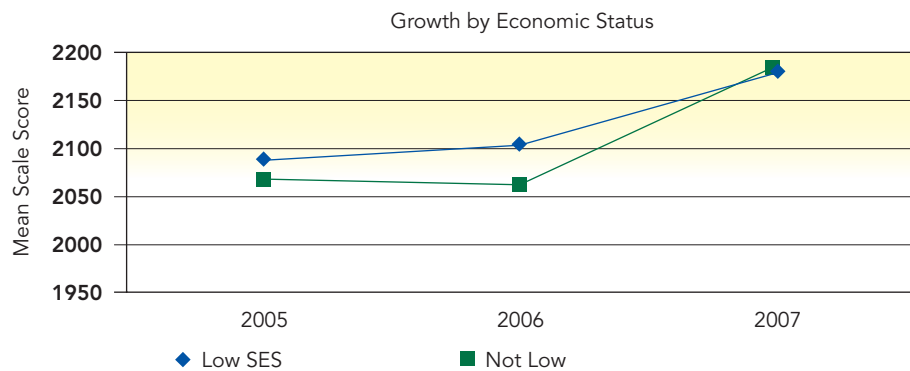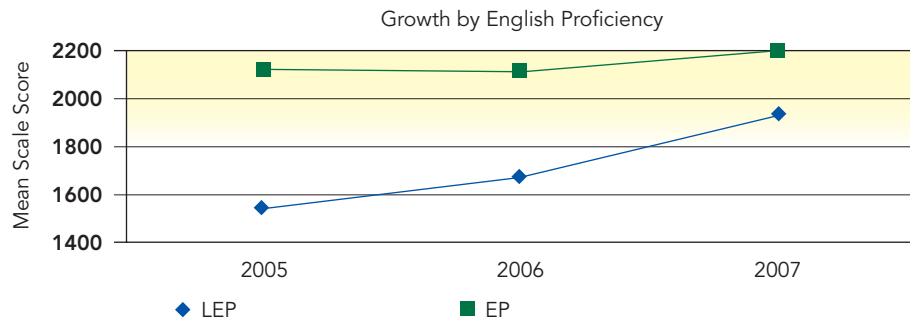
Differences by economic status

There was a significant within-subjects difference in TAKS scale score growth [$F_{(2, 194)}$ = 27.556, $p < .001$, $\eta^2 = .221$], but the between-subjects effects was not significant [$F_{(1, 97)}$ = 0.741, $p = .391$]. Students that did not have a low economic status had a slightly lower mean in 2006 than 2005, but their mean in 2007 was slightly higher than the low economic status students (Figure 5). Low economic status students made a slight gain from 2005 to 2006, but had a very large gain in 2007 after using *Passport Reading Journeys.*

FIGURE 5:

GROWTH ON TAKS READING SCALE SCORE BY PRETEST PERCENTILE BAND, ETHNICITY, LEP STATUS AND ECONOMIC STATUS

Growth by Pretest Percentile Band

Mean Scale Score

◆ 1st-25th   ■ 26th-39th

Growth by Ethnicity

Mean Scale Score

◆ African American   ■ Hispanic   ▲ White

FIGURE 5:
CONTINUED

Growth by English Proficiency

Mean Scale Score

2200
2000
1800
1600
1400

2005   2006   2007

◆ LEP   ■ EP

Growth by Economic Status

Mean Scale Score

2200
2150
2100
2050
2000
1950

2005   2006   2007

◆ Low SES   ■ Not Low

SUMMARY AND CONCLUSIONS

Passport Reading Journeys is a reading intervention system for adolescent students in the middle grades. Research-based assessments including the Lexile Framework for Reading, Reading Benchmarks and Vital Indicators of Progress (VIP) Reading Connected Text (RCT) are provided to assist the teacher to group for instruction and to monitor the progress of each student. In cooperation with the Dallas Independent School District (DISD), Voyager pilot tested the *Passport Reading Journeys* II program with two Reading Improvement teachers at W.T. White High School. Primarily, Reading Improvement is for students in Grades 7-9 that have a scale score below 2200 on TAKS reading.

A variety of data including *Passport Reading Journeys* assessments, Lexile, Reading Benchmark, Reading Connected Text (RCT), and Reading Status, and standardized test scores, TAKS ITBS, and Gates-MacGinite reading scores, were used to assess growth in reading for these DISD students. Repeated measures ANOVAs were generally used to assess growth.

Student Demographics

There were 168 students enrolled in nine sections of Reading Improvement. Five sections
(N = 94) were taught by a teacher who worked with the program in 2005-06, while the other four
(N = 74) were taught by a teacher using *Passport Reading Journeys* for the first time in
2006-07. Slightly more than half of the students were limited English proficient (LEP)
(28%) or exited LEP students (37%). The majority of students was Hispanic (79%) or
African American (14%) and had a low economic status based on free- or reduced-lunch
qualification (71%).

From spring 2006 ITBS assessment, most students scored below the 40th percentile in reading.
The Gates, administered in fall 2006, confirmed that most students read below grade level. Forty
percent had a grade equivalent in the grade 5 to grade 6 range. Only 45% had passed the spring
2006 TAKS reading subtest.

Passport Reading Journeys Assessments

Benchmark assessments are given three times per year. Reported results include Lexile, Reading
Connected Text and Reading Status.

LEXILES

At Benchmark One, there was only the possibility that 29% of the students fell in the Lexile
range for grade 9 (855L to 1165L), while by Benchmark Three, 72.5% of the students scored in
the grade 9 range. The rise in Lexile score was not as steep from the second to third benchmark as
it was from the first to second, however the amount of growth was statistically significant.

African American and Hispanic students had continuous growth from Benchmark One to Three,
while White students had similar Lexile scores at Benchmarks One and Two, but dropped in
score at Benchmark Three. LEP students had a continuous growth from Benchmark One to
Three, while English proficient (EP) students had a greater rate of growth from Benchmark One
to Two than from Benchmark Two to Three. Both low and not low economic status students
had continuous growth from Benchmarks One to Three, although low economic status students
always had a lower mean Lexile score.

READING CONNECTED TEXT

RCT determines a student's fluency rate when reading a grade-level passage that is predictive of a student's ability to read on grade level. From the first RCT assessment, 39% of the students read below the goal for Benchmark One (109 wpm). By Benchmark Two, 25% were below the goal of 120 wpm. However, by Benchmark Three, 49% had not reached the goal of 125 wpm.

Students pretesting in the 1st to 25th percentile band experienced a drop in mean RCT scores from Benchmark Two to Three, while those pretesting in the 26th to 39th percentile band had a continuous gain. Hispanic and White students had growth from Benchmark One to Two, but scores dropped at Benchmark Three. African American students maintained a minimal growth from Benchmark One to Three. Results by ethnicity should be interpreted with caution because 78% of the sample was Hispanic. Both English proficient and LEP students had similar growth patterns, increasing from Benchmark One to Two, but declining from Two to Three. English proficient students maintained higher scores than LEP students at each Benchmark. Scores were almost the same at each Benchmark, regardless of economic status.

Relationship of Standardized Assessments to *Passport Reading Journeys* Assessments

Correlations revealed that Benchmark Three assessments (Lexile and RCT) had stronger correlations with spring 2007 standardized assessments, such as TAKS and Gates, than did Benchmark One assessments. TAKS-assigned Lexiles were consistently higher than *Passport Reading Journeys* Lexile scores.

Growth on Standardized Assessments

GATES-MACGINITE READING COMPREHENSION

Rate of growth for Gates scale scores was similar for students scoring between the 1st to 25th and 26th to 39th percentile on the spring 2006 ITBS, although students pretesting in the 26th to 39th percentile band had higher mean scores for both fall and spring assessments.

The mean scale score for Hispanic students was 20 points lower than African American students in the fall, but was the same in the spring. Few White students had scores included in the analysis. Both EP and LEP students had growth from fall to spring. However, the rate of growth was much higher for LEP students. Both low and not low economic status students had growth from fall to spring, when scale score means were almost the same.

TAKS READING

Growth on TAKS Reading was assessed over three years: from 2005 to 2006 when students were in middle school and from 2006 to 2007 when students were using *Passport Reading Journeys* in Reading Improvement. Although students in the spring ITBS 26th to 39th percentile band had higher TAKS scale scores all three years, they did not make the pronounced growth in grade 9 of those that scored in the 1st to 25th percentile band in
spring 2006.

Hispanic students made little progress from grade 7 to grade 8, but made the greatest amount of growth in grade 9. Both EP and LEP students had growth from 2005 to 2007. However, the rate of growth was much higher for LEP students, although they remained behind EP students. Students that did not have a low economic status had a slightly lower mean in 2006 than 2005, but their mean in 2007 was slightly higher than the low economic status students. Low economic status students made a slight gain from 2005 to 2006, but had a significant gain in 2007 after using *Passport Reading Journeys*.

There are many results in this study that support the claim that using the *Passport Reading Journeys* program assists students in gaining reading proficiency. *Passport Reading Journeys* assessments showed continuous growth, while remaining consistent with standardized achievement measures. For this group of predominantly Hispanic, low income students, many with current or previous English language deficiencies, and the majority with well-defined reading inadequacies, increased reading proficiency was clearly demonstrated.

Students pretesting in the lowest percentile ranges (1st to 25th percentiles) on ITBS reading comprehension made significant gains in both *Passport Reading Journeys* and TAKS Lexile levels. However, after an initial growth in fluency, as measured by the RCT Benchmark, these students experienced a loss. It may be that the material was increasingly difficult, and it remained above their reading level, or the assessments may need to be more finely tuned to Lexile measurements. Students pretesting in the 26th to 39th percentiles also made gains, but they were not as marked as the lower scoring students.

All students made growth on the *Passport Reading Journeys* and TAKS Lexiles, Gates, and TAKS Reading scale score. When disaggregated by ethnicity, Hispanic students made particularly notable progress. There were few White students in the dataset, so their results should be interpreted cautiously. On the *Passport Reading Journeys* Lexile, Hispanic and African American students had similar growth patterns, with African American students having a consistently higher mean Lexile score. Growth on the Gates from fall to spring was greater for Hispanic students than African American. Three years of TAKS Reading data showed that Hispanic students particularly benefited from *Passport Reading Journeys* participation. Fluency measures dropped for Hispanic and African American students at Benchmark Three, perhaps for reasons mentioned above.

Limited English proficient students had lower mean starting scores than English proficient students for all measures. Yet, their growth was significant, particularly for TAKS Reading scale scores, which showed a large gain for the current year. Although their mean scale score was still at the level requiring Reading Improvement, it was well above the score required for passing TAKS (scale score of 2100).

There were fewer differences by economic status. Students that were low income or not both made significant gains on *Passport Reading Journeys* and TAKS Lexiles, Gates, and TAKS Reading scale score. Students that were low economic status had mean scores slightly below those that were not on all measures.

As growth was so evident on all standardized measures, it may be worth investigating more fully the Benchmark Three RCT fluency measure. Most student groups, regardless of disaggregation variable, experienced a decline in fluency at this Benchmark.

Because there is no control group for this study, the three-year TAKS Reading scale score analysis is particularly crucial. A repeated measures ANOVA provides a control, in that students become their own control. The change from 2005 to 2006, when students may or may not have been in a grade 8 Reading Improvement course, is contrasted with the change from 2006 to 2007, when students were enrolled in Reading Improvement and used the *Passport Reading Journeys* program. Additionally, variations among subjects due to individual differences, such as gender, ethnicity, etc., are completely removed from the analysis because every subject is compared only to him- or herself. This analysis revealed that the current year saw significant growth, particularly for those students for whom there is the most concern poor readers (pretesting below the 26th percentile), Hispanic, LEP, and low economic status.

VOYAGER SOPRIS
LEARNING®