HOUGHTON MIFFLIN HARCOURT

JOURNEYS

COMMON CORE

A STUDY ON THE EFFECTIVENESS OF *JOURNEYS* IN TEXAS USING EXISTING DATA SOURCES

Prepared By:

Miriam Resendez, M.A., Senior Researcher, PRES Associates
Mariam Azin, Ph.D., President, PRES Associates
March, 2012

HOUGHTON MIFFLIN HARCOURT

EXECUTIVE SUMMARY

The **Houghton Mifflin Harcourt *Journeys*** program is a research–based comprehensive K–6 reading/language arts program that targets key elements of literacy including reading comprehension, vocabulary, fluency, grammar, writing, and at grades K–2, phonics and phonemic awareness. The ***Journeys*** program has been adopted in over 700 schools/districts in the state of Texas as of the 2010–11 school year. for implementation in grades K-5 PRES Associates, an external, independent educational research firm with over twenty years of experience in applied educational research and evaluation, conducted Year 1 (2012) analyses for a multi–year quasi–experimental study on the effectiveness of the ***Journeys*** program in the state of Texas. This study looks at performance trends on Texas state assessment data among schools using ***Journeys***, as well as matched control schools using other literacy programs.

It should be noted that the findings contained herein should only be considered preliminary, given that schools have only recently adopted the ***Journeys*** program (as of Fall 2010). It takes time for teachers and schools to fully implement any newly adopted curriculum and for performance trends to subsequently manifest themselves in terms of state assessment data. Analyses from the subsequent 2011–2012 school year and onward will be more telling and sensitive to examining the longer term effects of the program.

With this caveat in mind, the purpose of this report is to present the interim Year 1 results of statistical analyses conducted on existing Texas state assessment data in order to evaluate the effectiveness of the ***Journeys*** program in helping Texas elementary school students attain vital literacy skills. Major findings arranged by evaluation questions include the following:

Are there significant changes in the reading performance of students who use *Journeys* over time?

Results showed significant growth over time after schools began using ***Journeys***—from Spring 2010 (pre) to Spring 2011 (post)—as measured by the TAKS reading scale score. Learning gains experienced by ***Journeys*** students were converted to percentile ranks; results showed a gain of 17 percentiles on the reading scale score. Furthermore, while there was relatively no change in the percentage of ***Journeys*** students meeting the state reading performance standard, there were increases in the percentage of ***Journeys*** students who showed commended performance.

Do growth patterns vary before and after schools began using *Journeys*?

Learning gains on the TAKS reading scale score are supported in part by the analysis of all students at each grade level and their performance before and after ***Journeys*** was introduced (i.e., crosssectional analysis). Among 4th and 5th graders, the highest reading score was observed in 2011, following ***Journeys*** usage, as compared to prior testing years. In contrast, 3rd graders showed the highest level of reading performance in 2010 (before ***Journeys***) as compared to 2011 (following ***Journeys***). With respect to writing, 4th grade students' performance on the writing assessment steadily declined over the three testing years (from 2009 to 2011). That said, the decline was smaller following ***Journeys*** usage than prior to exposure to this program. In sum, there is some inconsistency in the results obtained by grade level. Comparisons of growth rates among the same set of students (within subjects analysis) were also conducted. In particular, researchers compared growth from 3rd (2009) to 4th (2010) grade to growth from 4th (2010) to 5th (2011) grade. In the latter case, ***Journeys*** was introduced. Results showed that students experienced greater learning gains after their schools began implementation of the ***Journeys*** program as compared to before.

Is *Journeys* associated with improvements for various subgroups?

The following types of students showed significant changes in their reading performance following introduction of the ***Journeys*** program: Limited English Proficiency (LEP), Whites, and African Americans. Among Whites and African Americans, accelerated learning gains were observed following usage of ***Journeys***. A similar pattern was observed among Hispanics; however results were only marginally significant. In contrast, among LEP students, students showed greater learning gains prior to ***Journeys*** as compared to after ***Journeys***.

How does student achievement in reading on Texas state assessments differ across users and non-users of *Journeys*?

This set of analyses provides information on the relationship between ***Journeys*** and reading performance relative to schools that did not use ***Journeys***. Results from these analyses showed no significant differences between ***Journeys*** and non– ***Journeys*** students on the TAKS reading and writing scale score from both the longitudinal analysis (examining changes within students over time) and cross sectional analysis (examining performance at each grade level over time).

Do such findings vary across different subgroups of students?

Exploratory information indicates that females, Whites, Hispanics, African Americans, Limited English Proficient, and economically disadvantaged ***Journeys*** students had similar growth rates as compared to non– ***Journeys*** students who were in these subpopulations.

While the present study showed no overall effects of the ***Journeys*** program in Texas as compared to students using other literacy programs, it should be emphasized that students in ***Journeys*** schools had only been using the program for 7–8 months (testing occurs in April). Given that there is a learning curve when any new curriculum is introduced, it is not surprising that significant effects have not been obtained during this first year of analyses. Moreover, small effects are to be expected in any type of study evaluating entire curricula against one another; after all, literacy programs teach similar concepts (e.g., phonics, comprehension, etc..) Such overlap between curricula will reduce the likelihood of detecting effects. That said, as additional data is collected over the next two years (2012 and 2013), researchers will be able to look at more long-term effects of the program within Texas. Moreover, a national experimental study is being conducted (2011–12 and 2012–13 school years) which is designed to examine the efficacy of this program within more controlled settings.

In summary, the results of this study using state assessment data provide some preliminary support for a positive relationship between the ***Journeys*** program and elementary reading performance. However, more conclusive findings on the overall impact of the ***Journeys*** program will be obtained as additional data is collected.

TABLE OF CONTENTS

TABLE OF FIGURES AND TABLES

Figures

Tables

PROJECT BACKGROUND

> ***Literacy is the ability to identify, understand, interpret, create, communicate and compute using printed and written materials associated with varying contexts. Literacy involves a continuum of learning to enable an individual to achieve his or her goals, to develop his or her knowledge and potential, and to participate fully in the wider society.***
>
> United Nations Educational, Scientific and Cultural Organization (UNESCO)

According to the most recent National Assessment of Education Progress (2011), U.S. students continue to struggle with reading skills. Approximately 33% of our nation's 4th graders scored at levels at or above proficiency in 2011; this percentage remained relatively unchanged from 2009 when 34% of students were proficient. This is problematic as students who hope to succeed in future educational pursuits and career endeavors must have a strong reading foundation regardless of where their educational and career goals are headed. As so aptly stated by the American Federation of Teachers (2009), "reading is the fundamental skill upon which all formal education depends. Research clearly shows that children who are poor readers at the end of first grade are never likely to acquire the reading skills they need to successfully complete elementary school…any child who doesn't learn to read early and well will not easily master other skills and knowledge, and is unlikely to ever flourish in school or in life."

To help address the large gap in elementary students' literacy skills, **Houghton Mifflin Harcourt** developed a new reading/language arts program that has shown promise as an effective instructional program for elementary school students. The 2012 ***Journeys*** program is a research–based comprehensive K–5 literacy program that targets key elements of literacy including reading comprehension, vocabulary, fluency, grammar, writing, and at grades K–2, phonics and phonemic awareness. Designed to meet the diverse needs of all students, every lesson allows the student to develop comprehension and fluency by focusing on a target skill and target strategy in a relevant short story and non–fiction story companion.

Given how important literacy skills are to the future success of children, programs that can help in the development of these skills need to be looked at carefully to determine the extent to which they help students attain critical reading and writing skills. Planning, Research, and Evaluation Services (PRES Associates, Inc.)[1], conducted Year 1 (2012) analyses for a multi–year quasi–experimental study to examine the effectiveness of the ***Journeys*** program in helping elementary students in Texas improve their literacy skills and understanding. The quasi–experimental study will extend to the 2012–2013 school year as additional longitudinal data is obtained to examine the long term effects of the program. In the interim, the following report presents results for the first year of analyses.

1 PRES Associates is an external, independent, educational research firm with over 20 years of experience in applied educational research and evaluation.

PROJECT OVERVIEW

The purpose of this report is to present the results of statistical analyses conducted on existing Texas state assessment data in order to determine whether the ***Journeys*** program helps elementary school students attain important literacy skills. Specifically, the analyses are designed to address the following key evaluation questions:

1. Are there significant changes in the reading performance of students who use ***Journeys*** over time? Do growth patterns vary before and after schools began using ***Journeys***?
2. Is ***Journeys*** associated with improvements for various subpopulations of students?
3. How does student achievement in reading on Texas state assessments differ across users and non–users of ***Journeys***? Do such findings vary across different subgroups of students?

DESIGN AND **METHODOLOGY**

Research Design

A quasi–experimental design was used to evaluate the ***Journeys*** program among Texas elementary school students in Grades 3–5. To address these evaluation questions, PRES Associates gathered assessment data available from the Texas Education Agency (TEA) as a result of the No Child Left Behind Act (NCLB) of 2001, which requires measurement of school performance towards adequate yearly progress (AYP). As part of this legislation, states are required to administer reading assessments to students in Grades 3–8 (and during high school) and make school results available to the public.

Data for students from schools using ***Journeys*** and matched comparison schools were requested from the TEA. A detailed description of the measures and samples used follows.

Measures

The TAKS (Texas Assessment of Knowledge and Skills) is the primary state assessment of academic skills used from Spring 2003 to Spring 2011. TAKS is designed to measure core areas of the statemandated curriculum, the Texas Essential Knowledge and Skills (TEKS). The TAKS scale score became vertically scaled as of the Spring 2009 testing. This means that any TAKS scale scores from 2009 to 2011 can be compared as long as the comparison is for the same subject area. For example, vertical scale scores in reading can be compared each year, from Grades 3–5, but vertical scale scores in reading cannot be compared to vertical scale scores in mathematics. Since scores over recent years are comparable, data from Spring 2009 and 2010 testing serve as baseline (prior to ***Journeys***), and data from Spring 2011 serves as the first treatment year (***Journeys*** was first implemented in the 2010–11 school year).

Testing in reading occurs annually for all students in Grades 3–8. The writing assessment, however, is only administered at Grades 4 and 7. The following tables show the number items measured by objective for both reading and writing.

TABLE 1. TAKS BLUEPRINT FOR READING

Selections may be narratives, expository pieces, or mixed pieces, which combine two types of writing.	G.3	G.4	G.5
Objective 1—Basic understanding	15	15	13
Objective 2—Literary elements	7	8	8
Objective 3—Analysis using reading strategies	6	7	8
Objective 4—Analysis using critical thinking skills	8	10	13
Total number of items	**36**	**40**	**42**

TABLE 2. TAKS BLUEPRINT FOR WRITING: 4TH GRADE

Objective	Grade 4
Objective 1–Appropriate organization of ideas	4 multiple–choice items
Objective 2–Correct and effective sentence construction	8 multiple–choice items
Objective 3–Standard usage and appropriate word choice	8 multiple–choice items
Objective 4–Proofreading for punctuation, capitalization, spelling	8 multiple–choice items
Total number of items	**28 multiple-choice items**

The analyses described herein used the following outcome measures:

- 3rd–5th grade TAKS reading scale scores;
- 3rd–5th grade reading objectives (percent correct in each objective);
- 3rd–5th grade proficiency levels (percent meeting standards and percent who achieved commended performance); and
- 4th grade writing scale scores.

Samples

Elementary schools using the ***Journeys*** program in the 2010–11 school year were selected for inclusion in this study (n=55)[2]. Control sites[3] (n=50) were selected based on propensity scoring and matching methods.

This is described in more detail in Appendix A. This procedure matched each ***Journeys*** school with the closest non–***Journeys*** site based on the following school–level characteristics:

- Enrollment
- Percent Economically Disadvantaged
- Percent Limited English Proficient
- Percent Special Ed
- Percent Gifted
- Percent At–Risk (504)
- Mobility Rate
- Percent White
- Percent Hispanic
- Percent Black
- Percent Native American
- Percent Asian/Pacific Islander

As shown in Appendix A, the propensity matching procedure resulted in schools that were very similar with respect to the aforementioned demographic characteristics. Indeed, there were no significant differences observed between schools based on 2009–10 statistics.

Two cohorts of data were available. Cohort A and B had two and three years of longitudinal data available to examine student level change at the elementary level (i.e., Grades 3–5). Cohort A consists of elementary students who were in the 4th grade in the 2008–09 school year. Data was also obtained for these same students when they were in the 5th grade (2009–2010). Cohort B consists of 3rd grade students in the 2008–2009 school year. Data for these same students were also obtained when they were in the 4th grade (2009–2010) and 5th grade (2010–2011). Tables 3–4 display the data collection timeline for these samples and total sample sizes. As previously noted, analyses on the comparability of the ***Journeys*** and control sites at the school level showed no significant differences among 11 any of the measured demographic characteristics, p>.05.

2 Note that only schools confirmed to be ***Journeys*** users through contact with the school by an independent call center were included in this study. These schools had to have used the program in 75% or more of their classes.

3 Similarly, only schools confirmed to be non–***Journeys*** users by an independent call center were included in this study.

TABLE 3. TAKS-SAMPLE A (SAMPLE SIZE)

Group	Grade	08-09	09-10	10-11
***Journeys* Schools=52**	4	A (2444)		
	5		A (3230)	
Non-*Journeys* Schools=48	4	A (2239)		
	5		A (2857)	

TABLE 4. TAKS-SAMPLE B (SAMPLE SIZE)

Group	Grade	08-09	09-10	10-11
***Journeys* Schools=52**	3	B (2309)		
	4		B (2747)	
	5			B (3366)
Non-*Journeys* Schools=48	3	B (1833)		
	4		B (2193)	
	5			B (2797)

In addition to cohorts of students who can be followed over time, data was also collected from all students in Grades 3–5 at each of the participating schools for each school year from 2008–09 to 2010–11. The table below shows the counts for this sample available for cross–sectional analyses.

TABLE 5. TAKS-SAMPLE C (SAMPLE SIZE)

Group	Grade	08-09	09-10	10-11
***Journeys* Schools=55**	3	3418	3462	3487
	4	3053	3375	3330
	5	3033	3230	3366
Non-*Journeys* Schools=50	3	3046	3033	2959
	4	2893	B (2193)	
	5			B (2797)

Table 6 shows the number of students within each of the subpopulations for each study sample. For illustrative purposes, only the data for one grade level at a single school year is presented. As shown, while there are sufficient students within the vast majority of subpopulations, there are small numbers of students who are reported as "Other Race/Ethnicity" as well as Special Education. As a result, subpopulation analyses of these subgroups were not conducted.

TABLE 6. STUDENT LEVEL COUNTS BY SUBGROUP

	Group	Sample A 2010 5th Grade	Sample B 2011 5th Grade	Sample C 2011 4th Grade
Male	Control	1444	1386	1407
	Journeys	1624	1646	1607
Female	Control	1411	1411	1503
	Journeys	1606	1720	1723
White	Control	406	370	392
	Journeys	657	654	589
Hispanic	Control	2296	2240	2325
	Journeys	2160	2204	2330
African American	Control	142	145	158
	Journeys	328	396	349
Other Race	Control	13	42	35
	Journeys	85	112	62
Free/ Reduced Lunch	Control	2370	2325	2438
	Journeys	2324	2412	2463
LEP	Control	906	829	1188
	Journeys	749	866	1212
Special Ed	Control	10	13	5
	Journeys	6	17	14

Of note, the TEA blocks access to test scores for students within subpopulations (and combinations thereof) when there are less than five students. As a result, the aforementioned sample sizes are smaller than the actual data that is available for each school. Indeed, this "masking" process resulted in a loss of 6% to 25% of the sample (varies by year and sample). However, in order to comply with FERPA, the TEA must follow this process.

Curricula

Prior to discussing the varying curricula used by participating schools, it is important to note that teachers are all generally teaching similar reading/language arts concepts and, due to state and local curricular guidelines which are typically aligned to state

assessments, tend to cover similar content (e.g., phonics, fiction/nonfiction, etc..). Thus, there are similarities in content covered between treatment and control programs. That said, the focus of this study was to examine the effects of an entire core curriculum (***Journeys***) and as such, it must be compared to other core curricula that teach the same content area.

2012 *JOURNEYS* PROGRAM

The ***Journeys*** literacy program is a new, comprehensive K–6 literacy program that targets key elements of literacy including reading comprehension, vocabulary, fluency, grammar, writing, and at Grades K–2, phonics and phonemic awareness. Designed to meet the diverse needs of all students, every lesson allows the student to develop comprehension and fluency focusing on a target skill and target strategy in a relevant short story and non–fiction story companion. The ***Journeys*** Student Edition includes vocabulary instruction that takes students through key steps in acquiring, practicing, and applying a rich vocabulary. The ***Journeys*** program also includes weekly interactive lessons, Leveled Readers by Irene Fountas, Vocabulary Readers, and intervention support for struggling readers. For teachers, ***Journeys*** offers easy organization with Teacher's Editions that include whole and small group instruction and a Focus Wall that provides a blueprint for weekly instruction. The Grab and Go kit included in the program keeps classroom resources such as worksheets and transparencies all in one manageable location. Other key features of the program:

- ***Journeys*** Digital classroom that includes listening, writing, and reading supports to extend student learning and provide necessary skills for the 21st century.
- Reading Primary Toolkit provides instructional routines that reinforce and apply the principles of phonics, phonemic awareness, vocabulary, fluency, and comprehension.
- Small Group plans in the Teacher's Edition include Ready Made Workstation flip charts and a lesson plan for every leveled reader.

In Texas, sample schools began using the ***Journeys*** program in the 2010–2011 school year at grades K-5. Therefore, Spring 2011 is the first "post" year of data available. However, it is important to note that these schools had used the ***Journeys*** program for only approximately 7–8 months (state testing occurred in April). Therefore, during this first year of analyses of Texas state assessment data, it is unlikely that effects, if any are observed, will be meaningful. After all, it takes time for teachers and students to become accustomed to a new curricular program. Nevertheless, data was analyzed to determine whether preliminary patterns emerged.

CONTROL SITE CURRICULA

A total of six distinct reading/language arts curricula were employed by control schools. This includes one district–created curriculum. However, two reading programs were primarily used by control schools. Reading program 1 was used by 58% of the control sample; most schools used the 2010 and 2007 editions. This program is a comprehensive Reading Language Arts program that is designed to motivate students with engaging reading selections and improve test scores. The program provides K–6 students with fiction and non-fiction readers and a built-in technology component that includes games and animated comprehension activities. For teachers, the program offers instructional strategies for differentiation, a teaching management system, and small group lessons designed to address various reading levels. This program also includes a built-in writing program that allows the student to build writing fluency and make connections between reading and writing.

Reading program 2 is being used by 26% of the control schools. This program, designed for Grades Pre-K–6, includes scientifically research-based instruction and teaching tools, including leveled readers, trade books, ELL support, grammar and writing practice, benchmark assessments, and graphic organizers. This program also includes differentiated instruction activities with an emphasis on ongoing progress monitoring. Additionally, this program prioritizes skill instruction so teachers can focus on appropriate reading skills.

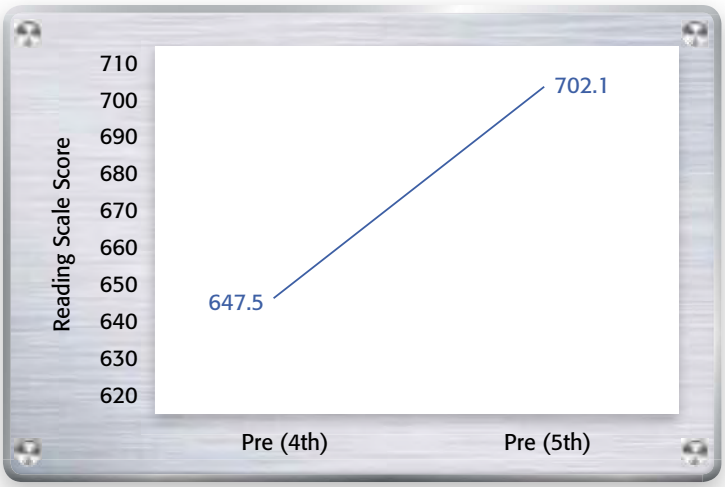
RESULTS

Are there significant changes in the reading performance of students who use *Journeys* over time?

To examine the extent to which students experienced learning gains while using the ***Journeys*** program, paired sample t–tests were performed[4]. Since schools began using ***Journeys*** in the 2010–11 school year, performance from the prior school year (Spring 2010, when students were in 4th grade) was compared to performance following ***Journeys*** usage (Spring 2011, when students were in 5th grade) from Sample B. This allowed researchers to examine if there was significant growth among the same group of students from one year to the next. Results showed significant growth in reading performance on the TAKS reading scale score, $p<.05$, $d=.45$. Conversion of these results to effect sizes shows that students grew 17 percentile points on the TAKS reading scale score.

4 Details on the statistical analyses performed are provided Appendix B.

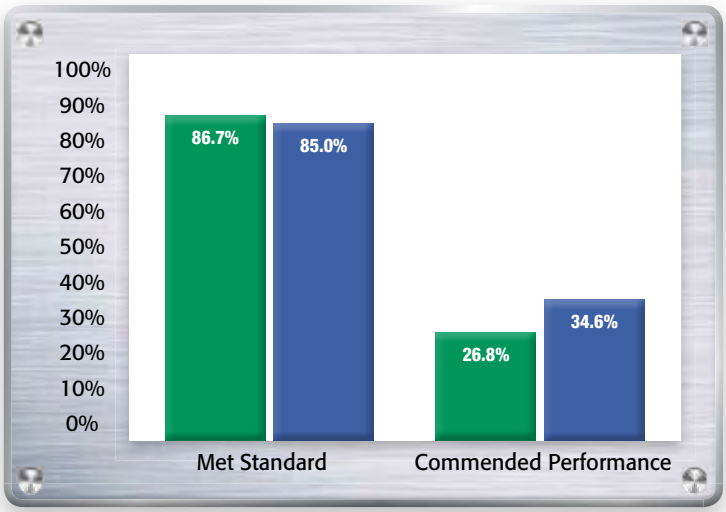
FIGURE 1. PRE AND POST TAKS READING PERFORMANCE OF *JOURNEYS* STUDENTS: SAMPLE B



*Students who used **Journeys** showed significant growth in reading performance as measured by the Texas state assessment (TAKS).*

Another way to examine whether positive changes were observed among students following usage of the ***Journeys*** program is to look at changes in the percent of students who met reading performance standards and the percent of students who were noted for commended performance. As shown in Figure 2, while there was relatively no change in the percent of students meeting the state reading performance standard, there were increases in the percent of ***Journeys*** students who showed commended performance. According to the TEA, students showing commended performance have "high academic achievement, considerably above state passing standards, and have a thorough understanding of the Texas reading curriculum."

FIGURE 2. PERCENT OF *JOURNEYS* STUDENTS MEETING TEXAS READING STANDARDS/COMMENDED PERFORMANCE: SAMPLE B



*While there was relatively no change on the percent of **Journeys** students who met the Texas reading standard, there was an increase in the percent of students who showed commended performance, the highest level of performance in Texas.*

Analyses were also performed on Sample C. As previously noted, this sample consists of all 3rd, 4th and 5th graders in participating schools in 2009, 2010, and 2011. Since schools began using ***Journeys*** in the 2010–11 school year, 2011 test data can be considered "post" and compared to performance levels prior to ***Journeys*** ("pre"). Results from ANOVA analyses showed variations in growth patterns by grade. As shown in Figures 3–6, 3rd graders showed the highest level of reading performance in 2010 (before ***Journeys***) as compared to 2011 (following ***Journeys***). In contrast, among 4th and 5th graders, the highest reading score was observed in 2011, following ***Journeys*** usage, as compared to prior testing years. With respect to writing, 4th grade students' performance on the writing assessment steadily declined over the three testing years. That said, the decline was smaller following ***Journeys*** usage than prior to exposure to this program. In sum, there is some inconsistency in the results obtained by grade level. As more data is obtained over the next two years, this may be clarified.

FIGURE 3. *JOURNEYS* SCHOOLS' TAKS READING PERFORMANCE FROM 2009–2011: GRADE 3

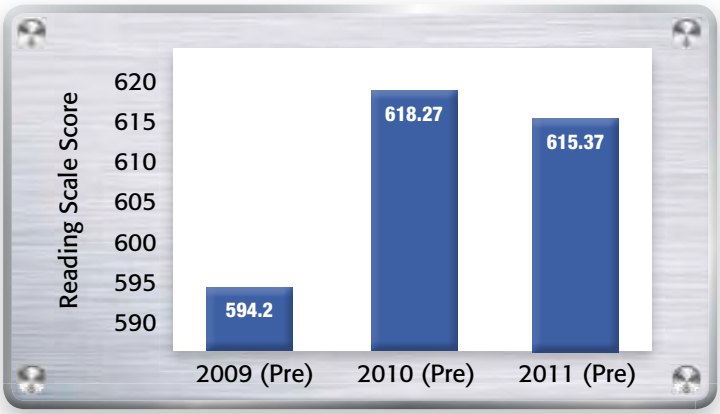


FIGURE 4. *JOURNEYS* SCHOOLS' TAKS READING PERFORMANCE FROM 2009–2011: GRADE 4

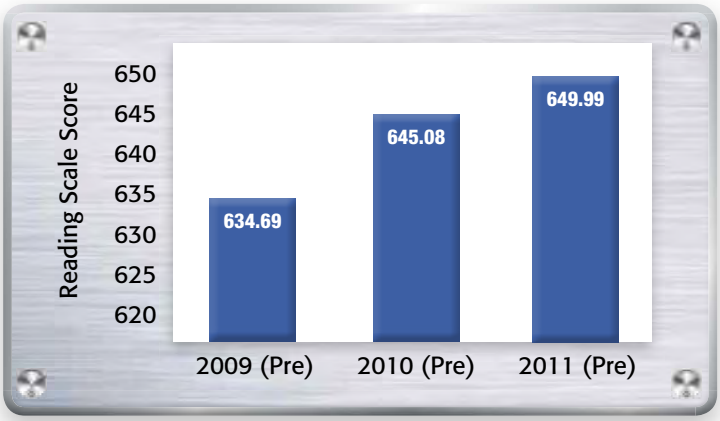


FIGURE 5. *JOURNEYS* SCHOOLS' TAKS WRITING PERFORMANCE FROM 2009–2011: GRADE 4

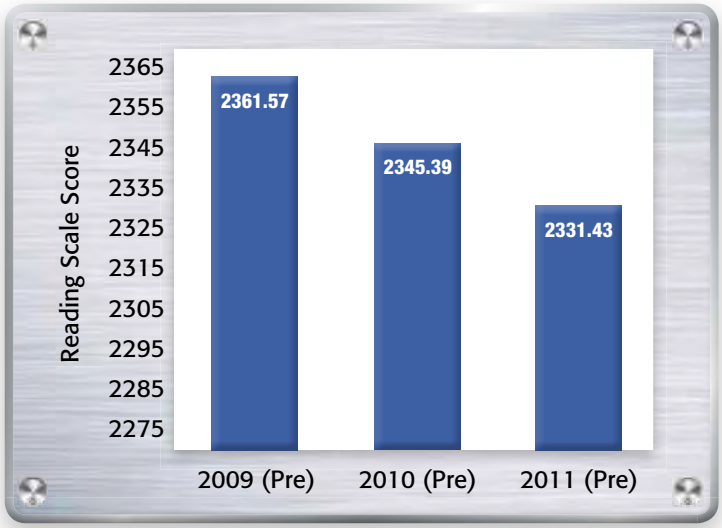
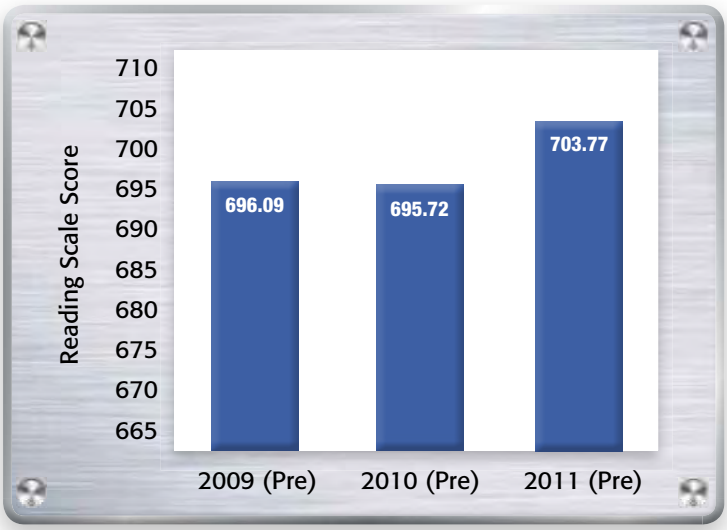


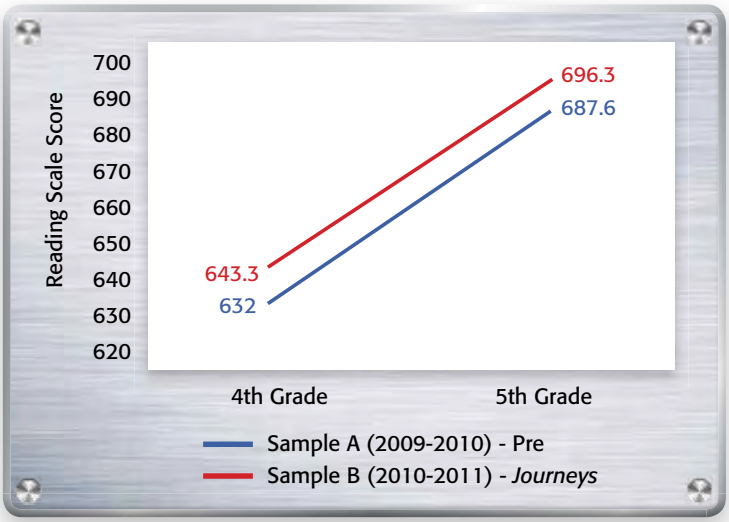
FIGURE 6. *JOURNEYS* SCHOOLS' TAKS READING PERFORMANCE FROM 2009–2011: GRADE 5



While 3rd grade students showed the highest level of reading performance in 2010, prior to ***Journeys****, 4th and 5th grade students showed the highest level following usage of* ***Journeys*** *in 2011.*

To more explicitly examine whether there were significant changes following usage of ***Journeys***, analyses were performed to examine whether gains were greater during the year prior to ***Journeys*** use (Spring 2009 to Spring 2010) or after ***Journeys*** use (Spring 2010 to Spring 2011). Two methods were employed. First, 4th to 5th grade growth from Spring 2009 to Spring 2010 (before ***Journeys***) was compared to 4th to 5th grade growth from Spring 2010 to Spring 2011 (following ***Journeys***). Note that while growth from different school years are being compared, the samples come from the same set of schools (i.e., schools that used ***Journeys*** in 2010–11). Results from this analysis showed no significant differences in the growth rates from both school years, $p>.05$. As shown in Figure 7, the lines demonstrating growth from both school years are fairly parallel.

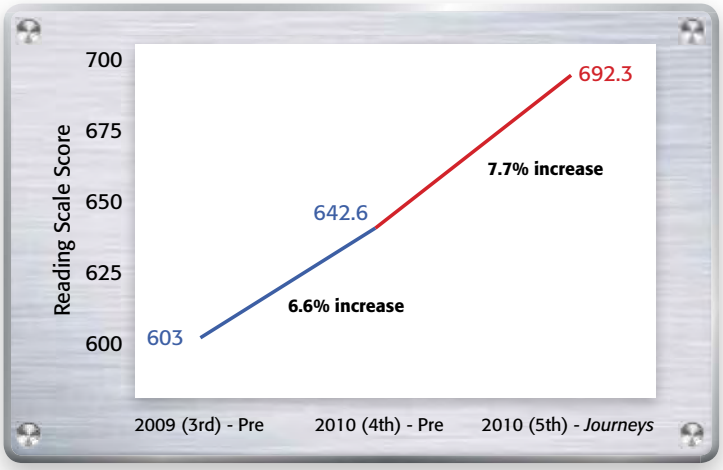
FIGURE 7. COMPARISON OF TAKS READING PERFORMANCE PRIOR TO *JOURNEYS* AND AFTER *JOURNEYS*



*Fourth to fifth grade growth rates were similar before and after the **Journeys** program was introduced at Texas schools.*

The second analysis performed consisted of comparing changes in growth rates among the same set of students (within subjects analysis). In particular, researchers compared growth from 3rd (2009) to 4th (2010) grade to growth from 4th (2010) to 5th (2011) grade. In the latter case, ***Journeys*** was introduced. Given that the TAKS is vertically scaled and rates of change should be constant, accelerated learning gains would be indicative of a positive impact. Results show that this was the case for the first year in which data is available, $p<.05$. Students experienced greater learning gains after their schools began implementation of the ***Journeys*** program as compared to before, see Figure 8.

FIGURE 8. DIFFERENCE BETWEEN PERFORMANCE PRIOR TO *JOURNEYS* AND AFTER *JOURNEYS*: WITHIN SUBJECT ANALYSIS



*Elementary students showed a significant increase in Texas state reading performance after their schools began implementing **Journeys** as compared to their performance prior to **Journeys**.*

Is *Journeys* associated with improvements for various subpopulations of students?

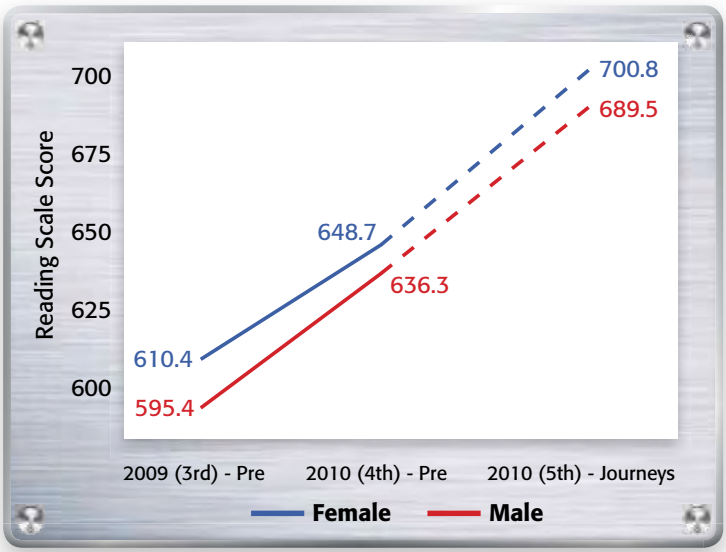
In order to address this question using the longitudinal TAKS Sample B data, HLM analyses were conducted. This analysis focused on the growth in performance from 3rd grade to 5th grade, and whether growth rates varied by subgroups following usage of ***Journeys***. In the two–level HLM analysis[5], change in performance from pre–testing (prior to ***Journeys***) to post–testing (following ***Journeys***) by the following student subgroups were analyzed: gender, economic disadvantage, Limited English Proficiency (LEP) status, and ethnicity.

Results of the HLM analysis on the TAKS data showed that the following variables were significantly related to accelerated learning gains: Limited English Proficiency, Whites, and African Americans. Among LEP students, students showed greater learning gains prior to ***Journeys*** than after.

5 It should be noted that given strong correlations between the various terms (associated with subgroups) and multicollinearities in the model, the subgroup effects were obtained by adding the interaction term(s) corresponding to each subgroup separately to the main effects model. Thus, separate models were run to obtain subgroup effects associated with each subgroup.

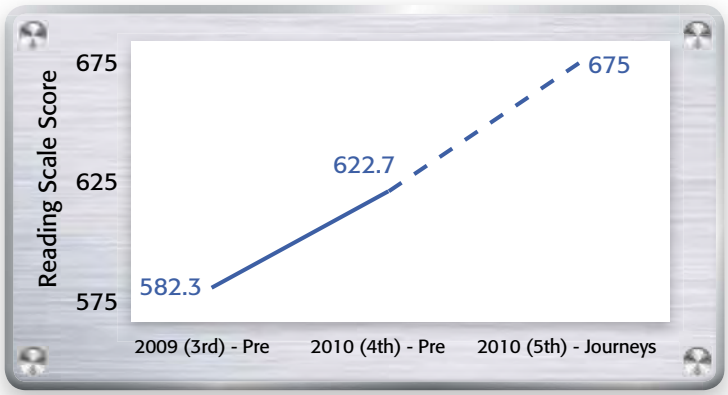
However, among Whites and African Americans, accelerated learning gains were observed following usage of ***Journeys***. A similar pattern was observed among Hispanics; however results were only marginally significant, p=.06. Figures 9–12 display gains made by subgroups before and after ***Journeys*** was introduced.

FIGURE 9. *JOURNEYS* STUDENTS' TAKS READING PERFORMANCE GROWTH BY GENDER



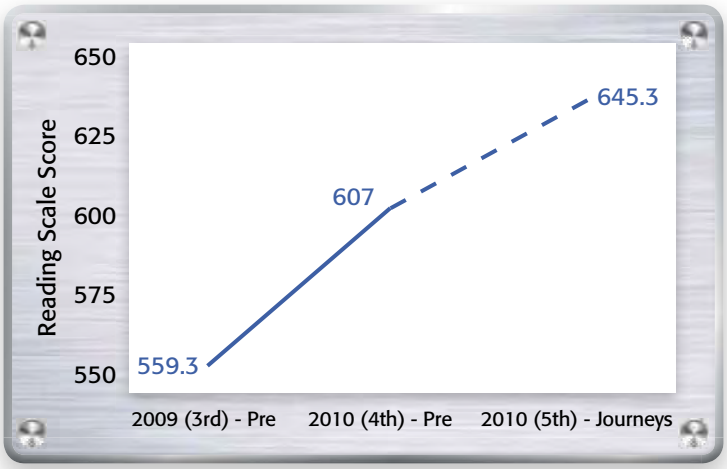
The annual rates of growth among female and male students were similar before (solid line) and after (dotted line) ***Journeys*** *was introduced.*

FIGURE 10. *JOURNEYS* STUDENTS' TAKS READING PERFORMANCE GROWTH BY ECONOMIC DISADVANTAGE



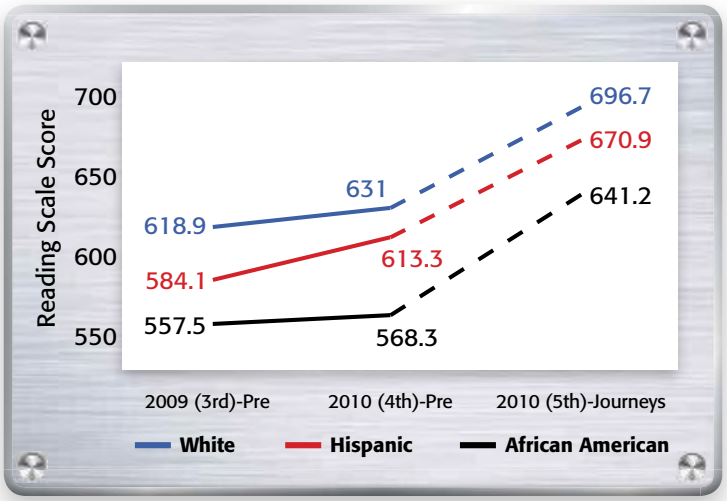
Among economically disadvantaged students, growth was similar before and after usage of ***Journeys****.*

FIGURE 11. *JOURNEYS* STUDENTS' TAKS READING PERFORMANCE GROWTH BY LIMITED ENGLISH PROFICIENCY STATUS



The annual rate of growth among Limited English Proficiency students was lower following ***Journeys*** *usage (see dotted line) as compared to before* ***Journeys*** *(solid line).*

FIGURE 12. *JOURNEYS* STUDENTS' TAKS READING PERFORMANCE GROWTH BY ETHNICITY

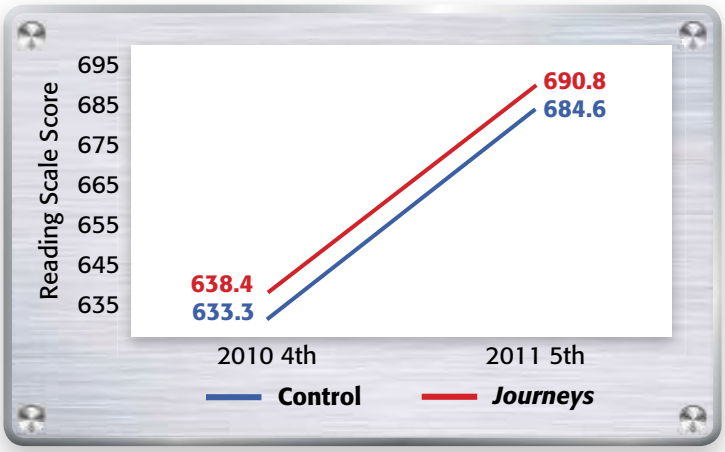


The annual rate of growth among African American and White students was significantly higher following usage of ***Journeys*** *than prior to* ***Journeys****. A similar pattern was observed among Hispanics though results were marginally significant.*

How does student achievement in reading on Texas state assessments differ across users and non-users of *Journeys*?

This set of analyses provides information on the relationship between ***Journeys*** and reading performance relative to schools that did not use ***Journeys***. Schools were coded based on whether they were a non–***Journeys*** school (0) or whether they used ***Journeys*** (1). Of interest in these analyses was whether or not group (***Journeys*** vs. non–***Journeys***) predicted growth in performance from Spring 2010 to Spring 2011[6]. Results showed that there were no significant differences in the growth in performance among ***Journeys*** and non– ***Journeys*** students, $p>.05$ [7]. Both types of students showed increases in reading performance, as measured by the TAKS reading scale score, see Figure 13.

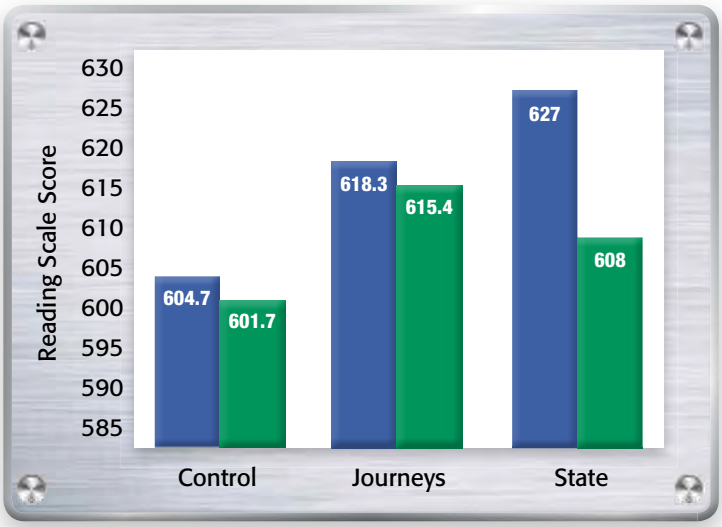
FIGURE 13. *JOURNEYS* AND CONTROL STUDENTS' READING PERFORMANCE FROM 2010 TO 2011



*Texas students using **Journeys** showed similar gains in reading as compared to students using other reading programs.*

Analyses were also performed to examine if there was a significant difference in the average reading/writing performance between ***Journeys*** and non–***Journeys*** students over time (from 2010 to 2011) by grade level. That is, for each grade level (3rd, 4th, and 5th) all students were compared to see if performance over the last two years varied by group[8]. In all cases, results showed no significant differences. That is, the change in performance from Spring 2010 to Spring 2011 among 3rd, 4th, and 5th grade ***Journeys*** students and non–***Journeys*** students was similar, $p<.05$. Figures 14–21 show these results, along with statewide results. In addition, the percentage of students who met the Texas reading/writing standard at both time periods for each grade level is provided for descriptive purposes.

FIGURE 14. TAKS READING SCALE PERFORMANCE FROM PRE TO POST BY GROUP AND STATE: 3RD GRADE



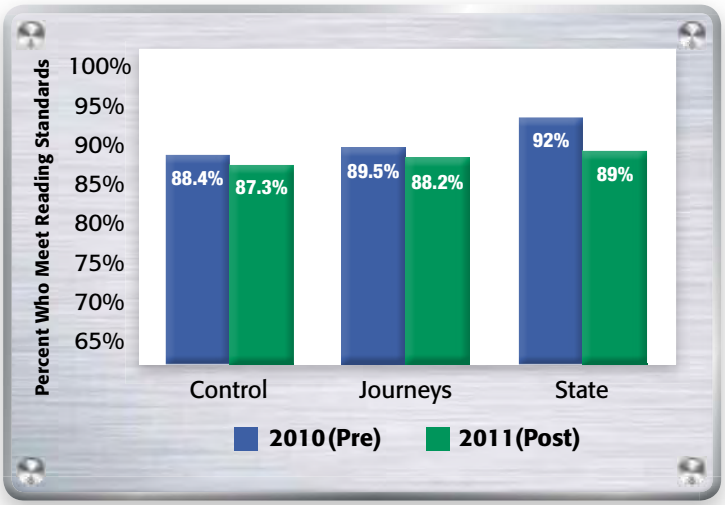
*While no differences were observed between **Journeys** and control 3rd grade students, the decline observed at the state level from 2010 to 2011 was much larger than that observed among study schools.*

6 Recall that ***Journeys*** was not introduced until the 2010–11 school year. Therefore, ***Journeys*** growth from 2010 to 2011 was compared to growth in control schools during the same years. Also note that cohort analysis of writing is not possible given that only 4th graders are assessed at the elementary level. However, crosssectional analyses were performed on the writing scale score and are presented subsequently.

7 Detailed statistical tables are presented in the Appendix B.

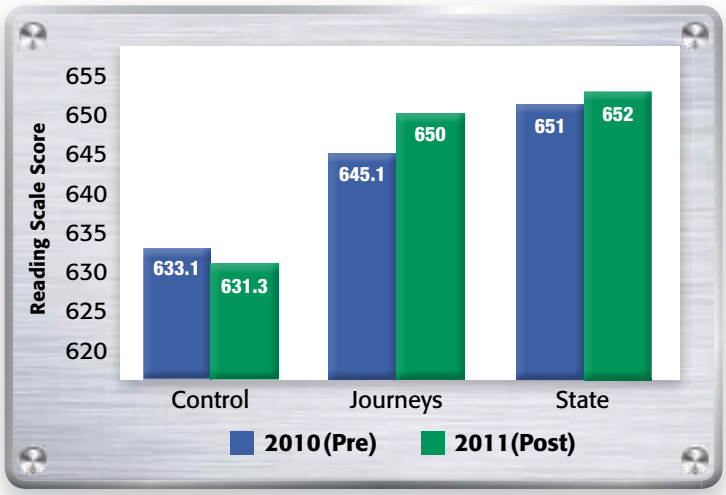
8 In these ANOVA analyses of the cross–sectional Sample C, the interaction between time and group is of primary interest. Appendix B contains the detailed statistics.

FIGURE 15. PERCENT OF STUDENTS WHO MET READING STANDARDS BY GROUP AND STATE: 3RD GRADE



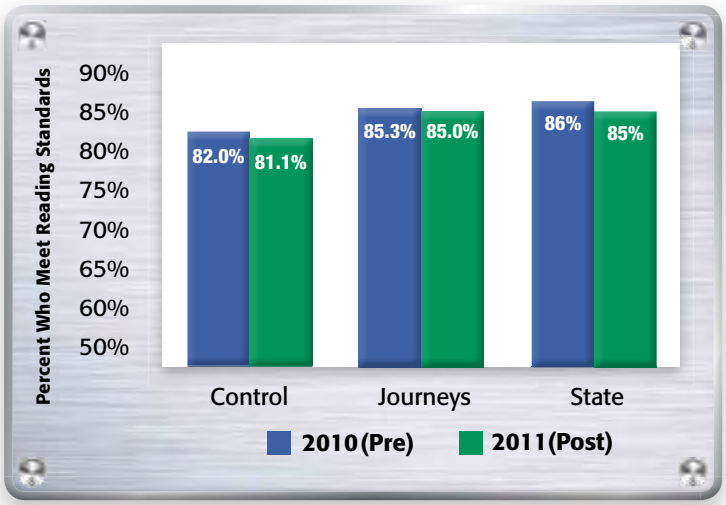
Among 3rd graders, the percentage of students who met the Texas reading standard was similar between groups during each test year.

FIGURE 16. TAKS READING SCALE PERFORMANCE FROM PRE TO POST BY GROUP AND STATE: 4TH GRADE



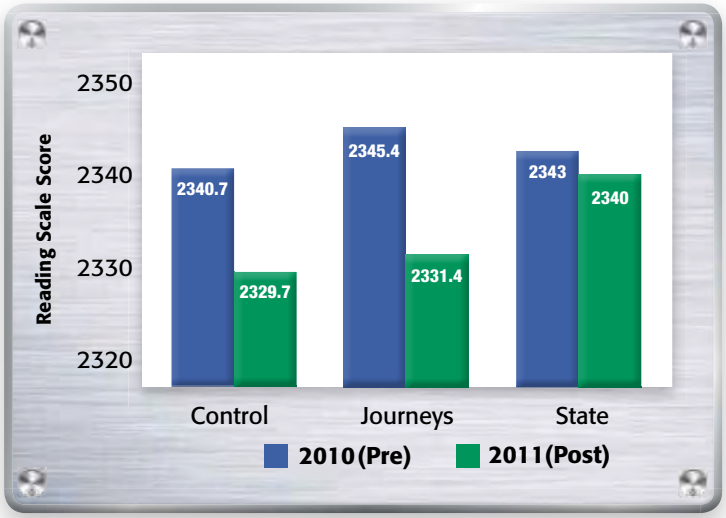
*While differences among 4th graders were not significant, a positive pattern was observed such that **Journeys** students made more positive changes in reading as compared to control students (and statewide).*

FIGURE 17. PERCENT OF STUDENTS WHO MET READING STANDARDS BY GROUP AND STATE: 4TH GRADE



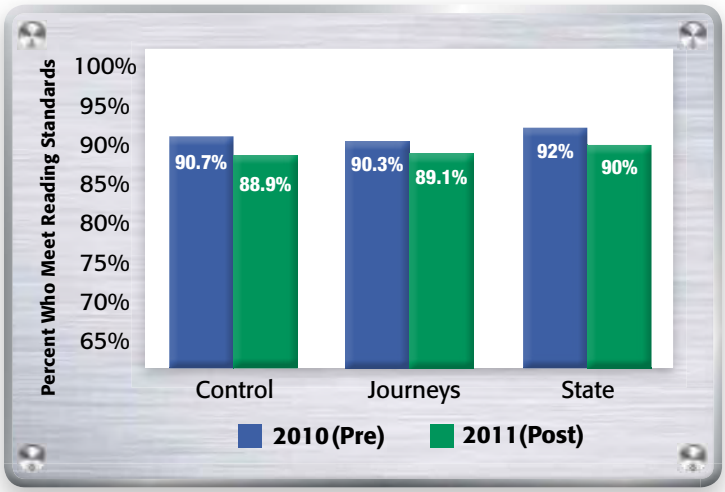
The percentage of 4th graders who met the Texas reading standard was similar across groups and over time.

FIGURE 18. TAKS WRITING SCALE PERFORMANCE FROM PRE TO POST BY GROUP AND STATE: 4TH GRADE



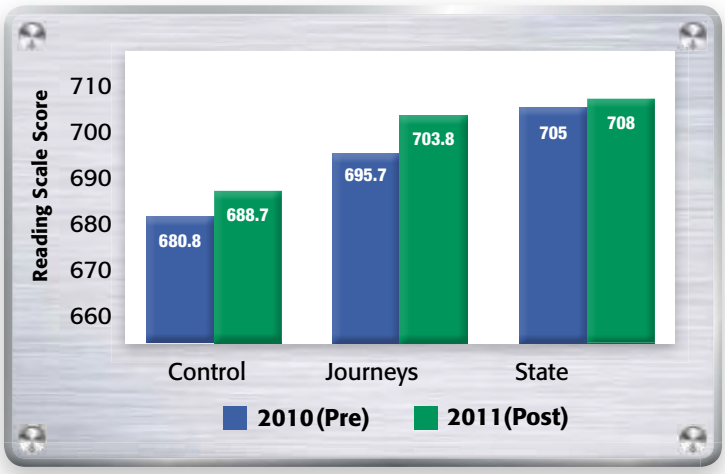
*Among all groups, there was a drop in 4th grade students TAKS writing scale score from 2010 to 2011. No significant differences were observed between control and **Journeys** students.*

FIGURE 19. PERCENT OF STUDENTS WHO MET WRITING STANDARDS BY GROUP AND STATE: 4TH GRADE



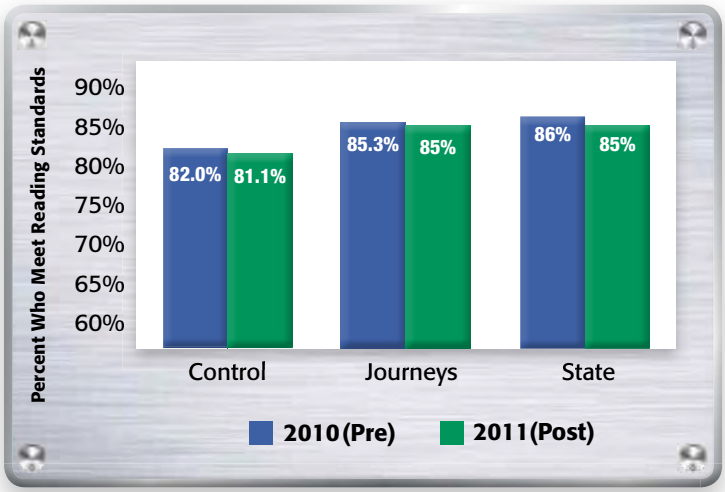
The percentage of 4th graders who met the Texas writing standard was similar for all groups.

FIGURE 20. TAKS READING SCALE PERFORMANCE FROM PRE TO POST BY GROUP AND STATE: 5TH GRADE



*Among 5th graders, the level of change from 2010 to 2011 in Texas reading performance was similar between **Journeys** and non-**Journeys** students.*

FIGURE 21. PERCENT OF STUDENTS WHO MET READING STANDARDS BY GROUP AND STATE: 5TH GRADE



*The percentage of 5th grade students who met the reading standard was similar for control and **Journeys** students (as well as statewide).*

DO FINDINGS VARY ACROSS DIFFERENT SUBGROUPS OF STUDENTS?

Exploratory[9] analyses by students' gender, ethnicity (White, Hispanic, African American), economically disadvantaged status, and LEP status were conducted to obtain preliminary information on whether there were significant differences between students in these subgroups who were in ***Journeys*** and non–***Journeys*** schools. HLM analyses were run separately for each subgroup using Sample B. In particular, gains from 2010 to 2011 were examined as a function of group and subgroup status. Results showed no consistent patterns among Sample B students as measured by the TAKS reading scale score. That is, ***Journeys*** and non–***Journeys*** students who were female, White, Hispanic, African American, economically disadvantaged, and LEP tended to perform similarly over time, p>.05. Figures 22–27 show subgroup results by group.

FIGURE 22. TAKS READING SCALE SCORE OF FEMALE STUDENTS BY GROUP AND TIME

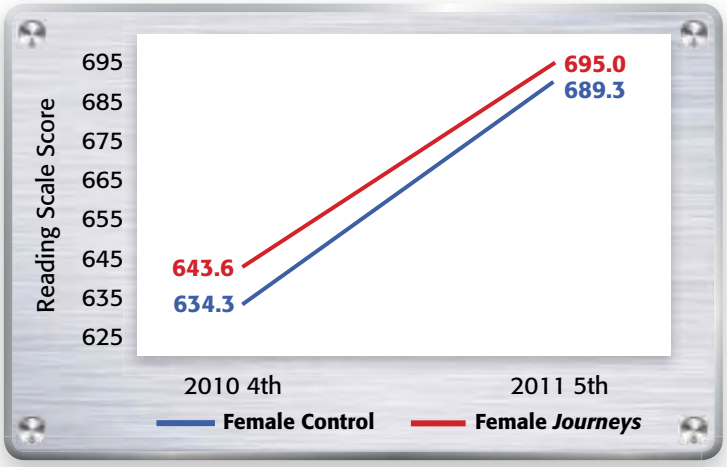


FIGURE 23. TAKS READING SCALE SCORE OF ECONOMICALLY DISADVANTAGED STUDENTS BY GROUP AND TIME

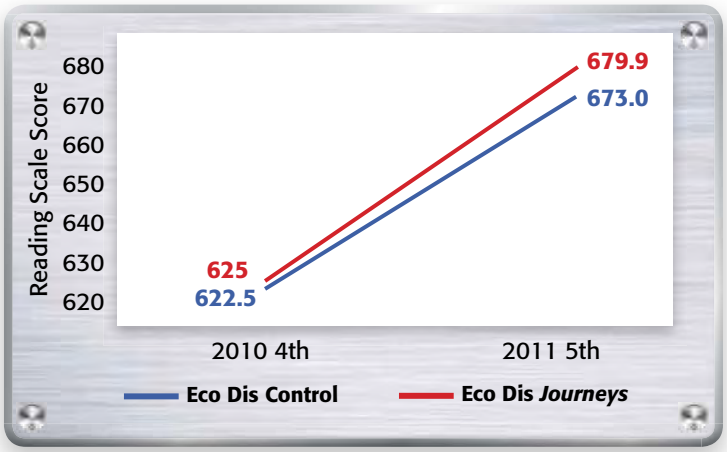


FIGURE 24. TAKS READING SCALE SCORE OF LIMITED ENGLISH PROFICIENT STUDENTS BY GROUP AND TIME

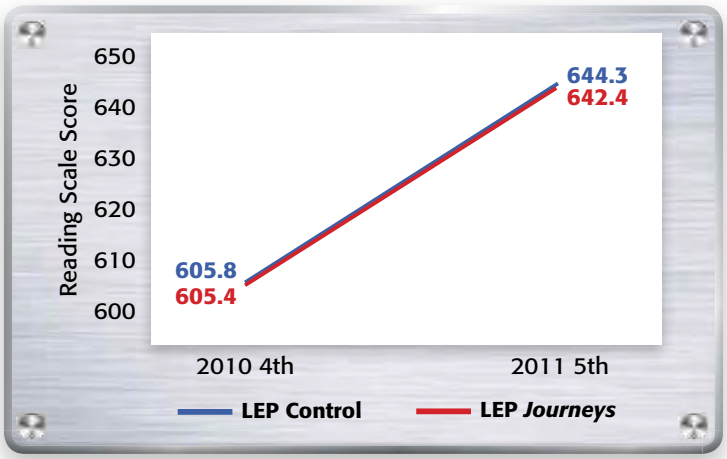
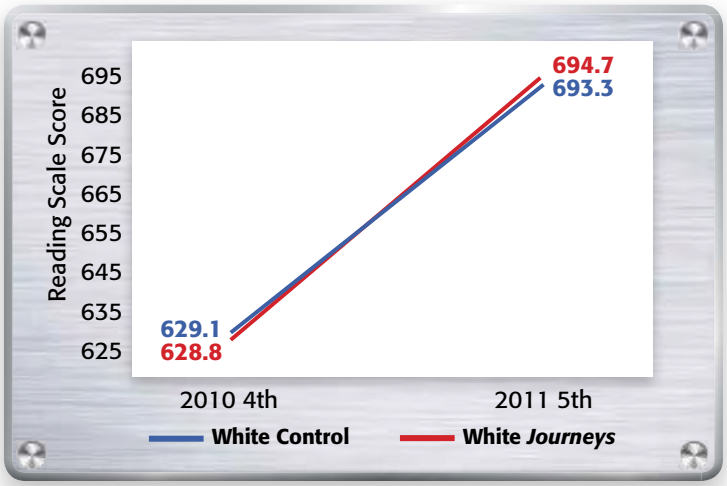


FIGURE 25. TAKS READING SCALE SCORE OF WHITE STUDENTS BY GROUP AND TIME



9 These analyses are exploratory because there have been very few studies that have examined subgroup effects relating to the curriculum of the ***Journeys*** program as well as elementary reading programs as a whole. In the absence of a strong program theory, the subgroup effects are viewed as empirical patterns that need theoretical frameworks and other rigorous experimental designs in the future to be estimated "causally." Further, analyses are based on smaller sample sizes.

FIGURE 26. TAKS READING SCALE SCORE OF HISPANIC STUDENTS BY GROUP AND TIME

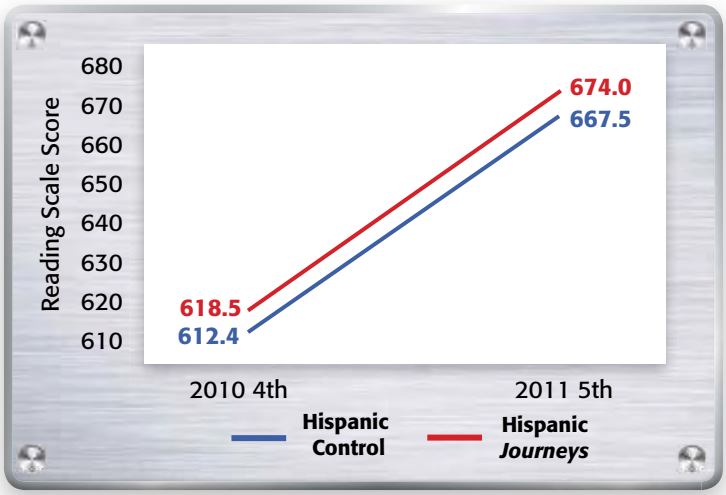
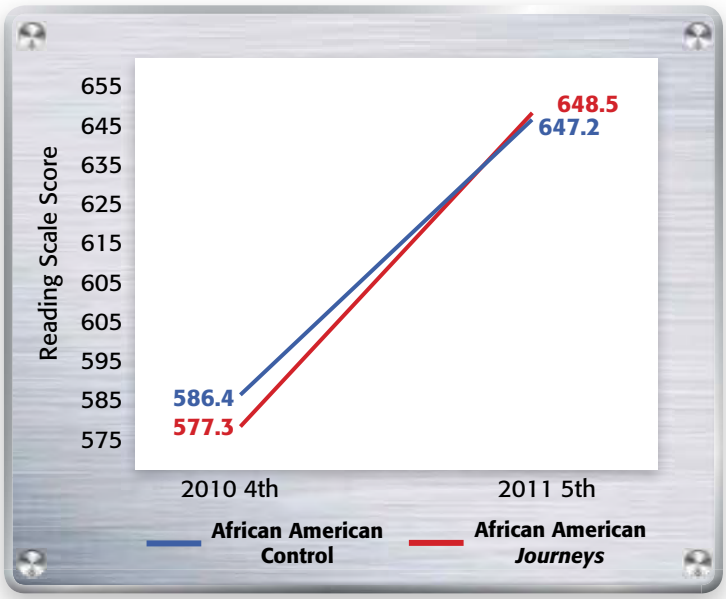


FIGURE 27. TAKS READING SCALE SCORE OF AFRICAN AMERICAN STUDENTS BY GROUP AND TIME



*Among all subgroups of students, no significant differences were observed between **Journeys** and non-**Journeys** students. Thus, females, Whites, Hispanics, African Americans, economically disadvantaged students, and LEP students who used and did not use **Journeys** showed similar rates of growth.*

CONCLUSION

Analyses of Texas state assessment data show that the ***Journeys*** program is associated with significant increases in student reading performance. Results showed a significant bump up in test scores following implementation of the ***Journeys*** program. These results are supported in part by the analysis of students at each grade level and their performance before and after ***Journeys*** was introduced (i.e., crosssectional analysis). Among 4th and 5th graders, the highest reading score was observed in 2011, following ***Journeys*** usage, as compared to prior testing years. In contrast, 3rd graders showed the highest level of reading performance in 2010 (before ***Journeys***) as compared to 2011 (following ***Journeys***).

In addition, results showed that students using ***Journeys*** and other literacy programs had similar levels of reading performance as measured by the TAKS reading scale score. That is, there were no significant differences observed among ***Journeys*** and non–***Journeys*** students. Students within these groups and in sub–populations (White, Hispanic, African American, economically disadvantaged status, and LEP) also showed no significant differences in learning gains. While the present study showed no overall effects of the ***Journeys*** program in Texas as compared to students using other literacy programs, it should be noted that students in ***Journeys*** schools had only been using the program for 7–8 months. Given that there is a learning curve when any new curriculum is introduced, it is not surprising that significant effects have not been obtained during the limited time period in which comparisons were made. As additional data is collected over the next two years (2012 and 2013), researchers will be able to look at more long term effects of the program within Texas. Moreover, a national experimental study is being conducted (2011–12 and 2012–13 school years) which is designed to examine the efficacy of this program within more controlled settings.

In summary, the results of this study using state assessment data provides some preliminary support for a positive relationship between the ***Journeys*** program and elementary reading performance. However, more conclusive findings on the overall impact of the ***Journeys*** program will be obtained as additional data is collected.

APPENDIX A: PROPENSITY SCORING MATCHING METHOD

The following three-step procedure was used to match the ***Journeys*** schools to the non–***Journeys*** schools:

Step 1. First the propensity to be a ***Journeys*** school is modeled as a function of school–level covariates. A logistic regression model is used to model the propensity to be a ***Journeys*** school. The predicted probability from the logistic regression serves as a measure of the propensity of being a ***Journeys*** school, and is also used as a distance measure to implement the matching described below. This predicted probability serves to reduce the multidimensional school–level characteristics into a single number that can be used to match ***Journeys*** and non– ***Journeys*** schools.

Variables included in the initial logistic regression model include:

- Enrollment
- Percent Economically Disadvantaged
- Percent Limited English Proficient
- Percent Special Ed
- Percent Gifted
- Percent At–Risk (504)
- Mobility Rate
- Percent White
- Percent Hispanic
- Percent Black
- Percent Native American
- Percent Asian/Pacific Islander

Step 2. Matches for the treatment group were obtained from the control group using a nearest neighbor algorithm (Ho et al., 2005, p. 9): "Matches are chosen for each treated unit one at a time, and at each matching step we choose the control unit that is not yet matched but is closest to the treated unit on the distance measure."

Step 3. Balance was assessed through t–tests of means to examine differences in means of each of the measures between the ***Journeys*** and non– ***Journeys*** schools. Tables A1–A3 describe the means for each of the measures in the two groups. No significant differences were obtained in any of the measures between the matched ***Journeys*** and non– ***Journeys*** schools.

TABLE A1. DIFFERENCES IN MEANS BETWEEN TEXAS *JOURNEYS* AND NON–*JOURNEYS* SCHOOLS

	Group	N	Mean	Standard Deviation	Standard Error Mean
School Enrollment	Control	50	536.76	179.15	25.34
	Journeys	57	547.72	223.65	29.62
Percent White	Control	50	19.36	26.10	3.69
	Journeys	57	21.57	27.15	3.60
Percent LEP	Control	50	36.18	25.07	3.55
	Journeys	57	33.69	28.33	3.75
Percent Economically Disadvantaged	Control	50	78.36	22.37	3.16
	Journeys	57	74.41	24.36	3.23
Percent At Risk	Control	50	60.05	22.50	3.18
	Journeys	57	52.11	26.93	3.57
Total Enrollment	Control	50	540.48	187.97	26.58
	Journeys	57	537.81	214.29	28.38
Percent African American	Control	50	8.15	14.99	2.12
	Journeys	57	11.32	18.62	2.47
Percent Hispanic	Control	50	70.42	31.79	4.50
	Journeys	57	64.25	34.25	4.54
Percent Native American	Control	50	.28	.74	0.10
	Journeys	57	.24	.33	0.04
Percent Asian/ Pacific Islander	Control	50	1.79	4.96	0.70
	Journeys	57	2.63	7.28	0.96
Percent Special Ed.	Control	50	7.25	2.57	0.36
	Journeys	57	6.57	2.52	0.33
Percent Gifted	Control	50	5.25	3.97	0.56
	Journeys	57	6.43	5.22	0.69
Percent Mobility	Control	50	17.57	6.51	0.92
	Journeys	54	16.21	4.98	0.68

Note: No differences were significant at p<.05 level.

APPENDIX B: STATISTICAL ANALYSES

Results for *Journeys* Students

Are there significant changes in the reading performance of students who use *Journeys* over time?

TABLE B1. PAIRED T-TEST FOR SAMPLE B *JOURNEYS* ONLY: CHANGE FROM PRE TO POST

Variables		N	Mean	Sd	t	df	Sig. Level
Reading Scale Score	Spring 2010	4667	647.50	120.86	–40.538	4666	.000
	Spring 2011	4667	702.06	119.16			

TABLE B2. RESULTS FOR SAMPLE C *JOURNEYS* ONLY: CROSS-SECTIONAL CHANGE

Variables		N	Mean	Sd	Statistic	df	Sig. Level
3rd Grade	2009	3418	594.20	119.30	F=25.817	2, 10364	.000
	2010	3462	618.27	165.95			
	2011	3487	615.37	164.28			
4th Grade	2009	3053	634.69	103.64	F=11.622	2, 9755	.000
	2010	3375	645.08	135.88			
	2011	3330	649.99	141.65			
4th Grade Writing	2009	3121	2361.57	244.60	F=13.847	2, 9830	.000
	2010	3379	2345.39	214.05			
	2011	3333	2331.43	231.48			
5th Grade	2009	3033	99.34	99.34	F=4.195	2, 9626	.015
	2010	3230	121.24	121.24			
	2011	3366	152.41	152.41			

Do growth patterns vary before and after schools began using *Journeys*?

TABLE B3. HLM RESULTS FOR SAMPLE A (2009-2010) VERSUS SAMPLE B (2010-2011) PERFORMANCE ON TAKS READING SCALE SCORE

Variables	Coefficient	Std. Error	t-ratio	Sig. Level
Intercept, β00	631.96	1.55	408.64	<.001
Intercept* Year	11.30	2.57	4.40	<.001
Time Slope	55.63	1.21	46.04	<.001
Year Slope	–2.63	2.01	–1.31	0.19

TABLE B4. HLM RESULTS FOR SAMPLE B PERFORMANCE ON TAKS READING SCALE SCORE (GROWTH BY PRE-POST *JOURNEYS* EXPOSURE)

Variables	Coefficient	Std. Error	t-ratio	Sig. Level
Intercept, β00	603.03	2.13	282.94	<.001
Time Slope	39.60	1.88	21.07	<.001
Pre-Post Slope	13.05	3.10	4.20	<.001

Is *Journeys* associated with improvements for various subpopulations of students?

TABLE B5. HLM RESULTS FOR *JOURNEYS* SUBGROUPS

Total Outcome Measures	Coefficient	Std. Error	t-ratio	Sig. Level
Reading Scale Score				
Baseline* Group	5.04	7.66	0.66	0.511
Change* Group	1.23	2.47	0.50	0.618

*Significant at the p<.05 level.

Results between *Journeys* and Control Students

How does student achievement in reading on Texas state assessments differ across users and non–users of *Journeys*?

TABLE B6. HLM RESULTS FOR *JOURNEYS* VS CONTROL SUBGROUPS: LONGITUDINAL SAMPLE B

Total Outcome Measures	Coefficient	Std. Error	t-ratio	Sig. Level
Reading Scale Score				
Baseline* Group	5.04	7.66	0.66	0.511
Change* Group	1.23	2.47	0.50	0.618

*Significant at the p<.05 level.

TABLE B7. ANOVA RESULTS FOR *JOURNEYS* VS CONTROL SUBGROUPS: CROSS-SECTIONAL SAMPLE C

Variables			N	Mean	Sd	F (time*Group)	df	Sig. Level
Grade 3 Reading	Control	2010	3462	618.27	165.945	F=.001	1, 12937	.976
		2011	3487	615.37	164.278			
	Journeys	2010	3033	604.72	116.907			
		2011	2959	601.67	131.383			
Grade 4 Reading	Control	2010	3375	645.08	135.875	F=2.08	1, 12483	.149
		2011	3330	649.99	141.654			
	Journeys	2010	2872	633.09	123.242			
		2011	2910	631.29	112.474			
Grade 4 Writing	Control	2010	3379	2345.39	214.046	F=0.14	1, 12523	.712
		2011	3333	2331.43	231.481			
	Journeys	2010	2897	2340.74	208.245			
		2011	2918	2329.68	219.820			
Grade 5 Reading	Control	2010	3230	695.72	121.241	F=.000	1, 12246	.984
		2011	3366	703.77	152.410			
	Journeys	2010	2857	680.60	111.947			
		2011	2797	688.74	129.749			

Do such findings vary across different subgroups of students?

TABLE B8. HLM RESULTS FOR *JOURNEYS* VS CONTROL SUBGROUPS: SAMPLE B

Total Outcome Measures	Coefficient	Std. Error	t-ratio	Sig. Level
Gender (Female)				
Baseline* Group	9.27	5.03	1.84	0.065
Change* Group	–3.65	3.45	–1.06	0.29
Economically Disadvantaged				
Baseline* Group	2.47	5.36	0.46	0.645
Change* Group	1.42	2.80	0.51	0.612
English Language Learners				
Baseline* Group	–0.36	6.10	–0.06	0.953
Change* Group	–1.60	3.97	–0.40	0.687
Ethnicity (Hispanic)				
Baseline* Group	7.46	6.52	1.14	0.253
Change* Group	–0.44	2.90	–0.15	0.88
Ethnicity (Black)				
Baseline* Group	–12.75	14.73	–0.87	0.387
Change* Group	10.47	9.46	1.11	0.268
Ethnicity (White)				
Baseline* Group	–3.55	9.68	–0.37	0.714
Change* Group	1.19	5.86	0.20	0.84

Evaluate *Journeys Common Core*

hmheducation.com/journeys

PRES ASSOCIATES
PO Box 10730
Jackson, WY 83002
Phone: 307-733-3255, Fax: 307-222-0312
info@presassociates.com