

## Final Report

# A Study on the Effectiveness of *Journeys* in Texas Using Existing Data Sources

### PREPARED BY:

MIRIAM RESENDEZ, M.A., SENIOR RESEARCHER  
MARIAM AZIN, PH.D., PRESIDENT  
JULY 2014



For inquiries, please contact PRES Associates at:  
[info@presassociates.com](mailto:info@presassociates.com)  
(307) 733-3255

© 2014 PRES Associates, Inc.

## Executive Summary

---

The Houghton Mifflin Harcourt Journeys program is a research-based comprehensive K-5 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 since the 2011-12 school year. PRES Associates, an external, independent educational research firm with over twenty years of experience in applied educational research and evaluation, conducted final year three (2013) analyses for a multi-year quasi-experimental study on the effectiveness of the Journeys program in the state of Texas. This study looked at three year performance trends on Texas state assessment data among schools using Journeys, as well as matched control schools using other literacy programs.

Major findings arranged by evaluation questions include the following:

### **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 showed a significant difference between Journeys students and students using other reading programs, after controlling for background variables including prior year(s) reading performance. Specifically, Journeys 3<sup>rd</sup> and 5<sup>th</sup> grade students (in 2013) demonstrated higher reading scores than non-Journeys students as measured by the new Texas state assessment – the STAAR.

No significant differences were observed among 4<sup>th</sup> grade students.

These findings are similar to the prior year's report (Year 2: 2012) in which significant differences were observed between Journeys and non-Journeys students on the STAAR reading scores. Thus, the present findings indicate that the positive effects observed in the prior year were sustained following almost three years of Journeys implementation.

This is further supported by the cross-sectional analyses conducted. In particular, a pattern was observed such that positive differences between Journeys and non-Journeys students at grades 3-5 were greatest in 2013 following three years of implementation as compared to 2011 and 2012 (following one and two years of Journeys implementation respectively). Moreover, this relationship was marginally significant among 4<sup>th</sup> graders. Such findings support the conclusion that as students use Journeys over time (across multiple school years), the positive impact of the program grows stronger.

### **Do such findings vary across different subgroups of students?**

Exploratory analyses revealed significant subgroup effects as well. Results showed that females, Hispanics, and economically disadvantaged Journeys students had higher STAAR reading scale scores as compared to non-Journeys students who were in these subpopulations, after controlling for background variables and prior year reading performance. These findings are also consistent with those obtained in 2012 (Year 2 report). No such differences were observed among Whites, African American or males. In sum, Journeys appears to be positively related to

improved test performance among important subgroups of students.

In summary, the results from the final year of analyses for this three year archival study provides additional support for a positive relationship between the Journeys program and elementary reading performance. Such findings are consistent with those obtained in a national randomized control trial conducted during the 2011-12 and 2012-13 school years. Analyses described in the first year RCT report showed that Journeys students significantly outperformed control students in the areas of reading comprehension, vocabulary, spelling, and word recognition. The consistency of positive effects in favor of the Journeys program across multiple outcomes and samples supports the conclusion that the Journeys program has a positive impact on student performance relative to other elementary reading/ language arts programs.

## Table of Contents

---

<b>Project Background .....</b>	<b>6</b>
<b>Project Overview.....</b>	<b>7</b>
<b>Design and Methodology.....</b>	<b>7</b>
Research Design.....	7
Measures .....	7
Samples.....	8
Curricula .....	10
<b>Results .....</b>	<b>12</b>
How does student achievement in reading on Texas state assessments differ across users and non-users of Journeys? .....	12
Do findings vary across different subgroups of students?.....	14
<b>Conclusion .....</b>	<b>16</b>
<b>Appendix A: Propensity Scoring Matching Method.....</b>	<b>17</b>
<b>Appendix B: Statistical Analyses.....</b>	<b>19</b>

## Table of Figures and Tables

---

### Figures

Figure 1. Texas Journeys and Control Students' STAAR Reading Performance (3 <sup>rd</sup> Grade – 2011) .....	13
Figure 2. Texas Journeys and Control Students' TAKS Reading Performance (4 <sup>th</sup> Grade – 2012) .....	13
Figure 3. Texas Journeys and Control Students' TAKS Reading Performance (5 <sup>th</sup> Grade – 2013) .....	12
Figure 4. Percent of 3rd Grade Students Meeting Reading Standards by Group and Year .....	14
Figure 5. Percent of 4 <sup>th</sup> Grade Students Meeting Reading Standards by Group and Year .....	14
Figure 6. Percent of 5 <sup>th</sup> Grade Students Meeting Reading Standards by Group and Year .....	14
Figure 7. STAAR Reading Scale Score of Female Students by Group (5 <sup>th</sup> Grade 2013).....	15
Figure 8. STAAR Reading Scale Score of Male Students by Group (5 <sup>th</sup> Grade 2013).....	15
Figure 9. STAAR Reading Scale Score of Economically Disadvantaged Students by Group (5 <sup>th</sup> Grade 2013).....	15
Figure 10. STAAR Reading Scale Score of Hispanic Students by Group (5 <sup>th</sup> Grade 2013) .....	15
Figure 11. STAAR Reading Scale Score of African American Students by Group (5 <sup>th</sup> Grade 2013) .....	15
Figure 12. STAAR Reading Scale Score of White Students by Group (5 <sup>th</sup> Grade 2013).....	16

### Tables

Table 1. TAKS Blueprint for Reading .....	8
Table 2. STAAR Blueprint for Reading.....	8
Table 3. Sample A (Sample Size) .....	9
Table 4. Sample B (Sample Size) .....	9
Table 5. Student Level Counts by Subgroup (Sample A: 2012-5 <sup>th</sup> Grade).....	10
Table A1. Differences in Means between Texas Journeys and Non- Journeys Schools .....	18
Table B1. Results for Journeys vs Control Subgroups: Longitudinal Sample A .....	19
Table B2. Results for Journeys vs Control Subgroups: Cross-sectional Sample B.....	20
Table B3. Results for Journeys vs Control Subgroups: Sample A .....	21

## Project Background

---

*“Literacy is a bridge from misery to hope. It is a tool for daily life in modern society. It is a bulwark against poverty, and a building block of development, an essential complement to investments in roads, dams, clinics and factories. Literacy is a platform for democratization, and a vehicle for the promotion of cultural and national identity. Especially for girls and women, it is an agent of family health and nutrition. For everyone, everywhere, literacy is, along with education in general, a basic human right.” - Kofi Annan*

Reading skills continue to be an area of difficulty for U.S. students (National Assessment of Education Progress, 2013). Only 35% of the nation’s 4<sup>th</sup> graders are performing at or above the proficient level, with little sign of change over the past several years (33% were proficient in 2011 and 34% were proficient in 2009). 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. Indeed 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 Publishers developed a new reading/language arts program that has shown promise as an effective instructional program for elementary school students. The 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 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.)<sup>1</sup> conducted Year 3 (2013) analyses examining the effectiveness of the Journeys program in helping elementary students in Texas improve their literacy skills and competency. The quasi-experimental study spanned the school years from 2010-2011 to 2012-2013, with cohort and longitudinal data presented for 3<sup>rd</sup> through 5<sup>th</sup> grade students each year. This report presents findings for all three school years to illustrate trends and long term effects of the program.

---

<sup>1</sup> 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 final 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. The year one report<sup>2</sup>, which presented results from 2008 to 2011, revealed significant increases in test scores following implementation of the Journeys program. Similar increases in reading performance (measured using the TAKS) were observed for students using Journeys and other literacy programs. In year 2, however, students using Journeys outperformed students using other reading programs as measured by the new Texas state assessment – the STAAR – after controlling for baseline reading scores.

The present analysis builds upon the prior analyses by examining the effects of the Journeys program following three years of implementation. Since teachers and students have more experience using the program, stronger effects were expected. Specifically, the present analyses were designed to address the following key evaluation questions:

1. How does student achievement in reading on Texas state assessments differ across users and non-users of Journeys?
2. Do findings vary across different subgroups of students?

---

<sup>2</sup> The Journeys Year One study report is available by request (email [info@presassociates.com](mailto:info@presassociates.com)).

## 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 to 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

Between spring 2003 and 2011, the TAKS (Texas Assessment of Knowledge and Skills) was the primary state assessment of academic skills. TAKS was designed to measure core areas of the state-mandated curriculum, the Texas Essential Knowledge and Skills (TEKS). The TAKS scale score became vertically scaled as of the Spring 2009 testing, which allows for growth analyses within the same subject area.

Beginning in spring 2012, the State of Texas Assessments of Academic Readiness (STAAR) replaced the TAKS. According to information from the TEA, the STAAR is more rigorous than previous state tests. It contains more test questions at most grade levels, and the majority of new STAAR assessments test material students studied that

year as opposed to being a cumulative measure. As well, for the first time since the state began its standardized testing program, the new tests have a time limit.

Under both assessments (TAKS and STAAR), testing in reading occurs annually for all students in grades 3-8. The following tables show the number items measured by objective for reading for both assessments. Of note, the STAAR and TAKS are not equated. As such, growth analyses cannot be done to measure change over the years between 2011 (last year of TAKS) and 2013 (second year of STAAR).

**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
<b>Total number of items</b>	<b>36</b>	<b>40</b>	<b>42</b>

**Table 2. STAAR Blueprint for Reading**

<b>Genres Assessed:</b> <b>Literary</b> • Fiction (Readiness) • Literary Nonfiction (Supporting) • Poetry (Supporting) • Drama (Ineligible) • Media Literacy (Embedded) <b>Informational</b> • Expository (Readiness) • Persuasive (Ineligible) • Procedural (Embedded) • Media Literacy (Embedded)	G. 3	G. 4	G. 5
Reporting Category 1—Understanding across genres	8	10	10
Reporting Category 2—Understanding/Analysis of literary texts	18	18	19
Reporting Category 3—Understanding/Analysis of informational texts	16	16	17
<b>Total number of items</b>	<b>40</b>	<b>44</b>	<b>46</b>

Previous analyses focused on directly comparing Journeys students with students using other reading programs at each year following adoption of the Journeys program (i.e., spring 2011 and 2012) while controlling for baseline performance (spring 2010). Baseline measures were established in 2010 for 3<sup>rd</sup> grade students with the Journeys program beginning implementation

in 2011. Unfortunately, there is no baseline for students entering 3<sup>rd</sup> grade in 2011 (the year 3 cohort); thus, the present analyses compare reading performance across 3 years, controlling for several factors that could influence reading to estimate the program effect.

The analyses described herein used the following outcome measures:

- 3<sup>rd</sup>-5<sup>th</sup> grade TAKS and STAAR reading scale scores,
- 3<sup>rd</sup>-5<sup>th</sup> grade proficiency levels (percent meeting standards).

## Samples

Elementary schools using the Journeys program in the 2010-11 school year were selected for inclusion in this study ( $n=55^3$ ). Control sites<sup>4</sup> ( $n=50$ ) 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

<sup>3</sup> 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.

<sup>4</sup> Similarly, only schools *confirmed* to be non-Journeys users by an independent call center, were included in this study.



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 samples were available. Sample A consisted of students who had longitudinal data available (from grades 3 to 5). Specifically, this sample consists of elementary students who were in the 3<sup>rd</sup> grade in the 2010-11 school year. Data for these same students were also obtained when they were in the 4<sup>th</sup> grade (2011-2012) and 5<sup>th</sup> grade (2012-2013). Table 3 displays the data collection timeline for this sample and total sample sizes.

**Table 3. Sample A (Sample Size)**

Group	Grade	10-11	11-12	12-13
Journeys schools= 55	3	A (2778)		
	4		A (3301)	
	5			A (3531)
Non-Journeys schools= 50	3	A (2293)		
	4		A (2661)	
	5			A (2795)

As previously noted, analyses on the comparability of the Journeys and control sites at the school-level showed no significant differences among any of the measured demographic characteristics when propensity matching was conducted using the baseline 2009-10 statistics,  $p > .05$ . However, at the student level and for the current sample A, student level differences did emerge among demographic and Year 1 (2010-11) reading performance. Given that baseline, pre-Journeys data was not available for this sample of students since assessment does not occur in the 2<sup>nd</sup> grade,

analyses controlled for both demographic and prior year performance to determine if effects observed in the prior year (2012) were sustained.

In addition to a cohort of students who was tracked over time, data was also collected from all students in grades 3-5 at each of the participating schools for each school year from 2009-10 to 2012-13. The table below shows the counts for this sample available for cross-sectional analyses.

**Table 4. Sample B (Sample Size)**

Group	Grade	09-10	10-11	11-12	12-13
Journeys schools= 55	3	3462	3487	3092	3475
	4	3375	3330	3244	3379
	5	3230	3366	3294	3529
Non-Journeys schools= 50	3	3033	2959	2710	2724
	4	2872	2910	2712	2806
	5	2857	2797	2775	2797

Table 5 shows the number of students within each of the subpopulations for each study sample. For illustrative purposes, only the data for a single grade level and 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.

Of note, the TEA blocks access to test scores for students within subpopulations (and combinations thereof) when there are less than 5 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.

**Table 5. Student Level Counts by Subgroup  
(Sample A: 2012-5<sup>th</sup> Grade)**

	Non-Journeys	Journeys
Male	1936	2210
Female	1884	2270
African American	277	532
Hispanic	2896	2881
White	535	837
Other Race/Ethnicity	113	230
Economically Disadvantaged	3028	3136
Special Ed	313	343

## 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/non-fiction, 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-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

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 include:

- 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 that 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. 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, analysis of spring 2011 data is less

powerful than analyses of data for spring 2012 and 2013 because of the learning curve that teachers are likely to experience during their first year of implementation. After all, it takes time for teachers and students to become accustomed to a new curricular program. Thus, analysis of spring 2012 and 2013 data is more sensitive to any treatment effects.

instruction activities with an emphasis on ongoing progress monitoring. Additionally, this program prioritizes skill instruction so teachers can focus on appropriate reading skills.

### **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 PreK – Grade 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

## Results

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

As previously reported in the Journeys Year One Report, no significant differences were observed between Journeys and non-Journeys students -- both types of students showed increases in reading performance from 2010 to 2011. However, since Journeys schools had only been using the program for 7-8 months, such findings were not surprising as it takes time for teachers and their students to learn a new curriculum. In the Year Two Report, students using Journeys for almost two years outperformed students using other reading programs as measured by the new Texas state assessment – the STAAR – after controlling for baseline reading scores. The present analyses focuses on whether any differences exist following three years of implementation in order to look at the more long term effects of the program within Texas.

Schools were coded based on whether they were a non-Journeys school (0) or whether they used Journeys (1). Because the 2012 STAAR data is not directly comparable to the TAKS data, longitudinal analyses was not possible. Instead, researchers examined whether differences existed within each grade cohort: the 2013 fifth graders (STAAR), 2012 fourth graders (STAAR), and 2011 3<sup>rd</sup> graders (TAKS) after controlling for demographics and previous TAKS/STAAR performance<sup>5</sup>, respectively. Of note, the student samples by grade vary from each previous year. The

students included in the current report must have been in the same school as 5<sup>th</sup> graders in the 2012-13 school year; this means that the samples for each grade are smaller in the Year Two and Year Three reports as compared to the Year One report, which did not have this requirement.

The analyses in the Year One Report were run using hierarchical linear modeling (HLM) to examine differences in growth rates whereas the Year Two Report analyses explicitly controlled for baseline performance (ANCOVA). The current analyses did not control for baseline performance (as there were no years when Journeys was not being implemented); instead, reading performance was compared for the Journeys and non-Journeys students, controlling for demographics and previous year(s) reading performance to determine if learning gains *persisted* over the years. Thus, any significant gains observed have occurred despite controlling for prior learning gains that may have been associated with the Journeys program.

First, students who were in 3<sup>rd</sup> grade during the 2010-11 school year (when Journeys was first implemented) were examined. Given the lack of baseline reading performance data (i.e., in 2009-10 when these students were in 2<sup>nd</sup> grade), analysis only control for demographic background variables. Results showed a statistically significant difference such that Journeys students outperformed non-Journeys students on the TAKS reading test,  $p < .05$ <sup>6</sup>, see Figure 1.

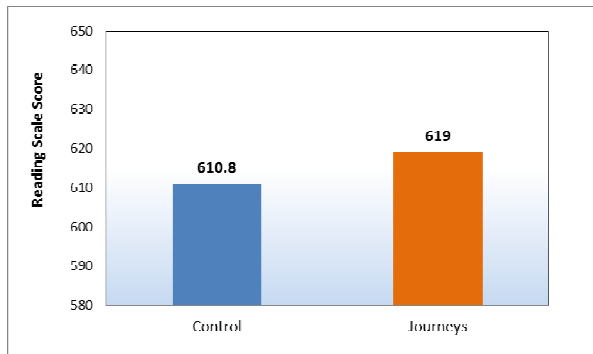
---

<sup>5</sup> Recall that Journeys was not introduced until the 2010-11 school year. Therefore, 3<sup>rd</sup> grade data (from the 2009-10) serves as baseline data since Journeys schools were not using the program at that time.

---

<sup>6</sup> Detailed statistical tables are presented in the Appendix B.

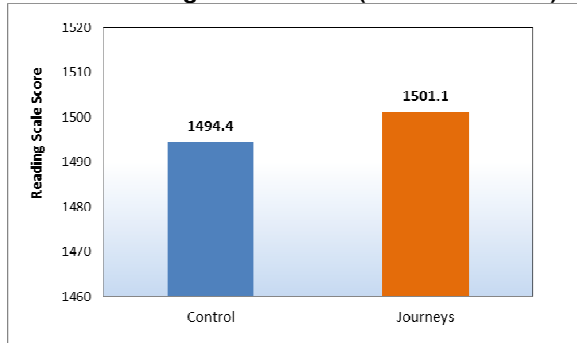
**Figure 1. Texas Journeys and Control Students' TAKS Reading Performance (3<sup>rd</sup> Grade – 2011)**



Covariates included: Gender, Ethnicity, Economic Disadvantage, Limited English Proficiency, and Special Education status (no test scores available in prior year when students were in 2<sup>nd</sup> grade).

Second, 4<sup>th</sup> graders who attended study schools during the 2011-12 school year (year two of Journeys implementation) were examined. Results showed that while Journeys students tended to outperform non-Journeys students on the new STAAR reading assessment, results were not statistically significant,  $p > .05$ , after controlling for demographics and 3<sup>rd</sup> grade (Spring 2011) reading performance.

**Figure 2. Texas Journeys and Control Students' TAKS Reading Performance (4<sup>th</sup> Grade – 2012)**

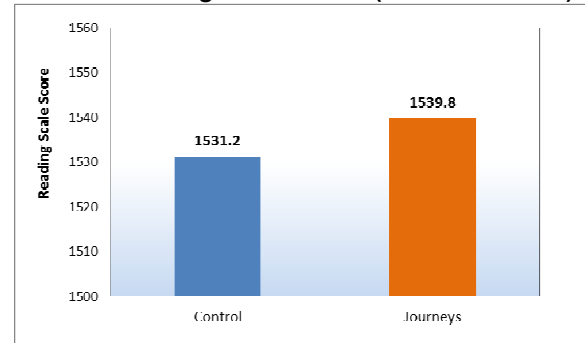


Covariates included: Gender, Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education status, and prior year reading scale score.

Similarly, 5<sup>th</sup> graders who attended study schools during the 2012-13 school year (year three of Journeys implementation) were examined to determine if significant differences were observed. Results showed that 5<sup>th</sup> grade Journeys students had significantly higher STAAR reading test

scores than non-Journeys students,  $p < .05$ , after controlling for demographics and 3<sup>rd</sup> grade (Spring 2011) and 4<sup>th</sup> grade (2012) reading performance. Thus, learning gains persisted even after controlling for prior positive program effects.

**Figure 3. Texas Journeys and Control Students' STAAR Reading Performance (5<sup>th</sup> Grade – 2013)**



Covariates included: Gender, Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education status, and prior years reading scale score.

---

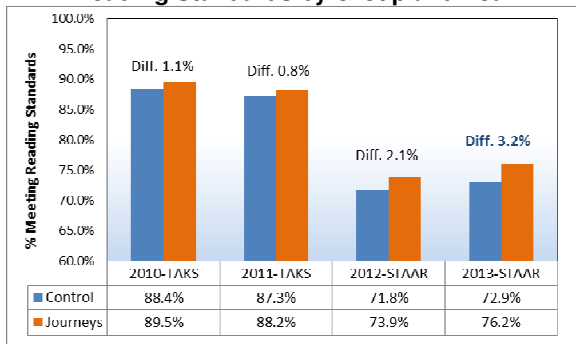
*Texas students using Journeys showed significantly higher STAAR and TAKS reading scores as compared to students using other reading programs for 3<sup>rd</sup> (2011) and 5<sup>th</sup> grade (2013).*

---

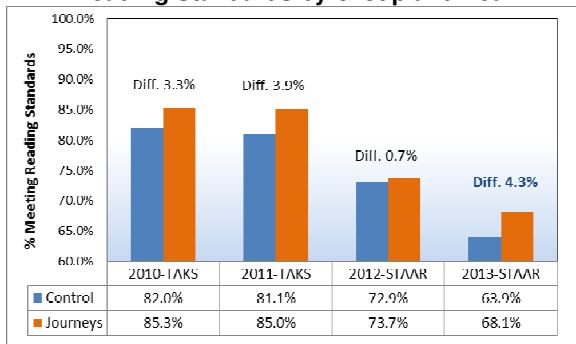
In addition, researchers examined whether or not differences observed at each grade had changed over the years, from baseline (2009-10 – no Journeys) to year one (2010-11), year two (2011-12), and year three (2012-13) of Journeys implementation. Given that teachers and students had more exposure to Journeys in year three, it was predicted that the most recent data would show the greatest differences. As a reminder, because the prior year TAKS data is not comparable to the STAAR, cross-sectional analysis (using logistic regression) was conducted separately for each grade level (3<sup>rd</sup> to 5<sup>th</sup>). This means that researchers examined different sets of students at each school year (see Figures 4-6 and Appendix B for detailed analyses). With this in mind, a pattern emerged in which Journeys students at each grade level were more likely to meet

reading standards as compared to non-Journeys students following three years of implementation (2013) as compared to one year (2011) or two years (2012) of implementation. That is, the difference between Journeys and non-Journeys students was greatest in 2013. Moreover, for 4<sup>th</sup> graders this pattern of greater performance following three years of Journeys implementation as compared to one or two years was marginally significant,  $p=.08$ .

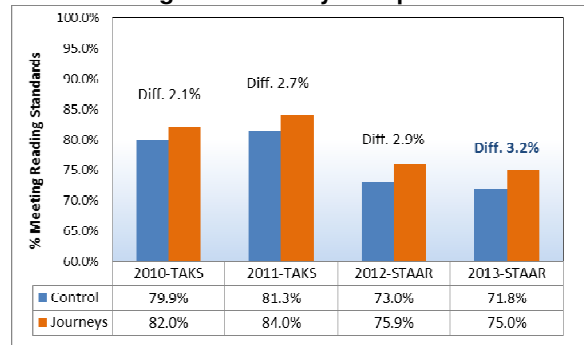
**Figure 4. Percent of 3rd Grade Students Meeting Reading Standards by Group and Year**



**Figure 5. Percent of 4<sup>th</sup> Grade Students Meeting Reading Standards by Group and Year**



**Figure 6. Percent of 5<sup>th</sup> Grade Students Meeting Reading Standards by Group and Year**



*Journeys students showed the highest levels of performance as compared to non-Journeys students (i.e., more students meeting Texas reading standards) following three years of using the program, after controlling for background variables. This pattern was statistically significant for 4<sup>th</sup> graders.*

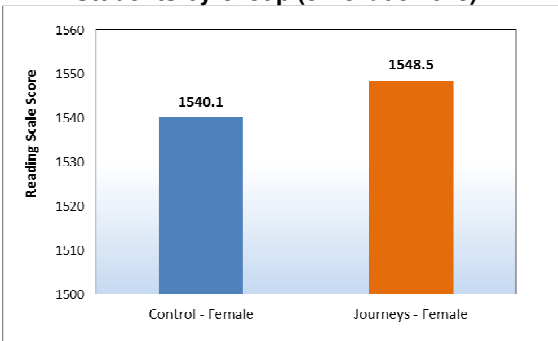
### Do findings vary across different subgroups of students?

Exploratory<sup>7</sup> analyses by students' gender, ethnicity (White, Hispanic, African American), and economically disadvantaged 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. ANCOVAs were run separately for each subgroup using Sample A. In particular, 5<sup>th</sup> grade STAAR reading performance was compared between Journeys and non-Journeys students within various subgroups (while controlling for prior year reading performance and other demographic variables).

<sup>7</sup> These analyses are exploratory because there have been very few studies that have examined subgroup effects relating to 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.

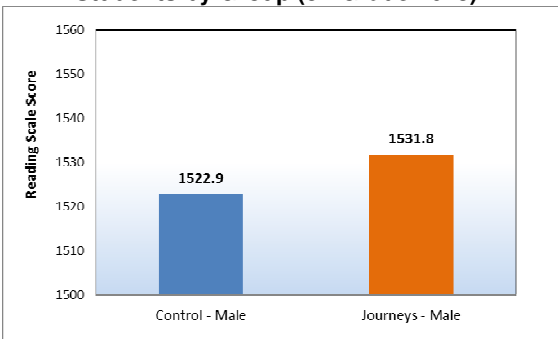
Results showed that Journeys students who were Hispanic or female performed significantly higher than non-Journeys students who were Hispanic or female,  $p < .05$ . As well, Journeys students who were economically disadvantaged outperformed non-Journeys students who were economically disadvantaged,  $p < .10$ . No significant differences were observed for male, White, or African American students; however it should be noted that Journeys students in these subgroups performed better than non-Journey students in the same subgroups, although the differences were not significant. Figures 7-12 show subgroup results by group.

**Figure 7. STAAR Reading Scale Score of Female Students by Group (5<sup>th</sup> Grade 2013)**



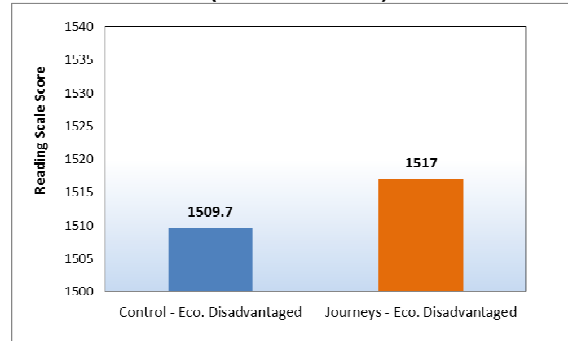
Covariates included: Prior year reading performance, ethnicity, economic disadvantage, Limited English Proficiency, Special Education

**Figure 8. STAAR Reading Scale Score of Male Students by Group (5<sup>th</sup> Grade 2013)**



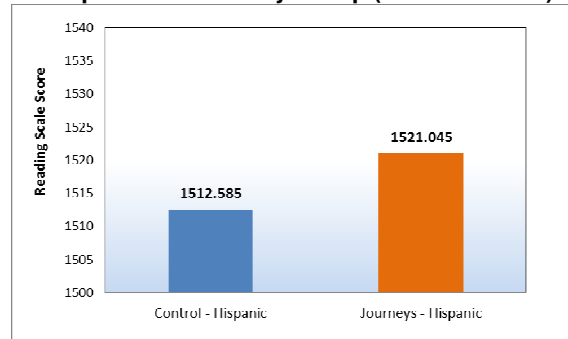
Covariates included: Prior year reading performance, ethnicity, economic disadvantage, Limited English Proficiency, Special Education

**Figure 9. STAAR Reading Scale Score of Economically Disadvantaged Students by Group (5<sup>th</sup> Grade 2013)**



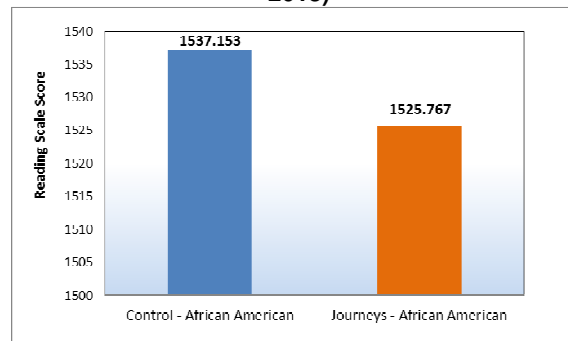
Covariates included: Prior year reading performance, gender, ethnicity, Limited English Proficiency, Special Education

**Figure 10. STAAR Reading Scale Score of Hispanic Students by Group (5<sup>th</sup> Grade 2013)**



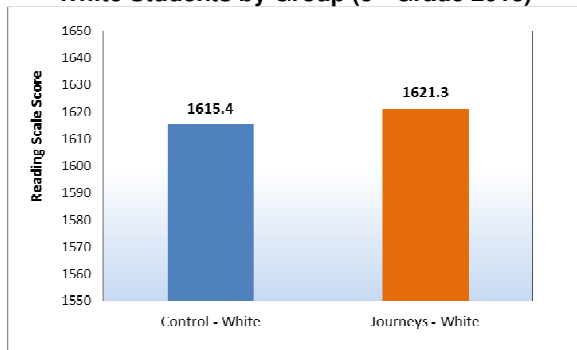
Covariates included: Prior year reading performance, gender, economic disadvantage, Limited English Proficiency, Special Education

**Figure 11. STAAR Reading Scale Score of African American Students by Group (5<sup>th</sup> Grade 2013)**



Covariates included: Prior year reading performance, gender, economic disadvantage, Limited English Proficiency, Special Education

**Figure 12. STAAR Reading Scale Score of White Students by Group (5<sup>th</sup> Grade 2013)**



Covariates included: Prior year reading performance, gender, economic disadvantage, Limited English Proficiency, Special Education

---

*Results showed that Hispanic, female, and economically disadvantaged students who used the Journeys program had higher STAAR reading scores as compared to students in these subgroups who used other elementary reading/language arts programs.*

---

## Conclusion

---

The Year 1 Report showed that the Journeys program is associated with significant increases in student reading performance. Specifically, a significant bump up in test scores was observed following implementation of the Journeys program; however, there were no differences in reading performance for Journeys students and students using other reading programs, as measured by the TAKS. Unlike in Year 1, the Year 2 Report demonstrated that students using Journeys *outperformed* students using other reading programs as measured by the new STAAR assessment and TAKS after controlling for baseline reading scores.

The present results in this Year 3 report reveal a similar picture. Journeys students outperformed non-Journeys students in STAAR and TAKS reading performance in 3<sup>rd</sup> and 5<sup>th</sup> grade. In addition, Journeys students were more likely than non-Journeys students to meet reading standards in 2013 (following three years of usage) as compared to 2012 (two years) and 2011 (one year), especially at the 4<sup>th</sup> grade. Among subgroups of students, Hispanic, female, and economically disadvantaged students who participated in Journeys had higher reading scores as compared to the same subgroup of students using other reading/language arts programs.

In summary, the results from Year 3 of this archival study using state assessment data provides additional support for a positive relationship between the Journeys program and elementary reading performance. The evidence suggests that the Journeys program improves elementary reading test scores, and the more students are exposed to the program, the more positive the effects observed.



## 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**

	<b>Group</b>	<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>t stat</b>
<b>School Enrollment</b>	Control	50	536.76	179.15	t(105)= -0.277, p=.78
	Journeys	57	547.72	223.65	
<b>Percent Whites</b>	Control	50	19.36	26.10	t(105)= -0.427, p=.67
	Journeys	57	21.57	27.15	
<b>Percent LEP</b>	Control	50	36.18	25.07	t(105)= 0.479, p=.63
	Journeys	57	33.69	28.33	
<b>Percent Economically Disadvantaged</b>	Control	50	78.36	22.37	t(105)= 0.869, p=.39
	Journeys	57	74.41	24.36	
<b>Percent At Risk</b>	Control	50	60.05	22.50	t(105)= 1.641, p=.10
	Journeys	57	52.11	26.93	
<b>Total Enrollment</b>	Control	50	540.48	187.97	t(105)= 0.068, p=.95
	Journeys	57	537.81	214.29	
<b>Percent African American</b>	Control	50	8.15	14.99	t(105)= -0.960, p=.34
	Journeys	57	11.32	18.62	
<b>Percent Hispanic</b>	Control	50	70.42	31.79	t(105)= 0.961, p=.34
	Journeys	57	64.25	34.25	
<b>Percent Native American</b>	Control	50	0.28	0.74	t(105)= 0.349, p=.73
	Journeys	57	0.24	0.33	
<b>Percent Asian/Pacific Islander</b>	Control	50	1.79	4.96	t(105)= -0.682, p=.50
	Journeys	57	2.63	7.28	
<b>Percent Special Ed.</b>	Control	50	7.25	2.57	t(105)= 1.384, p=.17
	Journeys	57	6.57	2.52	
<b>Percent Gifted</b>	Control	50	5.25	3.97	t(105)= -1.183, p=.24
	Journeys	57	6.43	5.22	
<b>Percent Mobility</b>	Control	50	17.57	6.51	t(105)= 1.205, p=.23
	Journeys	54	16.21	4.98	

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

## Appendix B: Statistical Analyses

### 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 B1. Results for Journeys vs Control Subgroups: Longitudinal Sample A**

#### **5<sup>th</sup> Graders (2013)**

Group	Mean (STAAR Reading SS)	Std. Error	N	Statistic	Sig.	Effect Size (eta-squared)
Control	1531.19	2.867	806	F Group (1, 1840)= 5.03	.03	.003
Journeys	1539.80	2.516	1043			

Covariates appearing in the model include: Gender, Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

#### **4<sup>th</sup> Graders (2012)**

Group	Mean (STAAR Reading SS)	Std. Error	N	Statistic	Sig.	Effect Size (eta-squared)
Control	1494.37	3.565	836	F Group (1, 1912)= 1.98	.160	.001
Journeys	1501.01	3.125	1084			

Covariates appearing in the model include: Gender, Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education, and 2011 3<sup>rd</sup> grade TAKS reading scale score

#### **3<sup>rd</sup> Graders (2011)<sup>8</sup>**

Group	Mean (TAKS Reading SS)	Std. Error	N	Statistic	Sig.	Effect Size (eta-squared)
Control	610.76	2.816	2284	F Group (1, 5051)= 4.67	.03	.001
Journeys	619.00	2.554	2774			

Covariates appearing in the model include: Gender, Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education

<sup>8</sup> Recall that the grade samples in the present sample vary somewhat from samples in the prior year's report in one important way: 1) the 3<sup>rd</sup> and 4<sup>th</sup> graders included in the current report must have been in the same school as 5<sup>th</sup> graders in 2013; this means that the sample is smaller than the students in the prior report who did not have this requirement.

**Table B2. Results for Journeys vs Control Subgroups: Cross-sectional Sample B**

**Grade 3**

Group	TEST YEAR	Mean* (percent meeting standards)	Std. Deviation	N	Statistic	Sig.
Control	2010	.88	.320	3037	F Group (1, 24943)= 0.091 F Year (3, 24943)= 321.095 F Group*Year (3, 24943)= 0.89	.764
	2011	.87	.332	2958		.000
	2012	.72	.450	2710		.443
	2013	.73	.444	2724		
Journeys	2010	.90	.306	3461		
	2011	.88	.323	3486		
	2012	.74	.439	3092		
	2013	.76	.426	3475		

\* Unadjusted statistics

**Grade 4**

Group	TEST YEAR	Mean* (percent meeting standards)	Std. Deviation	N	Statistic	Sig.
Control	2010	.82	.384	2871	F Group (1, 24626)= 4.884 F Year (3, 24626)= 266.329 F Group*Year (3, 24626)= 2.26	.027
	2011	.81	.392	2910		.000
	2012	.73	.444	2712		.079
	2013	.64	.480	2805		
Journeys	2010	.85	.354	3375		
	2011	.85	.357	3330		
	2012	.74	.440	3244		
	2013	.68	.466	3379		

\* Unadjusted statistics

**Grade 5**

Group	TEST YEAR	Mean* (percent meeting standards)	Std. Deviation	N	Statistic	Sig.
Control	2010	.80	.400	2855	F Group (1, 24642)= 1.569 F Year (3, 24642)= 81.502 F Group*Year (3, 24642)= 0.63	.210
	2011	.81	.390	2797		.000
	2012	.73	.444	2775		.598
	2013	.72	.450	2797		
Journeys	2010	.82	.384	3230		
	2011	.84	.366	3366		
	2012	.76	.428	3294		
	2013	.75	.433	3528		

\* Unadjusted statistics

Do such findings vary across different subgroups of students?

**Table B3. Results for Journeys vs Control Subgroups: Sample A**

**Females**

Group	Mean (Reading SS)	Std. Error	N	Statistic	Sig.
Control	1540.08	3.968	389	F Group (1, 882)= 2.49	.003
Journeys	1548.46	3.458	500		

Covariates appearing in the model include: Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

**Males**

Group	Mean (Reading SS)	Std. Error	N	Statistic	Sig.
Control	1522.94	4.112	417	F Group (1, 952)= 2.58	.109
Journeys	1531.80	3.593	543		

Covariates appearing in the model include: Ethnicity, Economic Disadvantage, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

**Economically Disadvantaged**

Group	Mean (Reading SS)	Std. Error	N	Statistic	Sig.
Control	1509.71	2.976	681	F Group (1, 1452)= 3.19	.074
Journeys	1517.02	2.783	778		

Covariates appearing in the model include: Gender, Ethnicity, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

**Hispanics**

Group	Mean (Reading SS)	Std. Error	N	Statistic	Sig.
Control	1512.59	3.003	659	F Group (1, 1371)= 4.12	.043
Journeys	1521.05	2.874	719		

Covariates appearing in the model include: Gender, Economic Disadvantage, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

## African Americans

Group	Mean (Reading SS)	Std. Error	N	Statistic	Sig.
Control	1537.15	16.430	25	F Group (1, 134)= 0.39	.534
Journeys	1525.77	7.481	115		

Covariates appearing in the model include: Gender, Economic Disadvantage, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

## Whites

Group	Mean (Reading SS)	Std. Error	N	Statistic	Sig.
Control	1602.49	9.483	106	F Group (1, 272)= 0.67	.796
Journeys	1605.64	7.367	173		

Covariates appearing in the model include: Gender, Economic Disadvantage, Limited English Proficiency, Special Education, 2011 3<sup>rd</sup>-grade TAKS reading scale score, and 2012 4<sup>th</sup>-grade STAAR reading scale score

## **PRES Associates**

PO Box 10730  
Jackson, WY 83002  
Phone: 307-733-3255, Fax: 307-733-2017  
[info@presassociates.com](mailto:info@presassociates.com)