A Study of Sadlier's *From Phonics to Reading Program*: Final Report

August 30, 2023

magnolia consulting

cultivating learning and positive change

www.magnoliaconsulting.org

Certified B Corporation

This company is part of the global movement for an inclusive, equitable, and regenerative economic system.

ACKNOWLEDGEMENTS

Magnolia Consulting appreciates the opportunity to conduct a study of *From Phonics to Reading*. We want to thank Sadlier for their collaborative approach to the study. We also want to thank the district for their collaboration in this study, in particular the district representatives who provided the study data. Last, we want to thank the Magnolia team members who supported this work.

The authors,

Jennifer A. Gruber, Ph.D.
Researcher and Evaluator

Lisa Shannon, Ph.D.
Director of Research and Evaluation

Holly Stokely, M.S.
Research Assistant

Magnolia Consulting, LLC
5135 Blenheim Rd.
Charlottesville, VA 22902
(ph) 855.984.5540 (toll free)
http://www.magnoliaconsulting.org

Preferred citation: Gruber, J. A., Shannon, L., & Stokely, H. (2023). *A Study of Sadlier's From Phonics to Reading Program: Final Report.* Magnolia Consulting, LLC.

EXECUTIVE SUMMARY

RESEARCH METHODS

Treatment-only pretest-posttest study design

347 K–2 students at two elementary schools that used *From Phonics to Reading*

Secondary analysis of Star Reading assessment data using multilevel modeling

William H. Sadlier, Inc. (Sadlier) developed *From Phonics to Reading*, a supplemental foundational reading skills program, to provide students with a strong foundation in early literacy. Sadlier contracted with Magnolia Consulting, LLC, an external consulting firm specializing in educational research and evaluation, to conduct a study of *From Phonics to Reading* among kindergarten, first-grade, and second-grade students at two elementary schools using data from the 2021/2022 school year.

Study findings indicate that as a group, students demonstrated improvements in reading scores from fall to spring. To understand how these improvements might differ from typical student growth within a school year on this assessment, researchers visually examined how students' mean reading scores compared with the Star Reading assessment performance criteria. These findings suggest that within each grade, the average reading performance of students in this study improved by a grade level or more from fall to spring. A summary of the three research questions and key findings from this study are presented below.

Did students who used *From Phonics to Reading* demonstrate improvements in reading performance from fall 2021 to spring 2022?

Key Findings

⇒ On average, kindergarten through second-grade students demonstrated a statistically significant improvement in reading performance from fall to spring.

⇒ On average, students within each grade demonstrated a statistically significant improvement in reading performance from fall to spring.

How did participating students' average fall 2021 and spring 2022 reading scores compare with assessment performance criteria?

Key Findings

⇒ In the fall, kindergarten students' mean Star Reading score was in the 25th–40th percentile and reflected a typical score for an average kindergarten student in the beginning of the school year. In the spring, their mean score was in the 50th–75th percentile and reflected a typical score for an average first-grade student in the fourth month of the school year.

⇒ In the fall, first-grade students' mean Star Reading score was in the 25th–40th percentile and reflected a typical score for an average kindergarten student in the fifth month of the school year. In the spring, their mean score was in the 40th–50th percentile and reflected a typical score for an average first-grade student in the fifth month of the school year.

⇒ In the fall, second-grade students' mean Star Reading score was in the 20th–25th percentile and reflected a typical score for an average first-grade student in the fifth month of the school year. In the spring, their mean score was in the 40th–50th percentile and reflected a typical score for an average second-grade student in the ninth month of the school year.

Did key subgroups of participating students demonstrate improvements in reading performance from fall 2021 to spring 2022?

Key Findings

⇒ On average, all subgroups examined (i.e., male students, female students, and students who qualified for Free or Reduced-Price Lunch) demonstrated a statistically significant improvement in reading performance from fall to spring.

It is important to note that because this study utilized a treatment-only, pretest-posttest design, apparent gains in student reading performance could be due to factors other than the use of From Phonics to Reading.

TABLE OF CONTENTS

INTRODUCTION

William H. Sadlier, Inc. (Sadlier) understands the importance of examining the efficacy of its programs. Therefore, Sadlier contracted with Magnolia Consulting, LLC (Magnolia), an external independent consulting firm specializing in educational research and evaluation, to conduct a study of their supplemental reading program, *From Phonics to Reading*. The study examined reading performance among kindergarten, first-grade, and second-grade students who used *From Phonics to Reading* using data from the 2021/2022 school year.

This report includes

- an overview of the Sadlier *From Phonics to Reading* program,
- the study design and methods,
- results related to student reading performance, and
- a discussion of the study findings.

PROGRAM DESCRIPTION

Reading achievement is an important predictor of academic outcomes. Students who are not reading proficiently by third grade are less likely to graduate high school compared with students who are reading proficiently (Fiester, 2010; Hernandez, 2011). Additionally, reading skills are associated with other important social and academic outcomes, such as classroom behavior and math achievement (Grimm, 2008; McIntosh et al., 2008). Early literacy skills (e.g., alphabet knowledge, phonological awareness) predict later reading achievement, even after accounting for demographic factors often associated with children's academic performance (e.g., socioeconomic status; Lonigan & Shanahan, 2009; Rabiner et al., 2016).

Given the importance of early literacy skills for later reading achievement, Sadlier developed *From Phonics to Reading* to provide students with a strong foundation in early literacy. *From Phonics to Reading* is a supplemental foundational skills program characterized by readiness skills instruction, a systematic scope and sequence, blending, dictation, high-frequency word instruction, reading decodable texts, and word awareness. The program covers the five pillars of reading: phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Reading Panel, 2000). The program follows a systematic scope and sequence, building from simple to complex skills and includes cycles of learning, reviewing, and practicing different literacy skills throughout the program (Banse, 2022), which is well supported by research (Bleses et al., 2018; Ortlieb & McDowell, 2016; Wonder-McDowell et al., 2011).

The *From Phonics to Reading* program includes student resources (e.g., workbooks), instructional guides for Tier I and Tier II reading instruction (including small group and whole group activities), formative summative, and cumulative assessments, teacher professional development, and an online learning platform with additional digital resources for both students and teachers.

STUDY PURPOSE, DESIGN, AND METHODS

This study examined students' reading performance (as reflected in their performance on a standardized reading assessment) in two schools that used *From Phonics to Reading* during the 2021/2022 school year.

The study addressed the following research questions:

1. Did students who used *From Phonics to Reading* during the 2021/2022 school year demonstrate a statistically significant improvement in reading performance from fall 2021 to spring 2022? If so, what was the magnitude of this improvement?
2. How did participating students' average fall and spring reading scores compare with the fall and spring scores of the assessment's norm group or with assessment performance criteria?
3. Did key subgroups of students who used *From Phonics to Reading* during the 2021/2022 school year demonstrate statistically significant improvements in reading from fall 2021 to spring 2022? If so, what was the magnitude of each improvement?

Study Design

Researchers used a treatment-only pretest-posttest design to conduct the study of *From Phonics to Reading*. Specifically, researchers conducted secondary data analysis of reading scores for students at two schools that used *From Phonics to Reading* during the 2021/2022 school year. Because *From Phonics to Reading* is designed for delivery at the classroom level, the study's sample selection occurred at the classroom level and did not require entire schools to participate in the study. Thus, the study's analytic sample was composed of kindergarten, first-grade, and second-grade teachers and students from classrooms in two schools that used *From Phonics to Reading* during the 2021/2022 school year.

This treatment-only study design had several advantages. It allowed teachers who were already using *From Phonics to Reading* to participate. It also allowed researchers to rely on existing data to conduct the study. A limitation of the study design is that the study did not include a comparison group. As such, any changes in reading performance from fall to spring could have been due to various factors, such as the use of *From Phonics to Reading,* the use of other curricula, teachers' instructional practices, maturation, or other factors.

Measures

For this study, researchers collected school, teacher, and student data. School data provided information about the study context. Teacher data included demographic data. Student data included demographic data, roster information (e.g., which classroom they were in, whether they left the school during the school year), and fall and spring Star Reading assessment scores.

Due to the grades included in the study (kindergarten through second grade), the data included student scores from multiple assessments: Star Early Literacy and Star Reading. The Star Early Literacy assessment is intended for students in Pre-K through third grade who may have limited reading ability (Renaissance, 2022a). The Star Reading assessment is intended for students in kindergarten through 12th grade (Renaissance, 2022b). Renaissance, the developer of the Star

Reading assessments, developed Unified Scale Scores for both assessments that enable the comparison of scores across different grades and assessments. These scores are composite scores and represent a student's reading performance across all the domains covered in the assessment. The Unified Scale Score from both assessments was the outcome of interest in this study.

Procedures

The study involved secondary data analysis. Researchers worked with district representatives to collect and prepare the data for the study. District representatives collected and securely shared the data with researchers in an appropriate electronic format. Researchers also followed up with the representatives with any questions related to the data or participants. Researchers received the data for the study on May 15, 2023, and followed specific procedures for receiving, managing, and preparing the data (Appendix A).

Analytic Approach

Multilevel Models. For the first and third research questions, researchers estimated multilevel models to determine whether students in the analytic sample who used *From Phonics to Reading* during the 2021/2022 school year demonstrated a statistically significant gain in their reading scores (based on Unified Scale Scores). Multilevel models are appropriate for analyzing data collected from students grouped within classrooms (i.e., nested) because the models account for the likelihood that students within the same classroom share experiences associated with the classroom that could affect their responses to curricula (Borman et al., 2005; Raudenbush & Bryk, 2002). For multilevel analyses, researchers considered findings statistically significant at an alpha level of .05. Thus, findings were considered statistically significant if the probability of the finding occurring by chance was less than 5%.

After estimating multilevel models, researchers followed What Works Clearinghouse (WWC, 2022) guidelines to calculate standardized effect sizes. These standardized effect sizes represent the magnitude of the difference between fall and spring scores, accounting for sample size and nesting. When interpreting the effect sizes presented in this study, it is important to consider that they correspond to fall-to-spring performance gains rather than program impact. Over the course of a school year, gains in reading performance are typically expected as students develop, and they can be attributed to various factors. Furthermore, effect sizes corresponding to learning gains can vary based on numerous factors including grade, content area, assessment, and demographic characteristics, among others (Scammacca, Fall, & Roberts, 2015).

Norms Comparison. For the second research question, researchers visually examined how the reading performance of students in the analytic sample who used *From Phonics to Reading* during the 2021/2022 school year compared with that of a nationally representative sample of students. Specifically, researchers calculated students' raw, unadjusted mean fall 2021 and spring 2022 Star Reading Unified Scale Scores and visually examined how they compared with the assessments' norm group scores using the technical manual for the Star Reading assessment (Renaissance, 2022b).[1] Specifically, researchers examined percentile ranks and

[1] The current Star Reading norms (dated 2022) are based on data collected during the 2018/2019 school year from students in grades Pre-K–3. The Star Reading norming sample included students from over 24,295 schools across the United States.

grade equivalent (GE) scores, which indicate the grade level at which an average student would typically earn a given score. These findings provide context regarding how the reading performance of students in this study compares with the reading performance of typical students in kindergarten, first, and second grades.

Study Setting

The study included two elementary schools from the same district. Researchers examined the characteristics of each school and the district using data from the National Center for Educational Statistics (NCES, n.d.).[2] The district is located in a remote area in the Northeastern United States. Both schools were public schools, enrolled Pre-K through third-grade students, and were eligible for Title I (Table 1).

Table 1. School Characteristics

School Characteristics	School 1	School 2
Type	Public School	Public School
Title I Eligible	Yes	Yes
Grades Enrolled	Pre-K–Third Grade	Pre-K–Third Grade
Number of Students Enrolled	300	335

Participants

This section includes an overview of study participants in the analytic sample. Additional details about the steps researchers took to determine the analytic sample are included in Appendix B.

Teacher Participants

The final teacher analytic sample (i.e., teachers whose students were included in the analyses) included teachers who taught kindergarten, first, or second grade at the two schools during the 2021/2022 school year. This included 28 teachers from the two schools. Twelve teachers taught at School 1 and 16 teachers taught at School 2. Teachers reported an average of 23.00 years of teaching experience, with an average of 20.14 years at their current school district (Table 2). One teacher held a bachelor's degree and 27 held a master's degree.

Table 2. Teacher Characteristics (n = 28)

Characteristics			
Teaching Experience	***n***	**Mean**	***SD***
Total years teaching	28	23.00	8.94
Years at current school district	28	20.14	8.61
Highest Degree	***n***	**Percent**	
Bachelor's degree	1	3.57%	
Master's degree	27	96.43%	

[2] These data are from the 2021/2022 school year.

Student Participants

The final student analytic sample included students who were enrolled in kindergarten, first-grade, and second-grade classrooms at the two schools for the 2021/2022 school year and who had complete data (i.e., fall and spring assessments, classroom roster information).[3] This included 347 students (157 at School 1 and 190 at School 2).

Students were divided almost evenly across kindergarten through second grades, with slightly more students in kindergarten (36.9%), followed by second grade (32.0%) and first grade (31.1%; Table 3). There were slightly more males (54.5%) than females (45.5%). Almost all students identified as not Hispanic or Latino (96.0%). The majority of students were White (85.0%), followed by multiracial (6.9%), Black/African American (4.3%), Asian (2.0%), American Indian/Alaska Native (1.2%), and Native Hawaiian/Other Pacific Islander (0.6%). Many students qualified for Free or Reduced-Price Lunch (60.8%) and the remaining 39.2% of students were not. Few students were English Learners (1.2%), received special education services (15.3%), or qualified for Section 504 (0.9%).[4]

[3] Researchers received data for 381 students, and removed 34 students (8.92%) from the analytic sample due to missing data.

[4] Section 504 is a federal law that provides funding and programs to public school districts to support students with disabilities: Protecting Students With Disabilities (ed.gov)

Table 3. Student Characteristics (n = 347)

Characteristics	*n*	Percent
Grade		
Kindergarten	128	36.9%
First	108	31.1%
Second	111	32.0%
Gender		
Male	189	54.5%
Female	158	45.5%
Ethnicity		
Hispanic or Latino	14	4.0%
Not Hispanic or Latino	333	96.0%
Race		
American Indian/Alaska Native	4	1.2%
Asian	7	2.0%
Black/African American	15	4.3%
Multiracial	24	6.9%
Native Hawaiian/Other Pacific Islander	2	0.6%
White	295	85.0%
Free or Reduced-Price Lunch Status		
No	136	39.2%
Yes	211	60.8%
English Learner Status		
No	343	98.8%
Yes	4	1.2%
Special Education Status		
No	294	84.7%
Yes	53	15.3%
504 Status		
No	344	99.1%
Yes	3	0.9%

STUDY FINDINGS

This section of the report begins with findings related to improvements in reading scores for students from fall 2021 to spring 2022 (as measured by Star Reading assessments). Next, it includes descriptions of the reading performance of students in the study compared with nationally representative samples of students in kindergarten, first grade, and second grade. Finally, it includes findings related to improvements in reading scores for subgroups of students (male students, female students, and students who qualified for Free or Reduced-Price Lunch) from fall 2021 to spring 2022.

Reading Performance Among Students Who Used *From Phonics to Reading* During the 2021/2022 School Year

Did students who used *From Phonics to Reading* during the 2021/2022 school year demonstrate a statistically significant improvement in reading performance from fall 2021 to spring 2022? If so, what was the magnitude of this improvement?

Key Findings

- ⇒ On average, kindergarten through second-grade students demonstrated a statistically significant improvement in reading performance from fall to spring. The mean gain in reading scores corresponded to an effect size of 1.12.
- ⇒ On average, students within each grade demonstrated a statistically significant improvement in reading performance from fall to spring. Grade-specific mean gains in reading scores corresponded to effect sizes ranging from 1.15 to 1.37.

As described previously, researchers estimated multilevel models to determine whether students in the analytic sample who used *From Phonics to Reading* during the 2021/2022 school year demonstrated a statistically significant gain in their reading scores (based on their Star Reading assessment). Specifically, researchers estimated multilevel models for the aggregate sample (i.e., students across grades) and within each grade.[5] Appendix C includes the raw, unadjusted means and standard deviations overall and by grade.[6]

When interpreting the effect sizes presented in this study, it is important to consider that they correspond to fall-to-spring reading performance improvement. Learning gains are typically expected over the course of a school year and their corresponding effect sizes can vary based on numerous factors (Scammacca et al., 2015). Thus, it is important to consider how students in this study performed relative to the Star Reading assessment's norm group. The next research question and associated findings provide this additional context.

[5] These models accounted for the nesting of students in classrooms. The difference in Star Reading Unified Scale Scores from fall 2021 to spring 2022 was the outcome of interest (i.e., gain scores).
[6] The raw, unadjusted means did not account for the data's nested structure or for any covariates.

Across grades, average reading performance improved from fall 2021 to spring 2022

Across grades, students gained an average of 120.55 points on the Unified Scale from the fall 2021 to the spring 2022 administration of the Star Reading assessment (Table 4). This gain was statistically significantly different from zero, meaning it was unlikely to have occurred by chance. The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.12.

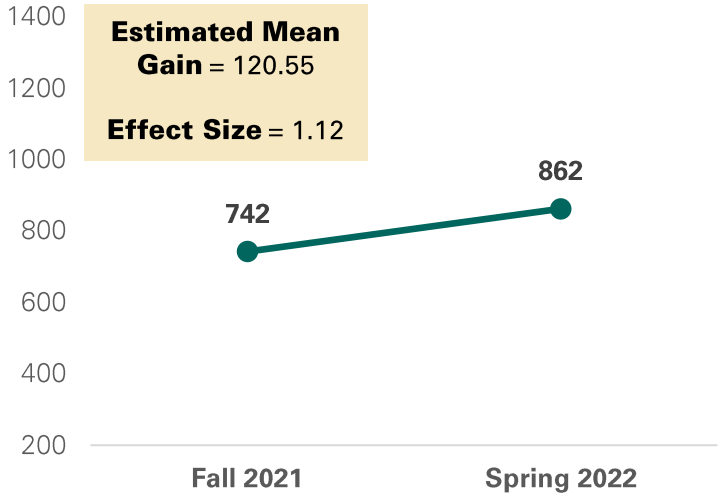
Table 4. Multilevel Model Results for the Aggregate Sample

	Coefficient	Standard Error	*t*-value	Approx. *df*	*p*-value	Effect Size
Star Reading Gain	120.55	7.07	17.06	26.81	<.001*	1.12

* Statistically significant at the .05 level.

As seen in Figure 1, after accounting for nesting, the estimated mean Unified Scale Scores of students in the analytic sample increased from fall to spring during the 2021/2022 school year.[7]

Figure 1. Estimated Unified Scale Score Means for the Analytic Sample



Within each grade, average reading performance improved from fall 2021 to spring 2022

Kindergarten students gained an average of 144.78 points on the Unified Scale from fall 2021 to spring 2022 (Table 5). This gain was statistically significantly different from zero, meaning it was unlikely to have occurred by chance. The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.37.

First-grade students gained an average of 109.65 points on the Unified Scale from the fall 2021 to the spring 2022 administration of the Star Reading assessment (Table 5). This gain was statistically significantly different from zero, meaning it was unlikely to have occurred by chance. The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.36.

[7] The range of possible Unified Scale Scores is 200-1100 for Star Early Literacy and 600-1400 for Star Reading.

Second-grade students gained an average of 104.01 points on the Unified Scale from fall 2021 to spring 2022 (Table 5). This gain was statistically significantly different from zero, meaning it was unlikely to have occurred by chance.[8] The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.15.

Table 5. Multilevel Model Results by Grade

	Coefficient	Standard Error	*t*-value	Approx. *df*	*p*-value	Effect Size
Kindergarten Star Reading Gain	144.78	15.41	9.40	8.27	<.001*	1.37
First Grade Star Reading Gain	109.65	7.68	14.28	7.17	<.001*	1.36
Second Grade Star Reading Gain	104.01	5.72	18.18	110.00	<.001*	1.15

* Statistically significant at the .05 level.

As seen in Figures 2–4, after accounting for nesting, the estimated mean Unified Scale Scores of students across grades increased from fall to spring during the 2021/2022 school year.

Figure 2. Estimated Unified Scale Score Means for Kindergarten Students

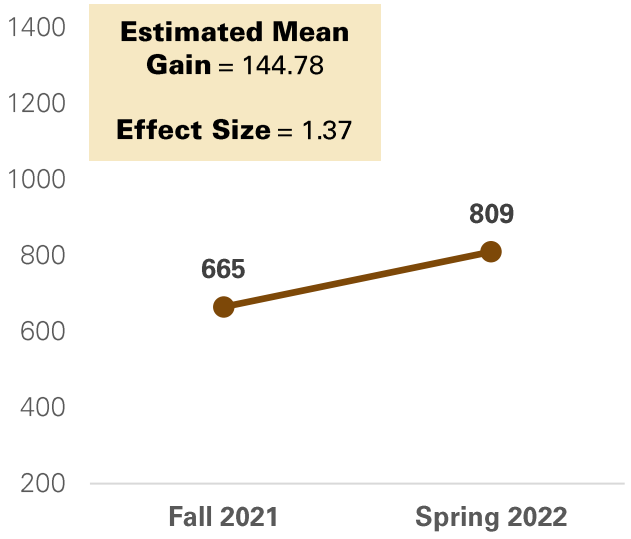
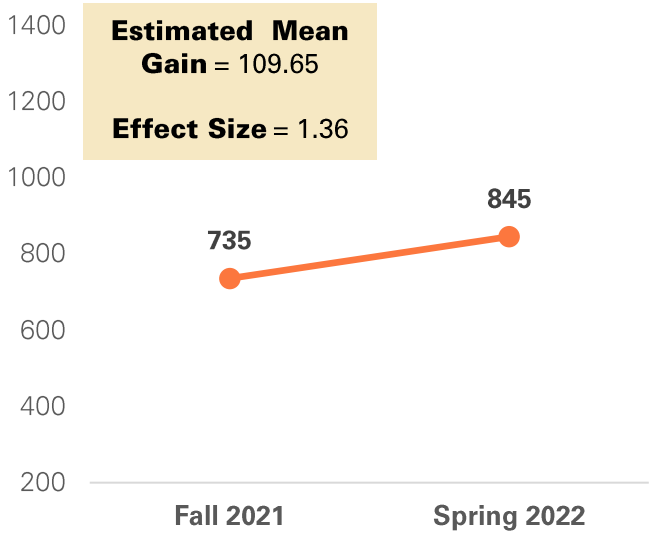
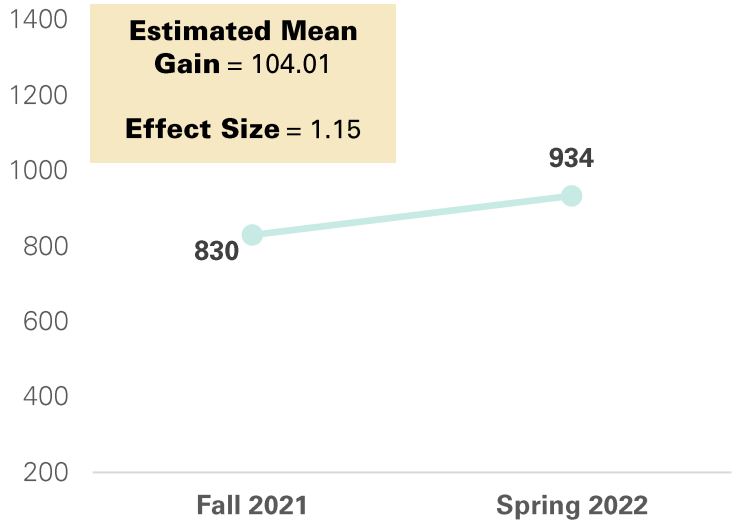


Figure 3. Estimated Unified Scale Score Means for First-Grade Students



[8] The intraclass correlation (i.e., the extent to which the reading performance of students in the same classroom scores is related to each other) for the multilevel model for second-grade students was 0. This impacts the estimated degrees of freedom (df) in the model.

Figure 4. Estimated Unified Scale Score Means for Second-Grade Students



Reading Performance Among Students Who Used *From Phonics to Reading* During the 2021/2022 School Year Compared With Nationally Representative Samples

How did participating students' average fall and spring reading scores compare with the fall and spring scores of the assessment's norm group or with assessment performance criteria?

Key Findings

- ⇒ In fall 2021, kindergarten students' mean Star Reading Unified Scale Score was between the 25th and 40th percentile and corresponded to a grade equivalent of 0.0. In spring 2022, their mean score was between the 50th and 75th percentile and corresponded to a grade equivalent of 1.4.
- ⇒ In fall 2021, first-grade students' mean Star Reading Unified Scale Score was between the 25th and 40th percentile and corresponded to a grade equivalent of 0.5. In spring 2022, their mean score was between the 40th and 50th percentile and corresponded to a grade equivalent of 1.5.
- ⇒ In fall 2021, second-grade students' mean Star Reading Unified Scale Score was between the 20th and 25th percentile and corresponded to a grade equivalent of 1.5. In spring 2022, their mean score was between the 40th and 50th percentile and corresponded to a grade equivalent of 2.9.

Researchers visually examined how the reading performance of students in the analytic sample who used *From Phonics to Reading* during the 2021/2022 school year compared with that of a nationally representative sample of students.

Within each grade level, researchers determined how students' mean fall 2021 and spring 2022 reading scores corresponded to national percentile ranks. Renaissance provides percentile rank

ranges corresponding to reading scores within each grade.[9] Then, within each grade level, researchers determined the extent to which students' mean fall and spring scores corresponded to the Star Reading GE scores. GE scores are based on a 10-month school year and indicate the grade level at which an average student would be expected to earn a given score. The digit to the left of the decimal specifies the school year, and the digit to the right specifies the month in the school year.[10]

Within each grade, percentile rank ranges corresponding to average reading performance increased from fall 2021 to spring 2022

Overall, visual examination of the percentile rank ranges corresponding to the mean Unified Scale score within each grade suggests they increased from fall 2021 to spring 2022 (Table 6):

- The mean kindergarten Unified Scale Score (*M* = 666.51, *SD* = 71.34) was between the 25th and 40th percentile in the fall, and the mean score (*M* = 816.58, *SD* = 101.15) was between the 50th and 75th percentile in the spring.
- The mean first-grade Unified Scale Score (*M* = 735.69, *SD* = 65.80) was between the 25th and 40th percentile in the fall, and the mean score (*M* = 845.66, *SD* = 89.61) was between the 40th and 50th percentile in the spring.
- The mean second-grade Unified Scale Score (*M* = 829.81, *SD* = 105.93) was between the 20th and 25th percentile in the fall, and the mean score (*M* = 933.82, *SD* = 103.57) was between the 40th and 50th percentile in the spring.

Table 6. Star Reading Assessment Percentiles by Grade

Grade	Fall 2021	Spring 2022
Kindergarten	25th–40th percentile	50th–75th percentile
First grade	25th–40th percentile	40th–50th percentile
Second grade	20th–25th percentile	40th–50th percentile

Within each grade, GE scores corresponding to average reading performance increased from fall 2021 to spring 2022 by a grade level or more

Overall, the GE scores associated with the mean Unified Scale Score for each grade increased from fall 2021 to spring 2022 by a grade level or more.

K

In fall 2021, the GE corresponding to kindergarten students' mean score (GE = 0.0) reflected a typical score for an average kindergarten student in the beginning of the school year. By spring 2022, the GE corresponding to their mean score (GE = 1.4) reflected a typical score for an average first-grade student in the fourth month of the school year. **This equates to an increase of over one grade level (1.4)**.

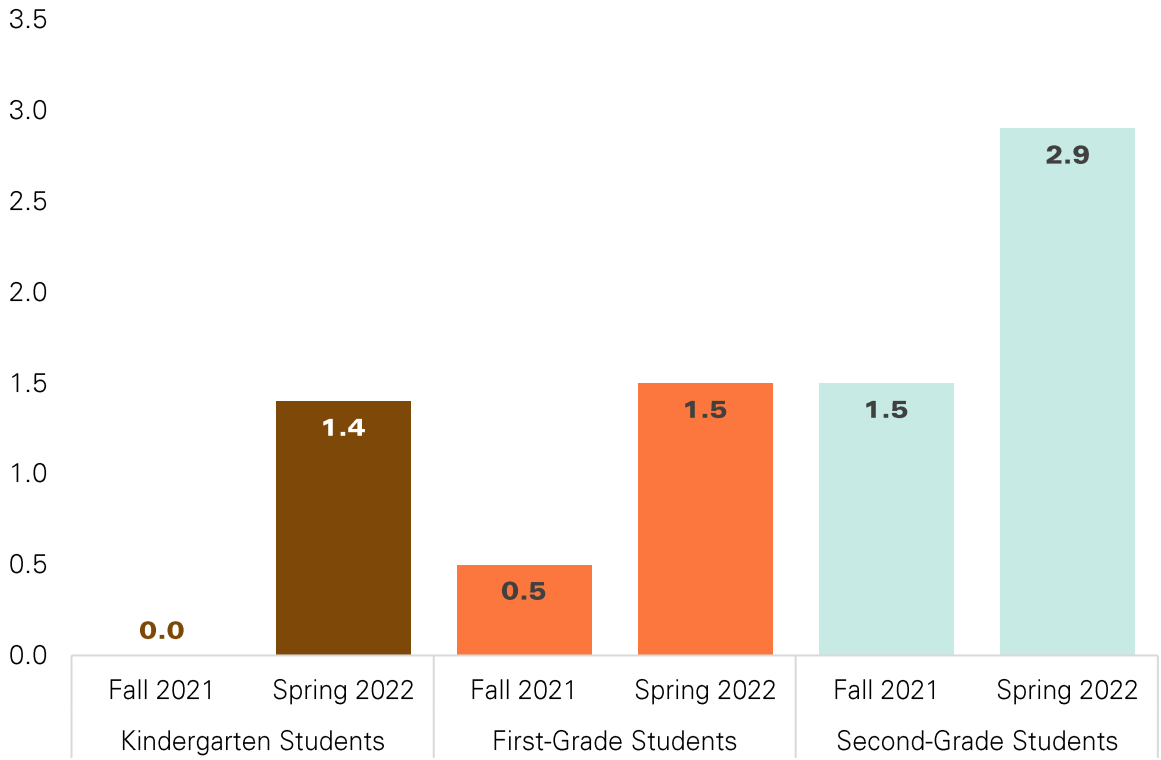
[9] Unified Scale Score to percentile rank range conversion table: SRUnifiedBenchmarksCutScores.pdf (widen.net)
[10] A GE score of 3.0 represents how an average student would be expected to perform on Star Reading at the start of third grade, and a GE score of 3.9 represents how an average student would be expected to perform on Star Reading in the ninth month of third grade.

1 In fall 2021, the GE corresponding to first-grade students' mean score (GE = 0.5) reflected a typical score for an average kindergarten student in the fifth month of the school year. By spring 2022, the GE corresponding to their mean score (GE = 1.5) reflected a typical score for an average first-grade student in the fifth month of the school year. **This equates to an increase of one grade level (1.0)**.

2 In fall 2021, the GE corresponding to second-grade students' mean score (GE = 1.5) reflected a typical score for an average first-grade student in the fifth month of the school year. By spring 2022, the GE corresponding to their mean score (GE = 2.9) reflected a typical score for an average second-grade student in the ninth month of the school year. **This equates to an increase of over one grade level (1.4)**.

As seen in Figure 5, the GE score associated with the mean Unified Scale Score for kindergarten students was below their grade-level benchmark in fall 2021, and above their grade level benchmark in spring 2022. The GE score associated with the mean Unified Scale Score for second-grade students was below their grade-level benchmark in fall 2021, and at their grade level benchmark in spring 2022. This indicates that—on average—these students began the school year with reading performance below their grade level, and ended the school year with reading performance at or above their grade level. For first-grade students, the GE score associated with their mean Unified Scale Score was below the grade-level benchmark in fall 2021 and spring 2022, but increased by one grade level.

Figure 5. GE Scores by Grade



Reading Performance Among Student Subgroups Who Used *From Phonics to Reading* During the 2021/2022 School Year

Did key subgroups of students who used *From Phonics to Reading* during the 2021/22 school year demonstrate statistically significant improvements in reading from fall 2021 to spring 2022? If so, what was the magnitude of each average improvement?

Key Findings

- ⇒ On average, male students demonstrated a statistically significant improvement in reading performance from fall to spring. The mean gain in reading scores corresponded to an effect size of 1.18.
- ⇒ On average, female students demonstrated a statistically significant improvement in reading performance from fall to spring. The mean gain in reading scores corresponded to an effect size of 1.08.
- ⇒ On average, students who qualified for Free or Reduced-Price Lunch demonstrated a statistically significant improvement in reading performance from fall to spring. The mean gain in reading scores corresponded to an effect size of 1.14.

Researchers estimated multilevel models to determine whether key subgroups of students in the analytic sample who used *From Phonics to Reading* during the 2021/2022 school year demonstrated a statistically significant gain in their reading scores (based on their Star Reading assessment). Specifically, researchers estimated multilevel models for the following student subgroups: male students, female students, and students who qualified for Free or Reduced-Price Lunch.[11] Appendix C includes the raw, unadjusted means and standard deviations overall and by student subgroup.[12]

Among male students, reading performance improved from fall 2021 to spring 2022

Male students gained an average of 126.04 points on the Unified Scale from the fall 2021 to the spring 2022 administration of the Star Reading assessment (Table 7; Figure 6). This gain was statistically significantly different from zero, meaning that it was unlikely to have occurred by chance. The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.18.

Table 7. Multilevel Model Results for Male Students

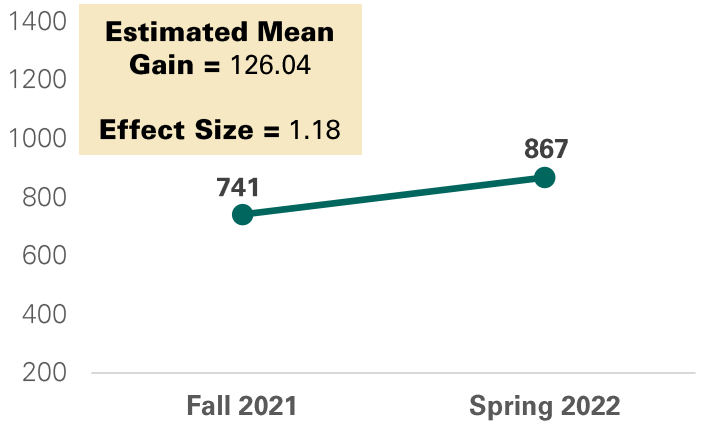
Outcome	Coefficient	Standard Error	*t*-value	Approx. *df*	*p*-value	Effect Size
Star Reading Gain	126.04	7.80	16.17	30.04	<.001*	1.18

* Statistically significant at the .05 level.

[11] These models accounted for the nesting of students in classrooms. The difference in Star Reading Unified Scale Scores from fall 2021 to spring 2022 was the outcome of interest (i.e., gain scores).
[12] The raw, unadjusted means did not account for the data's nested structure or for any covariates.

As seen in Figure 6, after accounting for nesting, the estimated mean Unified Scale Scores of male students increased from fall to spring during the 2021/2022 school year.

Figure 6. Estimated Unified Scale Score Means for Male Students



Among female students, reading performance improved from fall 2021 to spring 2022

Female students gained an average of 116.06 points on the Unified Scale from the fall 2021 to the spring 2022 administration of the Star Reading assessment (Table 8; Figure 7). This gain was statistically significantly different from zero, meaning that it was unlikely to have occurred by chance. The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.08.

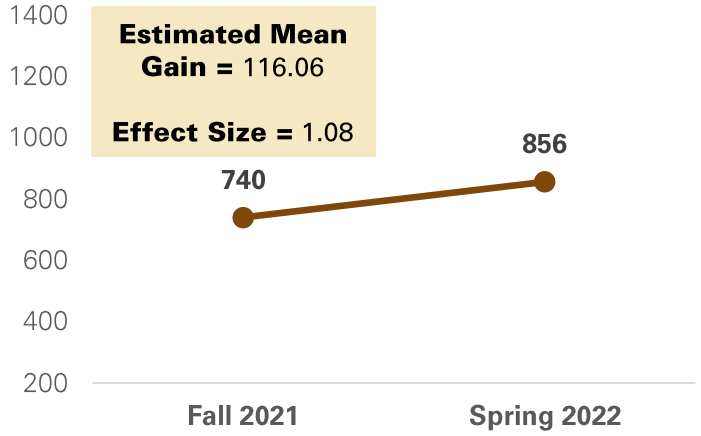
Table 8. Multilevel Model Results for Female Students

Outcome	Coefficient	Standard Error	*t*-value	Approx. *df*	*p*-value	Effect Size
Star Reading Gain	116.06	7.14	23.74	23.74	<.001*	1.08

* Statistically significant at the .05 level.

As seen in Figure 7, after accounting for nesting, the estimated mean Unified Scale Scores of female students increased from fall to spring during the 2021/2022 school year.

Figure 7. Estimated Unified Scale Score Means for Female Students



Among students who qualified for Free or Reduced-Price Lunch, reading performance improved from fall 2021 to spring 2022

Students who qualified for Free or Reduced-Price Lunch gained an average of 120.24 points on the Unified Scale from the fall 2021 to the spring 2022 administration of the Star Reading assessment (Table 9; Figure 8). This gain was statistically significantly different from zero, meaning that it was unlikely to have occurred by chance. The standardized effect size reflecting the magnitude of the mean fall-to-spring gain was 1.14.

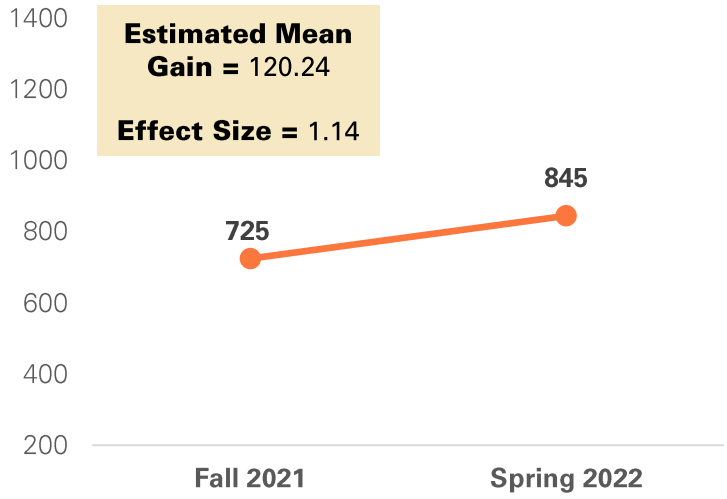
Table 9. Multilevel Model Results for Students Who Qualified for Free or Reduced-Price Lunch

Outcome	Coefficient	Standard Error	*t*-value	Approx. *df*	*p*-value	Effect Size
Star Reading Gain	120.24	7.69	15.63	24.67	<.001*	1.14

* Statistically significant at the .05 level.

As seen in Figure 8, after accounting for nesting, the estimated mean Unified Scale Scores of students who qualified for Free or Reduced-Price Lunch increased from fall to spring during the 2021/2022 school year.

Figure 8. Estimated Unified Scale Score Means for Students Who Qualified For Free or Reduced-Price Lunch



SUMMARY & DISCUSSION

Researchers conducted a secondary data analysis of reading scores for students at two schools that used *From Phonics to Reading* during the 2021/2022 school year. The analyses included 347 students from 29 classrooms in two schools.

Summary of Findings

Overall, findings from the multilevel models indicated that students in this study demonstrated statistically significant improvements in reading scores from fall to spring (i.e., their mean fall-to-spring gain was statistically significantly different from 0). This was also true within each grade (K–2) and within each examined student subgroup (male students, female students, students who qualified for Free or Reduced-Price Lunch). Findings from the visual examination of assessment norms suggest that the average reading performance of students within each grade improved by a grade level or more from fall to spring.

The table below includes a summary of the findings for each research question.

Did participating students demonstrate a statistically significant improvement in reading performance from fall to spring?	On average, kindergarten through second-grade students demonstrated a statistically significant improvement in reading performance from fall 2021 to spring 2022. The mean gain in reading scores corresponded to an effect size of 1.12. On average, within each grade, students demonstrated a statistically significant improvement in reading performance from fall 2021 to spring 2022. Grade-specific mean gains in reading scores corresponded to effect sizes ranging from 1.15 to 1.37.
How did participating students' reading performance compare with the assessment's norm group and performance criteria?	Within each grade, ranges of percentile ranks corresponding to students' average reading performance increased from fall 2021 to spring 2022. Within each grade, grade equivalence scores corresponding to students' average reading performance increased from fall 2021 to spring 2022 by a grade level or more. For kindergarten and second-grade students, average reading performance was below their grade-level benchmark in fall 2021 and increased to be at or above their grade-level benchmark by spring 2022.
Did key subgroups of participating students demonstrate statistically significant improvements in reading performance from fall to spring?	On average, each subgroup examined (i.e., male students, female students, and students who qualified for Free or Reduced-Price Lunch) demonstrated a statistically significant improvement in reading performance from fall 2021 to spring 2022. Subgroup-specific mean gains in reading scores corresponded to effect sizes ranging from 1.08 to 1.18.

Study Strengths and Limitations

This study of *From Phonics to Reading* used a valid, reliable measure of student reading performance. The pretest-posttest design allowed researchers to estimate fall-to-spring gains, which offers more information about student reading performance than a single time point. It also included appropriate analyses that accounted for students being nested within classrooms. However, the study was a treatment-only design and did not include a comparison group. As such, the gains in student reading performance cannot be uniquely attributed to the use of *From Phonics to Reading.* Additionally, the study included a small number of classrooms from two schools with little demographic diversity, which limits the generalizability of the study findings.

Recommendations

This treatment-only, pretest-posttest study provides important findings regarding reading performance improvement among students who used *From Phonics to Reading.* There are additional areas that merit future exploration beyond the scope of this study. To strengthen the research base for *From Phonics to Reading*, researchers recommend the following:

- A study with a comparison group (e.g., a randomized control trial or quasi-experimental study), as well as a larger and more demographically diverse sample, would provide evidence regarding the efficacy of *From Phonics to Reading*.
- A study that examines the implementation of *From Phonics to Reading* in school settings (such as through teacher implementation logs and surveys) would enable researchers to assess how teachers implement the program and their perceptions of the program.

REFERENCES

Banse, H. W. (2022). *From Phonics to Reading Foundational Research.* William H. Sadlier, Inc.

Bleses, D., Højen, A., Dale, P. S., Justice, L. M., Dybdal, L., Piasta, S., ... & Haghish, E. F. (2018). Effective language and literacy instruction: Evaluating the importance of scripting and group size components. *Early Childhood Research Quarterly, 42*, 256–269.

Borman, G. D., Slavin, R. E., Cheung, A., Chamberlain, A. M., Madden, N. A., & Chambers, B. (2005). Success for All: First-year results from the national randomized field trial. *Educational Evaluation and Policy Analysis, 27*(1), 1-22.

Fiester, L. (2010). *Early Warning! Why Reading by the End of Third Grade Matters. KIDS COUNT Special Report.* Annie E. Casey Foundation.

Grimm, K. J. (2008). Longitudinal associations between reading and mathematics achievement. *Developmental Neuropsychology, 33*(3), 410–426.

Hernandez, D. J. (2011). *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation.* Annie E. Casey Foundation.

Lonigan, C. J., & Shanahan, T. (2009). Developing early literacy: Report of the National Early Literacy Panel. Executive summary. *A Scientific synthesis of early literacy development and implications for intervention.* National Institute for Literacy.

McIntosh, K., Horner, R. H., Chard, D. J., Dickey, C. R., & Braun, D. H. (2008). Reading skills and function of problem behavior in typical school settings. *The Journal of Special Education, 42*(3), 131–147.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

NCES. (n.d.) U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Retrieved from Search for Public Schools (ed.gov)

Ortlieb, E., & McDowell, F. D. (2016). Looking closer at reading comprehension: Examining the use of effective practices in a literacy clinic. *English Teaching: Practice & Critique, 15*(2), 260–275.

Rabiner, D. L., Godwin, J., & Dodge, K. A. (2016). Predicting academic achievement and attainment: The contribution of early academic skills, attention difficulties, and social competence. *School Psychology Review, 45*(2), 250–267.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Renaissance. (2022a). *Star Assessments for Reading Technical Manual: Star Early Literacy*. Renaissance Learning.

Renaissance. (2022b). *Star Assessments for Reading Technical Manual: Star Reading*. Renaissance Learning.

Scammacca, N. K., Fall, A., & Roberts, G. (2015). Benchmarks for expected annual academic growth for students in the bottom quartile of the normative distribution. *Journal of Research on Educational Efficacy, 8*, 366–379.

What Works Clearinghouse (WWC). (2022). *What Works Clearinghouse Procedures and Standards Handbook, Version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE).

Wonder-McDowell, C., Reutzel, D. R., & Smith, J. A. (2011). Does instructional alignment matter? Effects on struggling second graders' reading achievement. *The Elementary School Journal*, *112*(2), 259–279.

APPENDIX A: DATA PREPARATION

Researchers followed specific protocols for data cleaning and preparation, including calculating and examining descriptive statistics and identifying any potential outliers.

Specifically, researchers did the following:

- Maintained a methodological log detailing all actions related to the study data (e.g., when data was received, included variables, recoding variables, merging databases, etc.)
- Conducted an audit of the study data to confirm all necessary variables were included
- Conducted descriptive statistics before and after merging and recoding the data to ensure there were no errors
- Checked for missing data and followed up with the involved district to confirm these cases
- Examined outliers in the Unified Scale Scores for the Star Reading assessment data using the z-score approach, which indicated there were no outliers in the study data[13]

[13] Researchers calculated z scores (i.e., each value's standardized deviation from the mean) for each Unified Scale Score observation. Researchers considered z scores > |3.29| as outliers.

APPENDIX B: DETERMINING THE ANALYTIC SAMPLE

Researchers received data for 381 students. Data came from three sources (fall and spring Star Reading assessments, and classroom roster information). Some students took the assessments multiple times within the assessment benchmark window. For these cases, researchers included students' earliest score for the fall and latest score for the spring in the analytic sample.[14]

Across the three sources, 34 students (8.92%) were missing data from one or more of the three data sources (Table B1). Researchers removed cases where students were missing data from any of these data sources, resulting in an analytic sample of 347 students.

Table B1. Missing Data

	n	Percent Missing
Roster Information	2	0.52%
Fall 2021 Assessment	29	7.61%
Spring 2022 Assessment	4	1.50%

Note. One student was missing both demographic and spring assessment data.

Researchers were able to justify some of the cases of missing data. For example, seven students enrolled into one of the two schools in the district after the fall assessment benchmark window (January 2022 or later), therefore did not have fall assessment data. Additionally, one student left one of the two schools in the district before the spring assessment benchmark window in December 2021, therefore did not have spring assessment data.

Six students enrolled in one of the two schools in the district in November or December 2021. Researchers requested information from the district to ensure these six cases were outside of the benchmark window, but these cases were not confirmed.

[14] In fall 2021, five students took the same assessment on different days. Twelve students took the same assessment on the same day. Four students took both assessments within the fall benchmark window. One student took both assessments on the same day.

In spring 2022, four students took the same assessment on different days. Two students took the same assessment twice on one day and once on a later date. Eighteen students took the same assessment on the same day. Seven students took both assessments within the spring benchmark window. One student took both assessments on the same day.

Researchers observed a unique pattern of scores throughout the fall assessment data. There were 18 cases with Unified Scale Scores of 615 (with a Scaled Score of 8). Researchers requested further information from the district to ensure these cases were not errors. The district contacted the assessment developer's Help Desk who confirmed a Scaled Score of 8 is a valid score. Therefore, researchers retained the earliest or latest score for a student in the analytic sample.

APPENDIX C: STUDENT PERFORMANCE RESULTS

These tables include the Unified Scale Score means and standard deviations for the Star Reading assessments by data collection point for the overall analytic sample, as well as by grade and key student subgroups (Table C1).

Table C1. Students' Unified Scale Score Means and Standard Deviations

	Fall Unified Scale Score			Spring Unified Scale Score		
	n	**Mean**	***SD***	***n***	**Mean**	***SD***
All students	347	740.28	105.93	347	863.13	103.57
By Grade						
Kindergarten students	128	666.51	71.34	128	816.58	101.15
First-grade students	108	735.69	65.80	108	845.66	89.61
Second-grade students	111	829.81	105.93	111	933.82	103.57
By Subgroups						
Male students	189	737.62	105.76	189	866.26	103.05
Female students	158	743.46	106.38	158	859.39	104.40
Students who qualified for Free or Reduced-Price Lunch	211	718.19	101.83	211	839.87	104.09