

A Comparison of the Responsive Reading and
Reading Recovery Interventions
for the
Pleasant Valley School District

Edward S. Shapiro, Ph.D.

and

Virginia R. Hampton, Ph.D.

Center for Promoting Research to Practice

Lehigh University

Technical Report # 2010-2



August 23, 2010

Executive Summary

The purpose of this evaluation was to examine the effectiveness of the Responsive Reading intervention in helping first grade students gain literacy skills and to determine whether students participating in the program have significant advantages over students participating in the Reading Recovery program. The evaluation sought to answer four questions:

- (1) Do students demonstrate differences in selected literacy skills at the completion of the intervention?
- (2) Do students demonstrate differences in selected literacy skills at the end of the year?
- (3) Are there group differences in student achievement as measured by the DIBELS?
- (4) Are there group differences in student performance as measured by the DIBELS during second grade?

Classrooms were randomly selected for participation in either the Responsive Reading or Reading Recovery program. A total of 82 low-achieving students participated, with 30 students receiving the Responsive Reading intervention and 52 students receiving the Reading Recovery intervention. The Responsive Reading group had a slightly lower percentage of Hispanic students and Black students, and a slightly higher percentage of White and Asian/Pacific Islander students. In addition, the Responsive Reading group began with slightly higher risk profiles as indicated by higher percentages of students receiving free and reduced lunch, and receiving ESOL, speech and special education services.

The intervention occurred during first grade, with follow-up data collected in second grade. Students began the Responsive Reading or Reading Recovery intervention at either the beginning or the middle of the year in first grade, and remained in the program for an average of 16.3 weeks for Responsive Reading and 17.5 weeks for Reading Recovery. Assessments of selected literacy skills were given at the beginning and end of the intervention, as well as at the end of the year for students who completed the intervention at mid-year. These assessments included Observation Survey measures and a high-frequency word list. In addition, students received DIBELS assessments at three points across the school year (fall, winter, and spring). These included Letter Naming Fluency in the fall, Phoneme Segmentation Fluency and Nonsense Word Fluency in the fall, winter and spring, and Oral Reading Fluency in the winter and spring. A second assessment of Oral Reading Fluency was conducted using a set of identical passages administered in the fall, winter, and spring. As a follow-up to the intervention,

DIBELS assessments in second grade were also examined. These assessments included Nonsense Word Fluency in the fall and Oral Reading Fluency in the fall, winter and spring.

Across all measures, the scores of students in both groups increased across the year, indicating that all students demonstrated progress in their literacy skills. However, differences in performance between the groups were found. Results from the post-intervention assessments indicated that the students in the Responsive Reading program had significantly higher scores on the high frequency word list and all Observation Survey measures except for Writing Vocabulary. Further, end of year assessments suggested that Responsive Reading students continued to have significantly higher skills in recognition of high frequency words, letter identification, concepts about print, phoneme awareness, and oral reading fluency. Among the subgroup of students who began the intervention at the beginning of the year, the Responsive Reading group had significantly higher skills in recognition of high frequency words, concepts about print, spelling, phoneme awareness and oral reading fluency at the end of the year. An examination of end-of-year first grade DIBELS scores found that both groups had end of year scores above the benchmark goal. When comparing the groups, no significant differences between the groups were found for Nonsense Word Fluency and Phoneme Segmentation Fluency. However, students in the Responsive Reading group had significantly higher scores for Oral Reading Fluency at the end of first grade than did the Reading Recovery group. Another assessment of Oral Reading Fluency using identical passages found that the Responsive Reading students also had significantly higher end of year scores for Oral Reading Fluency in first grade after controlling for students' skills in reading the same passage at the beginning of the year. In addition, the Responsive Reading group ended the year with a higher percentage of students considered *Low Risk* and a lower percentage of *At Risk* students than the Reading Recovery group. It should be noted that the Responsive Reading group also began the year with more *Low Risk* students and less *At Risk students* than the Reading Recovery group, however.

In second grade, students from both groups continued to demonstrate progress in their skills across the year. Both groups began the year with scores above benchmark for NWF. With ORF, the Responsive Reading group had scores above benchmark across the year. The Reading Recovery group began the year slightly below the benchmark goal and had winter and spring scores slightly above benchmark for ORF. In addition, the Responsive Reading students had significantly higher scores in Oral Reading Fluency in the fall and winter. No significant differences were found between the groups for Nonsense Word Fluency or for spring Oral Reading Fluency. The Responsive Reading group entered second grade with a higher percentage of *Low Risk* students and a lower percentage of *At Risk* students less risk than the Reading Recovery group. This pattern continued across the year. Of note is the difference between the groups in the change that

occurred between the end of first grade and the beginning of second grade. The Responsive Reading group demonstrated a smaller decrease in the percentage of *Low Risk* students and a smaller increase in the percentage of *At Risk* students over the summer than the Reading Recovery group.

Overall, the results indicate that although both intervention programs resulted in substantial gains for low achieving first grade students, the Responsive Reading program provided students with significantly more skills than the Reading Recovery program. These differences were found immediately following the intervention as well as at the end of first grade. The Responsive Reading students had a higher percentage of students considered *Low Risk* and a lower percentage of *At Risk* students based on DIBELS scores in first grade. A follow-up of these students in second grade found that the students who had participated in Responsive Reading had significantly higher Oral Reading Fluency scores in the fall and winter and more *Low Risk* and less *At Risk* students throughout second grade. .

Introduction

The purpose of this evaluation was to determine whether students in the Responsive Reading intervention gain significantly more literacy skills than students in the Reading Recovery intervention. First grade classrooms were randomly assigned to either the Reading Recovery or Responsive Reading intervention. Low-achieving students participated in either Reading Recovery or Responsive Reading, depending on their classroom. Students were identified as low-achieving through a process that began with informal assessments and an examination of DIBELS scores at the end of kindergarten, followed by a more formal assessment process at the beginning of first grade. . Students with the lowest rankings based on these criteria were included in these programs. The intervention began at the beginning of the year for some students and at mid-year for other students. The amount of time for each intervention varied, based on student progress. Students were assessed on selected literacy skills before beginning the intervention and at the end of the intervention, using the Observation Survey and a high frequency word list. For those students who completed the intervention at mid-year, these assessments were again administered at the end of the year. In addition, students had DIBELS assessments which included LNF, PSF and NWF in the fall, and PSF, NWF and ORF in the winter and spring. Second grade DIBELS assessments were also examined, which included fall NWF and ORF, and winter and spring ORF.

Measures

To evaluate progress, students were assessed with the following measures:

- *Woodcock Johnson words.*

This is a 15 item list of high frequency words, taken and modified from the Woodcock Johnson, to assess students' recognition of common words. This assessment was given to students prior to beginning the intervention and at the end of the intervention. For those students who completed the intervention at mid-year, this word list was also administered at the end of the year. Students were scored based on the number of words read correctly, without being timed.

- *Observation Survey*

The Observation Survey includes 6 measures. *Letter Identification* examines children's ability to identify lowercase and uppercase letters. *Word Test* assesses students' skills in reading high frequency words. The *Concepts About Print* measure evaluates students' understanding about print, such as one-to one correspondence between what is printed and what is spoken, and the meaning of punctuation. The *Writing Vocabulary* measure evaluates spelling. *Hearing and Recording Sounds in Words* assesses phoneme awareness. The *Text Reading* measure examines oral reading fluency and indicates

students' reading level. These assessments were administered to students before the intervention began and at the completion of the intervention. Those students who finished the intervention in the middle of the year also were assessed with the Observation Survey measures at the end of the year.

- **Dynamic Indicators of Basic Early Literacy Skills (DIBELS)**

Four DIBELS measures were included. *Letter Naming Fluency* (LNF) examines students' ability to identify uppercase and lowercase letters. *Phoneme Segmentation Fluency* (PSF) measures phonological awareness, and *Nonsense Word Fluency* (NWF) measures letter-sound correspondence. In addition, oral reading fluency is assessed by the *Oral Reading Fluency* (ORF) measure. The DIBELS were administered three times per year. First grade assessments include LNF, PSF and NWF in the fall, and PSF, NWF, and ORF assessments in the winter and spring. An additional assessment of *Oral Reading Fluency* was administered by having students read identical passages in the fall, winter and spring.

In second grade, fall assessments include NWF and ORF, whereas the winter and spring assessments only include the ORF measure.

Demographic Characteristics

The Reading Recovery group included 52 first grade students and the Responsive Reading group had 30 first grade students. Gender was equally represented across the two groups. The racial and ethnic composition of the two groups was similar, with a slightly higher percentage of Hispanic and Black students and a slightly lower percentage of White and Asian/Pacific Islander students in the Reading Recovery group (see Table 1).

Table 1. Gender and Race/Ethnicity By Group

Group	Gender		Race/Ethnicity			
	Male	Female	Hispanic	White	Black	Asian/Pacific Islander
Responsive Reading	51.6%	48.4%	16.1%	74.2%	3.2%	6.5%
Reading Recovery	51.8%	48.2%	21.4%	66.1%	10.7%	1.8%

The Responsive Reading group had a higher percentage of students receiving speech, ESOL, and special education services (see Table 2). In addition, the percentage of students receiving free and reduced lunch was higher for the Responsive Reading group. The data suggests that the Responsive Reading group had a higher percentage of students likely to have higher risk of failure than the Reading Recovery group.

Table 2. Services and Meal Status Percentages

Group	Services			Meal Status
	Speech	ESOL	Special Education	Free or Reduced Lunch
Responsive Reading	22.6%	6.5%	3.2%	32.2%
Reading Recovery	10.7%	3.6%	0%	21.5%

Intervention

The Reading Recovery intervention targets low-achieving students to help them gain skills in reading and writing. With Responsive Reading, the emphasis is on reading fluency and comprehension for children having difficulty learning to read. Students in Reading Recovery met daily with their teacher on an individual basis for 30 minutes per day. For Responsive Reading, groups of three students met daily with their teacher for 40 minutes per day.

As seen in Table 3, half of the Reading Recovery students began the program at the beginning of the year and half began at mid-year. For Responsive Reading, 64.5% began at the beginning of the year and 35.5% began mid-year. The average amount of time students participated in Reading Recovery was 17.5 weeks and the average for Responsive Reading was 16.3 weeks (see Table 4).

Table 3. Percentage of Students at Intervention Starting Point

Group	Percentage of Students Starting Intervention at Each Time Point	
	Beginning of the Year	Middle of Year
Responsive Reading	64.5%	35.5%
Reading Recovery	50%	50%

Table 4. Total Weeks of Intervention By Group

Group	Total Weeks of Intervention	
	Average	Range
Responsive Reading	16.3	8-26
Reading Recovery	17.5	10-33

Evaluation Questions

The purpose of the evaluation was to determine whether students gain significantly more skills than when participating in the Responsive Reading versus the Reading Recovery interventions. Four specific questions were examined:

- (1) Do students demonstrate differences in selected literacy skills at the completion of the intervention?
- (2) Do students demonstrate differences in selected literacy skills at the end of the year?
- (3) Are there group differences in student achievement as measured by the DIBELS?
- (4) Are there group differences in student performance during second grade?

Evaluation Question 1: Do students demonstrate differences in selected literacy skills at the completion of the intervention?

Assessments using the Woodcock Johnson words and the Observation Survey were administered to students before the intervention began and at the conclusion of the intervention. Pretest scores reflect students' skills prior to beginning the intervention and were administered at the beginning of the year for those students who began in the fall, and in mid-year for students who began the intervention in the middle of the year. The posttest scores indicate students' scores at the completion of the intervention and therefore could have occurred at mid-year or at the end of the year. Table 4 presents the means and standard deviations for the pretest and posttest scores for the Woodcock Johnson words and the Observation Survey measures. As the table indicates, both groups demonstrated gains at the end of the intervention. The posttest scores for the Responsive Reading group are generally higher than the scores for the Reading Recovery students, with only slightly higher scores for Letter ID, Writing Vocabulary, and Hearing and Recording Sounds.

Table 5. Means and Standard Deviations of WJ Words and Observation Survey By Group

Measure	Responsive Reading				Reading Recovery			
	Pretest N=26		Posttest N=25		Pretest N=52		Posttest N=49	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
WJ Word	4.73	4.41	14.44	1.70	5.10	4.48	11.51	3.34
OS Letter ID	48.50	5.20	53.84	0.37	48.58	4.94	53.04	1.44
OS Word	6.04	7.09	19.28	1.67	6.98	7.37	16.86	3.95
OS Concepts About Print	12.88	4.98	23.04	1.40	13.21	4.71	19.71	2.75
OS Writing Vocabulary	14.58	12.89	44.72	9.81	17.31	14.45	43.16	13.31
OS Hearing and Recording Sounds	24.35	9.53	36.72	0.84	23.02	10.90	35.43	1.95
OS Text	3.23	4.56	17.96	4.90	3.06	3.32	12.96	5.07

To determine whether there were any significant differences between groups in students' growth during the intervention period, a Multiple Analysis of Covariance (MANCOVA) was conducted. The pretest scores were used as covariates, to account for any differences among the groups in skills at the beginning of the intervention. The Pillai's Trace multivariate test of overall differences between groups was statistically significant ($p < .001$). Even after controlling for pretest scores, there was a significant difference in performance between the groups. Univariate between-subject tests indicated significant differences between the groups for all measures except Writing Vocabulary. At the end of the intervention, the Responsive Reading students had significantly higher scores on the Woodcock Johnson words and all Observation Survey measures except for Writing Vocabulary when controlling for differences in scores prior to beginning the intervention. Figures 1 through 7 show the growth of students across the intervention period for each measure and indicate any significant differences in posttest scores between the groups. Note that with the MANCOVA procedure, comparisons are made among adjusted means (i.e., the means that occur after controlling for differences in all of the covariates). In other words, this procedure examines the posttest scores based on the groups being equal on all the covariates. This can explain a significant difference when only a small difference appears between group means for a particular measure. See Table 6 for the adjusted means for the posttest scores.

Figure 1. Mean Scores for Woodcock-Johnson Words at Pretest and Posttest By Group

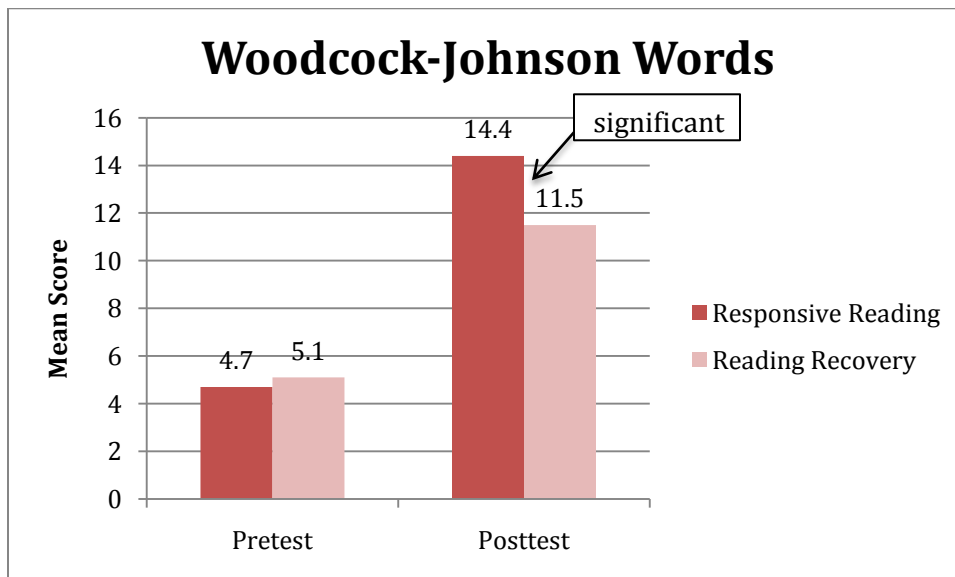


Figure 2. Mean Scores for Observation Survey: Letter Identification at Pretest and Posttest By Group

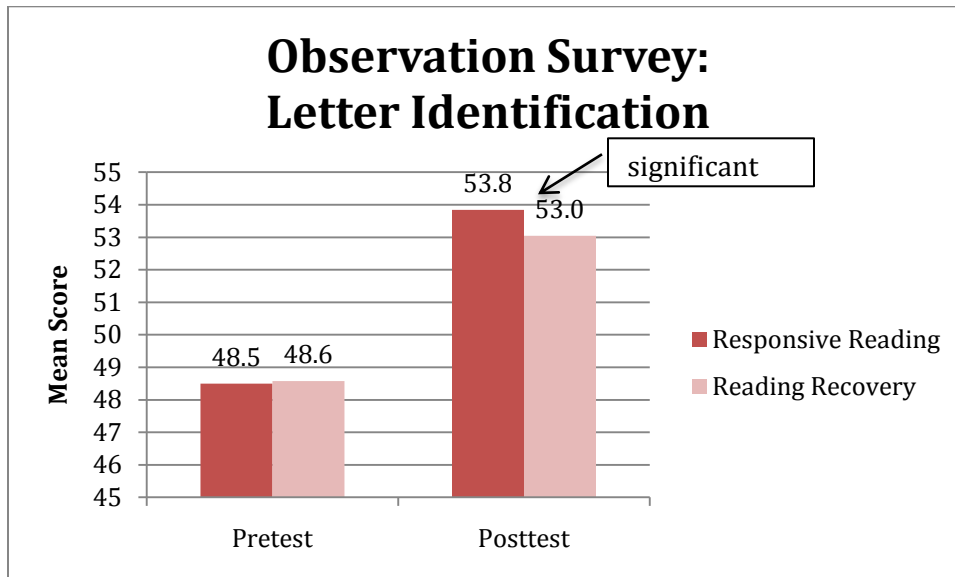


Figure 3. Mean Scores for Observation Survey: Word Test at Pretest and Posttest By Group

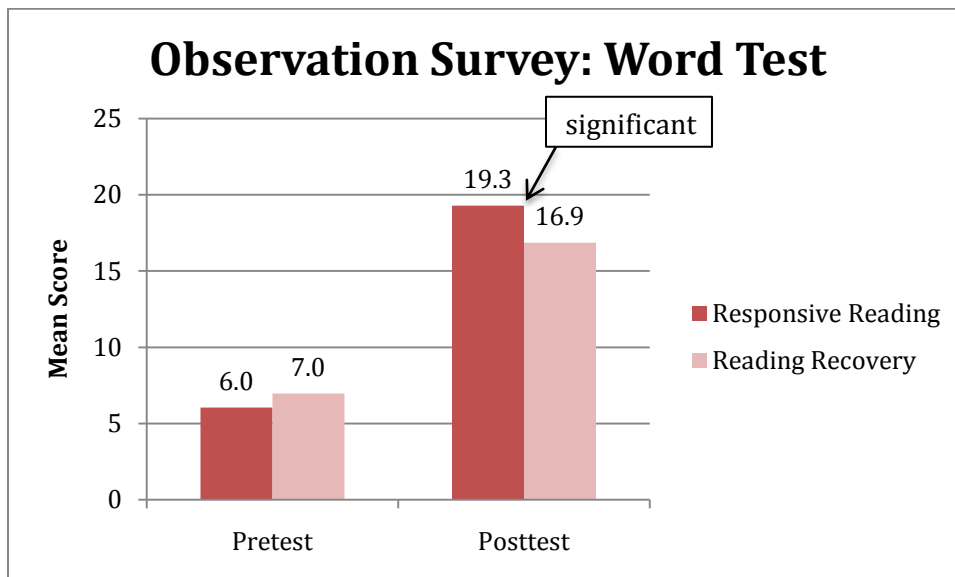


Figure 4. Mean Scores for Observation Survey: Concepts About Print at Pretest and Posttest By Group

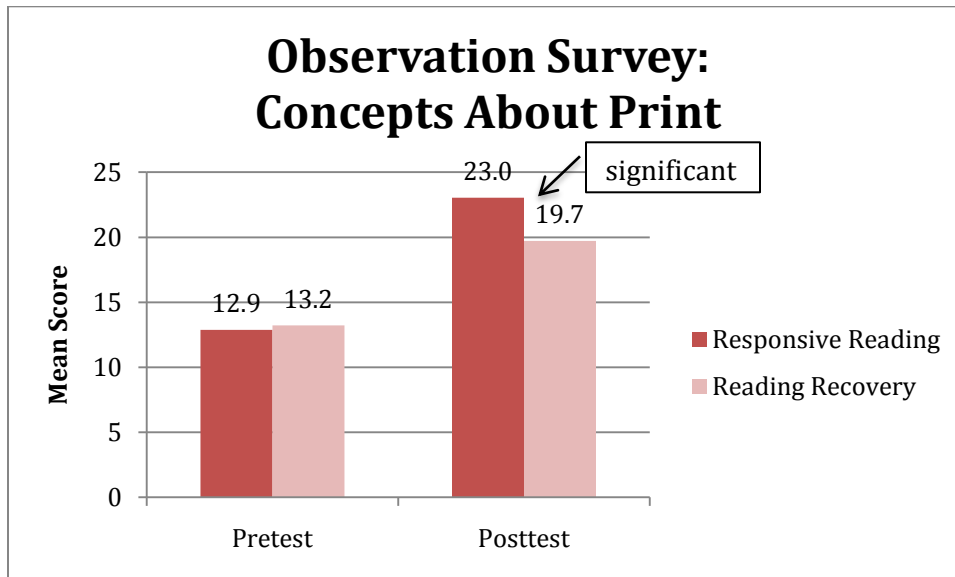


Figure 5. Mean Scores for Observation Survey: Writing Vocabulary at Pretest and Posttest By Group

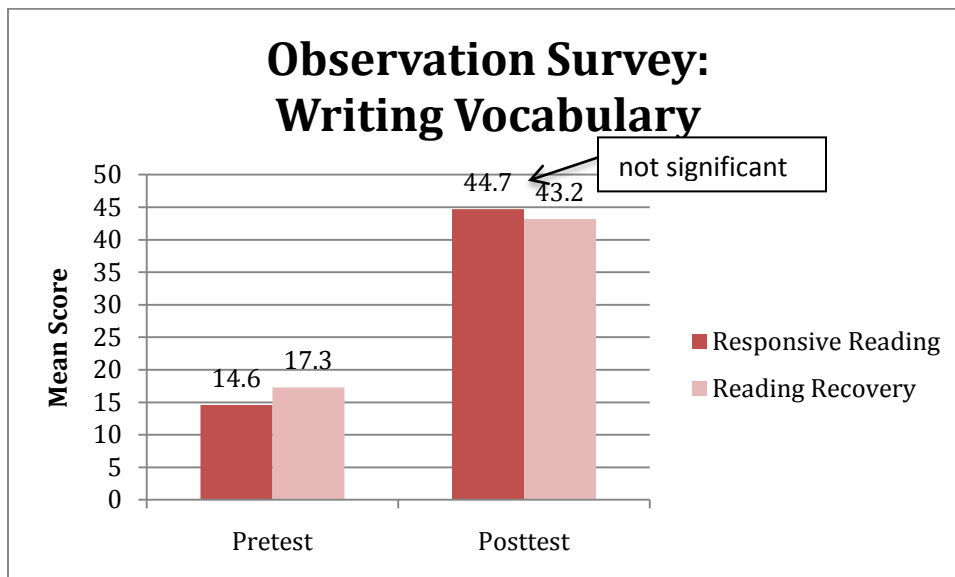


Figure 6. Mean Scores for Observation Survey: Hearing and Recording Sounds in Words at Pretest and Posttest By Group

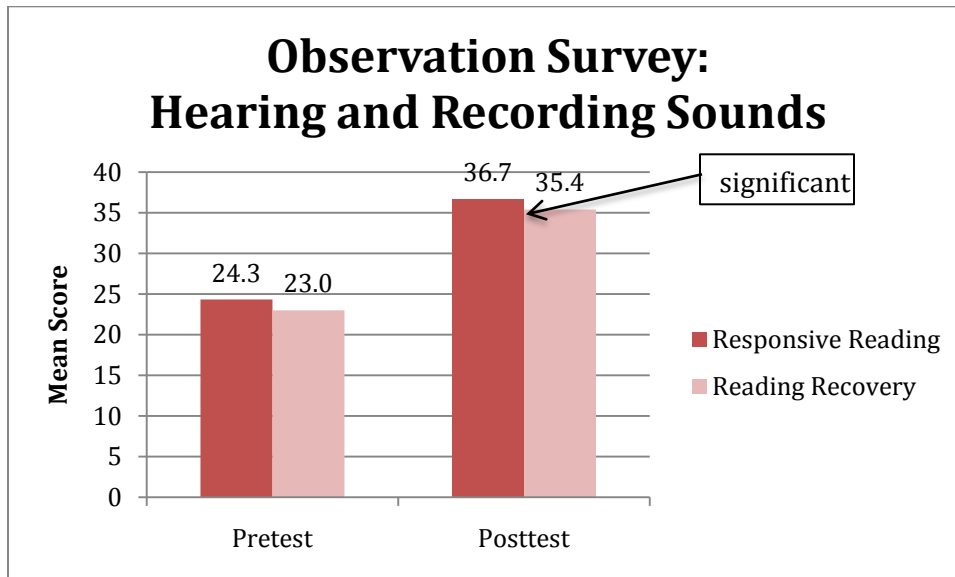


Figure 7. Mean Scores for Observation Survey: Text Reading at Pretest and Posttest By Group

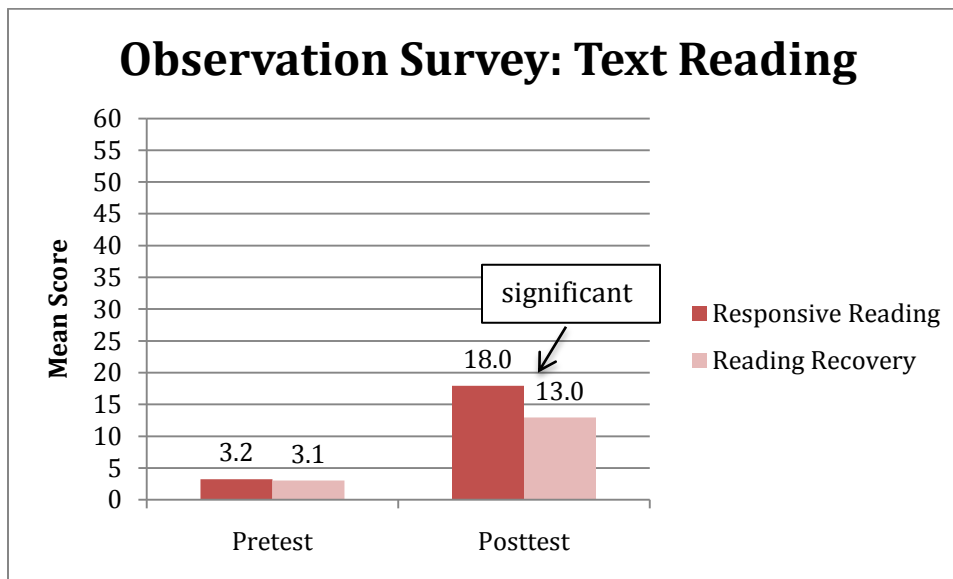


Table 6. Adjusted Means and Standard Errors for WJ Words and Observation Survey at Posttest By Group

Measure	Responsive Reading		Reading Recovery	
	Adjusted Mean	Standard Error	Adjusted Mean	Standard Error
WJ Word	14.54	0.49	11.46	0.34
OS Letter ID	53.84	0.25	53.04	0.17
OS Word	19.35	0.60	16.82	0.42
OS Concepts About Print	23.02	0.38	19.72	0.27
OS Writing Vocabulary	46.41	2.18	42.30	1.52
OS Hearing and Recording Sounds	36.77	0.34	35.40	0.24
OS Text	18.07	0.70	12.90	0.49

In summary, selected literacy skills were assessed before beginning the intervention and at the end of the intervention period, using Woodcock Johnson words and Observation Survey measures to examine the effects of the two interventions. Both groups demonstrated gains across all literacy skills. However, when controlling for differences in pretest scores, the performance of students at the end of the intervention period was significantly higher for students in the Responsive Reading group as compared to students in the Reading Recovery group on the Woodcock Johnson words and on the Observation Survey measures of Word Test, Letter Identification, Concepts About Print, Hearing and Recording Sounds in Words, and Text Reading.

Evaluation Question 2: Do students demonstrate differences in selected literacy skills at the end of the year?

Although the timing of the pretest and posttest assessments varied across the students, the end-of-year score reflects an assessment that occurred at the same point for all students. Students who completed the intervention at mid-year were assessed again at the end of the year. For those students who completed the intervention at the end of the year, the end of the year score is also the posttest score because the post-intervention assessment occurred at the year's end. Table 7 presents the means and standard deviations of students from both groups at the beginning of the intervention and the end of the year. As seen below, both groups demonstrated gains across all measures. The end of year scores for the Responsive Reading students are higher across all measures except for Writing Vocabulary, in which both groups have the same end of year score.

Table 7. Pretest and End of Year Means and Standard Deviations of WJ Words and Observation Survey Measures By Group

Measure	Responsive Reading				Reading Recovery			
	Pretest N=26		End of Year N=30		Pretest N=52		End of Year N=48	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
WJ Word	4.73	4.41	14.93	0.25	5.10	4.48	13.27	2.33
OS Letter ID	48.50	5.20	53.93	0.25	48.58	4.94	53.46	0.83
OS Word	6.04	7.09	19.97	0.18	6.98	7.37	18.62	2.29
OS Concepts About Print	12.88	4.98	23.53	0.86	13.21	4.71	20.54	2.65
OS Writing Vocabulary	14.58	12.89	47.53	10.53	17.31	14.45	47.62	11.73
OS Hearing and Recording Sounds	24.35	9.53	36.97	0.18	23.02	10.90	35.50	2.54
OS Text	3.23	4.56	22.77	4.08	3.06	3.32	16.50	4.44

To determine whether the differences in performance were significant, a Multiple Analysis of Covariance (MANCOVA) was conducted using pretest scores as covariates. The Pillai's Trace multivariate test of overall differences between groups was statistically

significant ($p < .001$). Univariate between-subject tests found significant differences between the groups for all measures except for Writing Vocabulary. Figures 8 through 14 display the pretest and end of year scores for each measure, and indicate any significant differences. See Table 8 for the adjusted means for the end of year scores.

Figure 8. Mean Scores for Woodcock-Johnson at Pretest and End of Year By Group

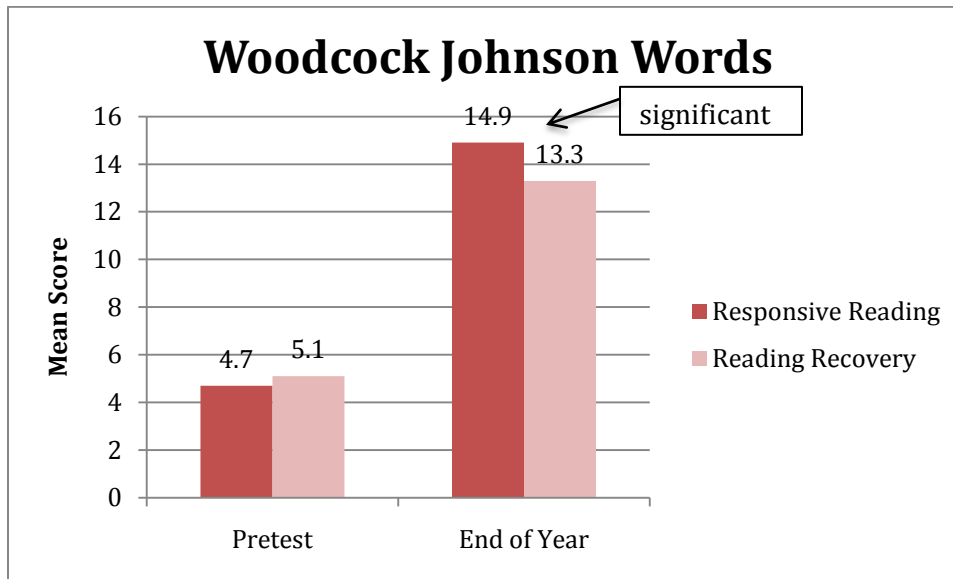


Figure 9. Mean Scores for Observation Survey: Letter Identification at Pretest and End of Year By Group

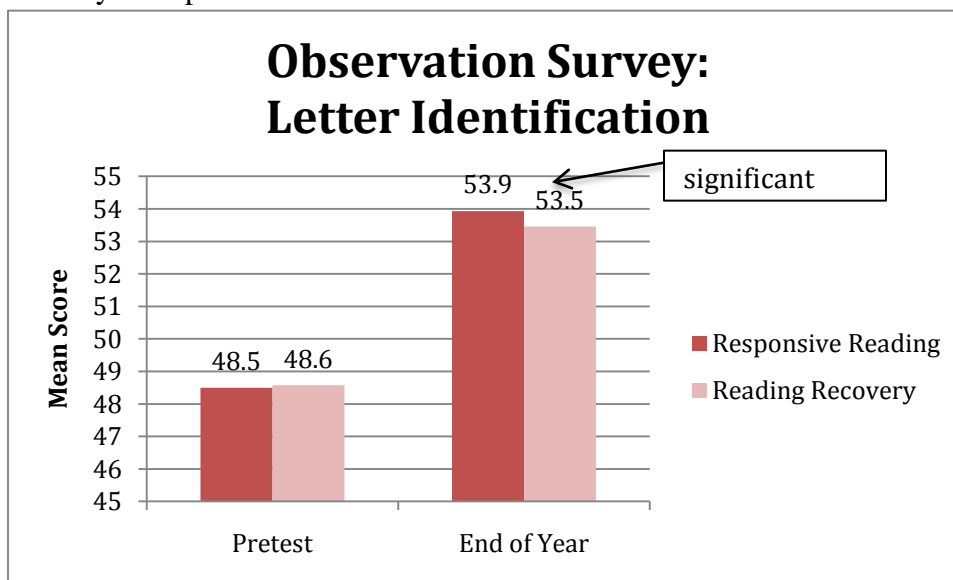


Figure 10. Mean Scores for Observation Survey: Word Test at Pretest and End of Year By Group

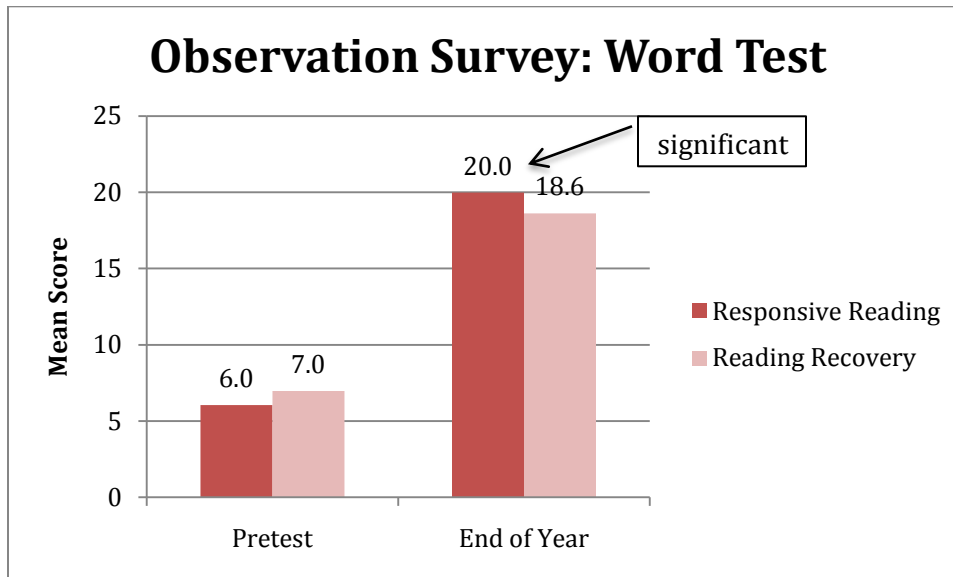


Figure 11. Mean Scores for Observation Survey: Concepts About Print at Pretest and End of Year By Group

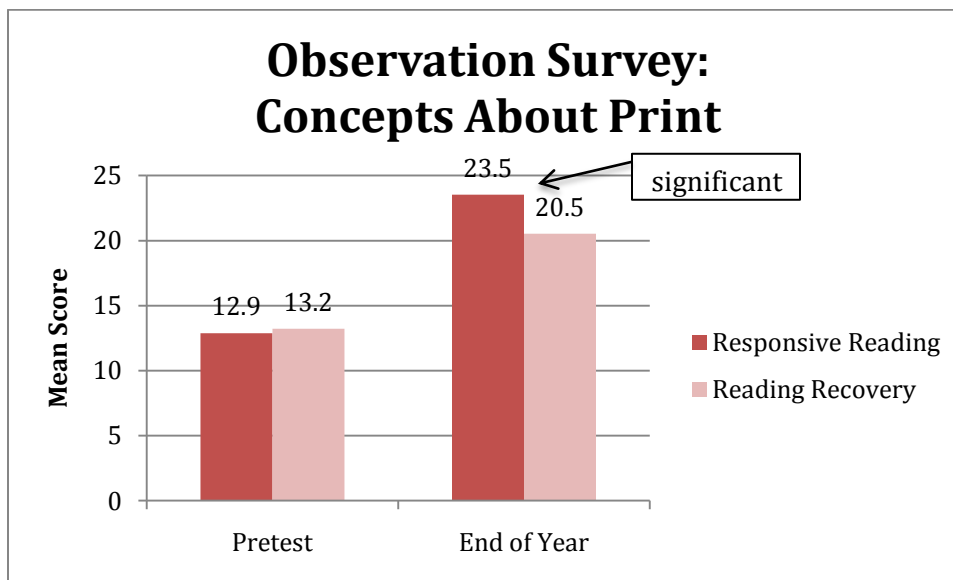


Figure 12. Mean Scores for Observation Survey: Writing Vocabulary at Pretest and End of Year By Group

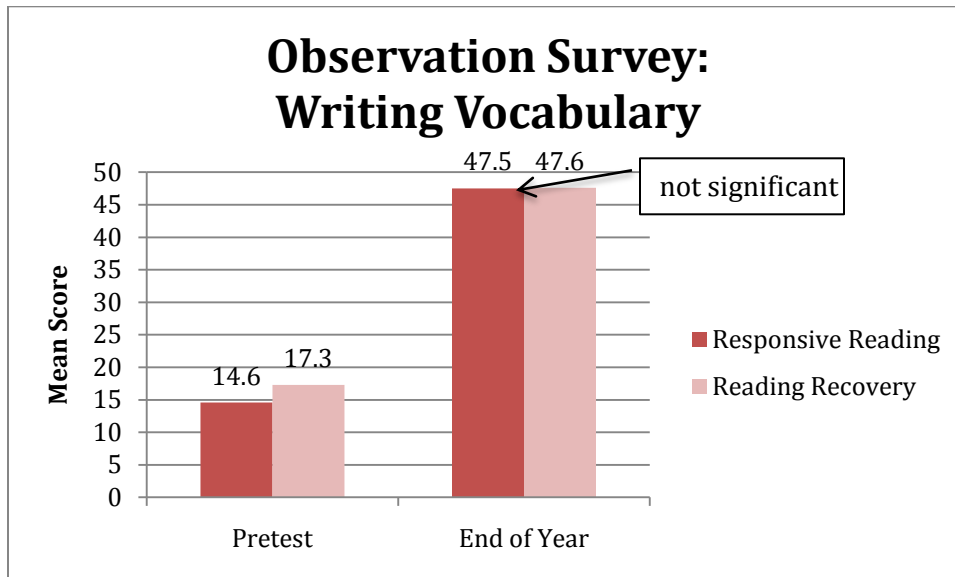


Figure 13. Mean Scores for Observation Survey: Hearing and Recording Sounds in Words at Pretest and End of Year By Group

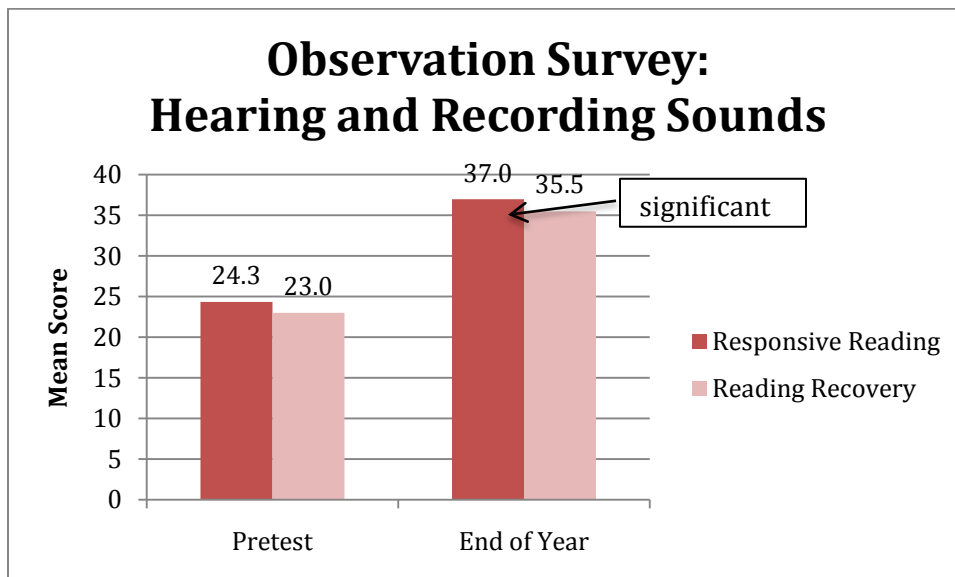


Figure 14. Mean Scores for Observation Survey: Text Reading at Pretest and End of Year By Group

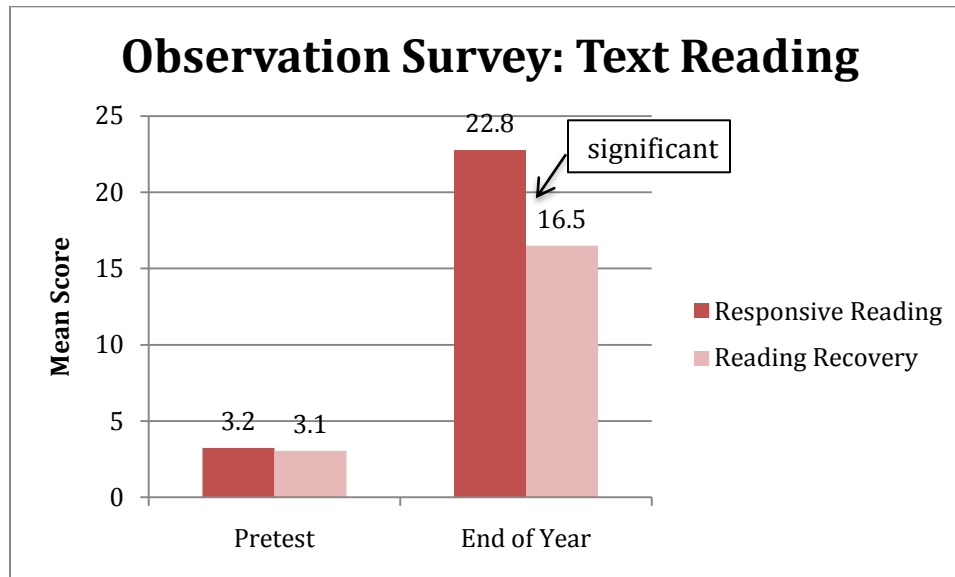


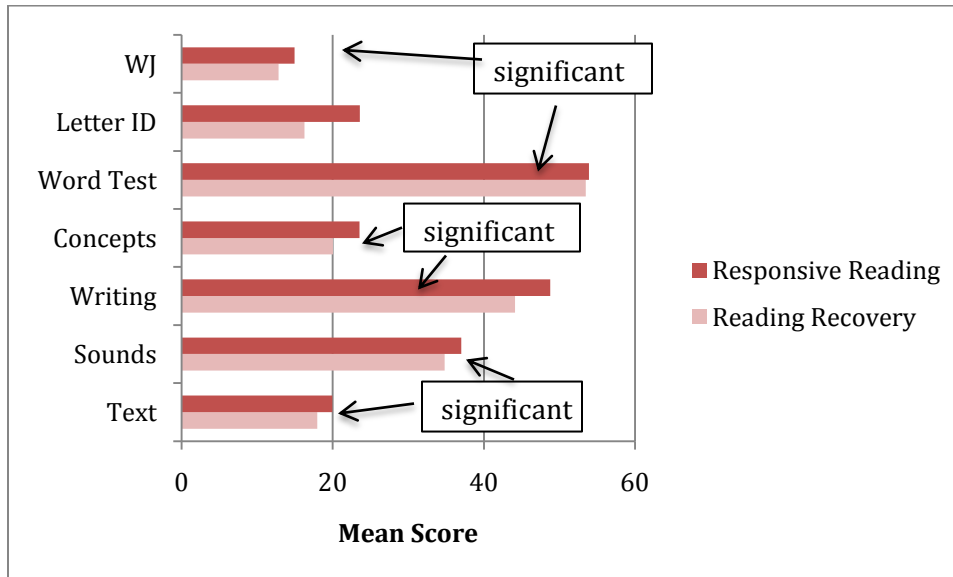
Table 8. Adjusted Means and Standard Errors for WJ Words and Observation Survey at End of the Year By Group

Measure	Responsive Reading		Reading Recovery	
	Adjusted Mean	Standard Error	Adjusted Mean	Standard Error
WJ Word	15.00	0.39	13.16	0.28
OS Letter ID	53.93	0.15	53.43	0.11
OS Word	19.97	0.38	18.54	0.27
Os Concepts About Print	23.51	0.45	20.50	0.32
OS Writing Vocabulary	50.46	2.17	46.55	1.53
OS Hearing and Recording Sounds	37.11	0.45	35.34	0.32
OS Text	22.76	0.91	16.68	0.64

In addition, a MANCOVA was conducted for the subgroup of students who began the intervention at the beginning of year to examine whether differences in outcomes were maintained over a longer period of time. Given that some students had posttest assessments at the end of the year and therefore the end of year score reflected the students' skills at the end of the intervention, this analysis examined the end-of-year skills of students who had completed the intervention at mid-year (N=19 for Responsive Reading, N=24 for Reading Recovery). Figure 15 displays the results. The Pillai's Trace

test of multivariate differences between groups was statistically significant ($p < .05$). This indicates a significant difference in end of year performance between the groups. Univariate tests of between-subject effects found significant differences between the group on the word list and all Observation Survey measures except for Letter Identification.

Figure 15. Mean Scores for WJ and Observation Survey measures at End of Year by Group for Students Who Began Intervention at Beginning of Year



To summarize, the same assessments that were administered at the end of the intervention were given at the end of the year for students who completed the intervention in mid-year. For those students who finished the intervention at the end of the year, the posttest score was also considered the end of year score. The performance of all students was examined to determine whether there were any significant differences between the groups at the end of the year. When controlling for differences in pretest scores, students in the Responsive Reading group had significantly higher growth in their skills as measured by Woodcock Johnson words, and the Observation Survey measures of Letter Identification, Word Test, Concepts About Print, Hearing and Recording Sounds, and Text Reading. Given that the change from pretest to end of year reflected two different time periods (i.e., from beginning of year to end of year for the students who began in the fall, and from mid-year to end of year to students who began mid-year), analyses were also conducted to examine the growth of the subgroup of students who began the intervention in the fall. For this subgroup, the Responsive Reading group had significantly higher performance on all measures except for Observation Survey Letter Identification.

Evaluation Question 3: Are there group differences in student achievement as measured by DIBELS?

To further evaluate whether there were differences in student achievement, the DIBELS scores of each group were examined. Table 9 presents the means and standard deviations for each group at the fall, winter and spring assessments. The Responsive Reading students have higher mean scores for each measure across the year.

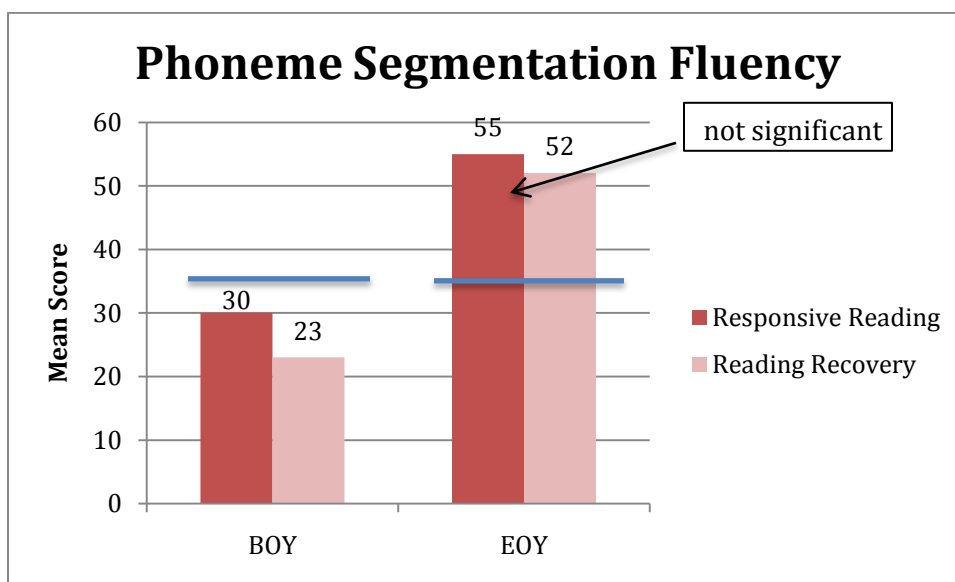
Table 9. Means and Standard Deviations of First Grade DIBELS Scores By Group

Measure	Responsive Reading N=30		Reading Recovery N=50	
	Mean	Standard Deviation	Mean	Standard Deviation
<i>Fall DIBELS</i>				
Letter Naming Fluency	34.74	11.83	30.48	9.39
Phoneme Segmentation Fluency	29.65	15.12	23.43	14.74
Nonsense Word Fluency	22.00	10.87	15.30	8.05
<i>Winter DIBELS</i>				
Phoneme Segmentation Fluency	57.07	7.94	54.68	7.31
Nonsense Word Fluency	57.00	10.52	48.81	11.69
Oral Reading Fluency	23.40	13.70	15.86	6.60
<i>Spring DIBELS</i>				
Phoneme Segmentation Fluency	54.50	5.92	51.73	8.24
Nonsense Word Fluency	76.07	21.04	69.56	19.63
Oral Reading Fluency	58.10	19.14	44.50	15.39

Figures 16 through 18 display the scores and benchmark goals at the beginning of the year (BOY) and at the end of the year (EOY), and indicate any significant differences between groups in the scores. Both groups began the year below benchmark for PSF and NWF, and were above benchmark at the end of the year for PSF, NWF, and ORF. To assess whether the groups had significant differences in their spring DIBELS scores, a

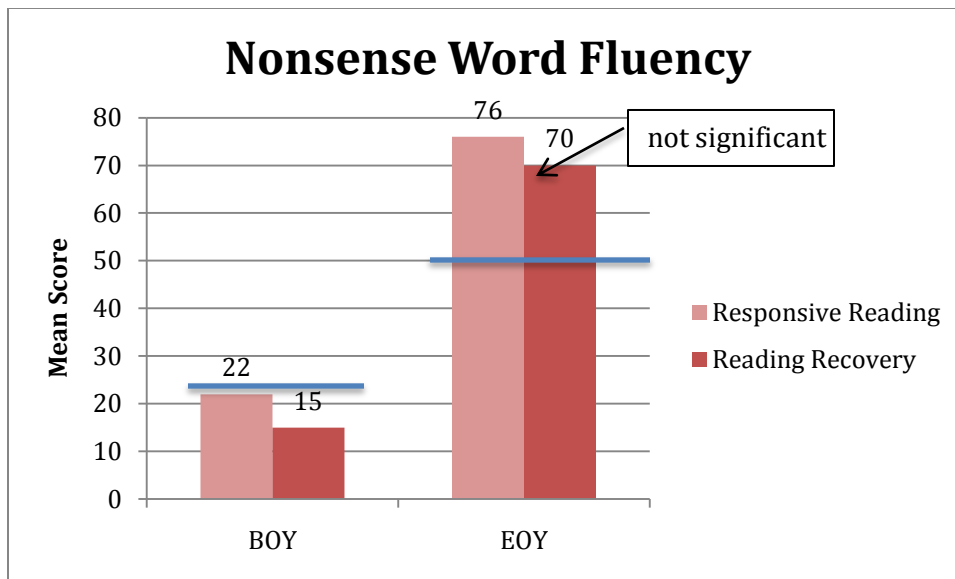
MANCOVA was conducted using the fall PSF score as a covariate. Preliminary MANCOVA analyses revealed that the fall LNF and NWF scores did not significantly contribute to the end of DIBELS scores, so they were dropped as covariates. Differences among the winter scores were not included in these analyses since some students began their intervention after the winter DIBELS assessments, whereas other students would have ended their intervention during that same time period. The Pillai's Trace multivariate test of overall differences between groups was statistically significant ($p < .01$). Univariate tests of between subjects effects found a significant difference between the groups on ORF. This indicates that when controlling for students' DIBELS scores at the beginning of the year, there was a significant difference between the groups for ORF.

Figure 16. First Grade DIBELS PSF Scores By Group



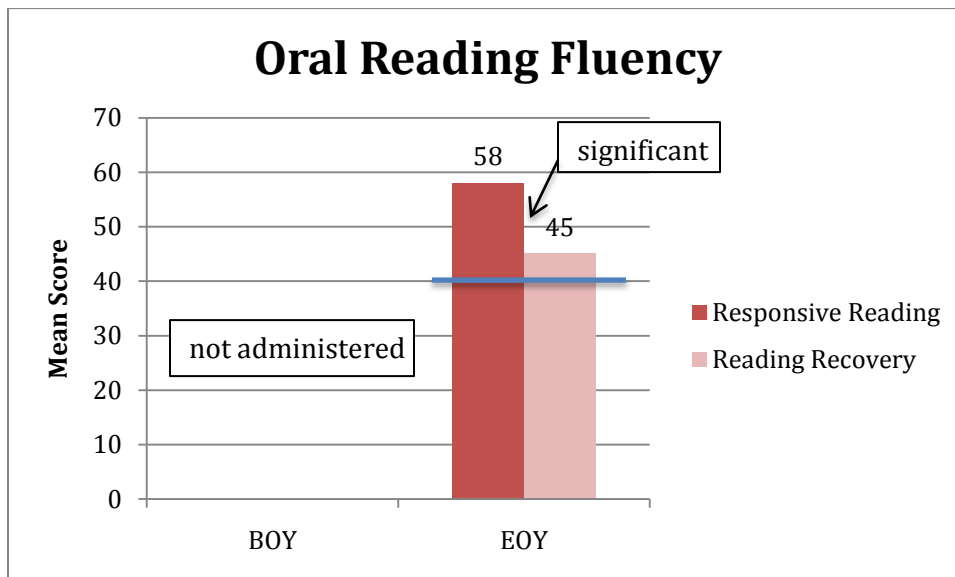
Note. Blue lines indicate benchmark goals for each assessment point.

Figure 17. First Grade DIBELS NWF Scores By Group



Note. Blue lines indicate benchmark goals for each assessment point.

Figure 18. First Grade DIBELS Oral Reading Fluency Scores By Group

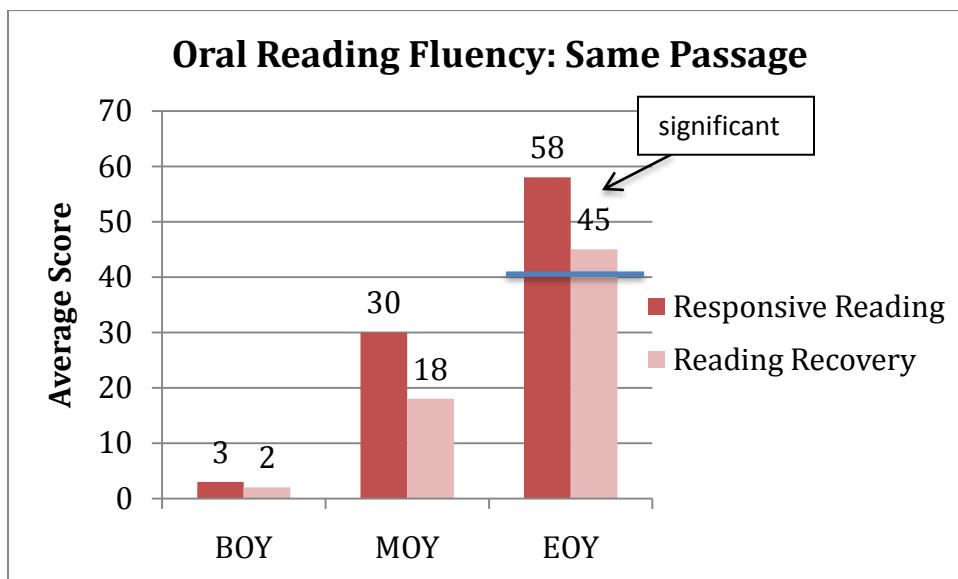


Note. Blue line indicates benchmark goal.

An additional assessment of Oral Reading Fluency was conducted. The typical administration of the DIBELS ORF occurs in the winter and spring assessments, with different passages given at each benchmark point. However, the passages used at the end of the year were also administered to students at the fall and winter assessments, to

examine students' progress. As seen in Figure 19, the groups began the year with similar mean scores, but at the winter and spring assessment points, the Responsive Reading students had higher scores than the students in the Reading Recovery group. An Analysis of Covariance (ANCOVA) was conducted for the spring ORF, using the beginning of year ORF scores as the covariate. (The ANCOVA was not conducted for the winter scores because some students were just beginning the intervention and some students were ending the intervention at that point.) The spring ORF scores of the Responsive Reading students were statistically significantly higher than the scores of Reading Recovery students when taking into account their fall ORF scores using the same reading passages ($p < .001$ level).

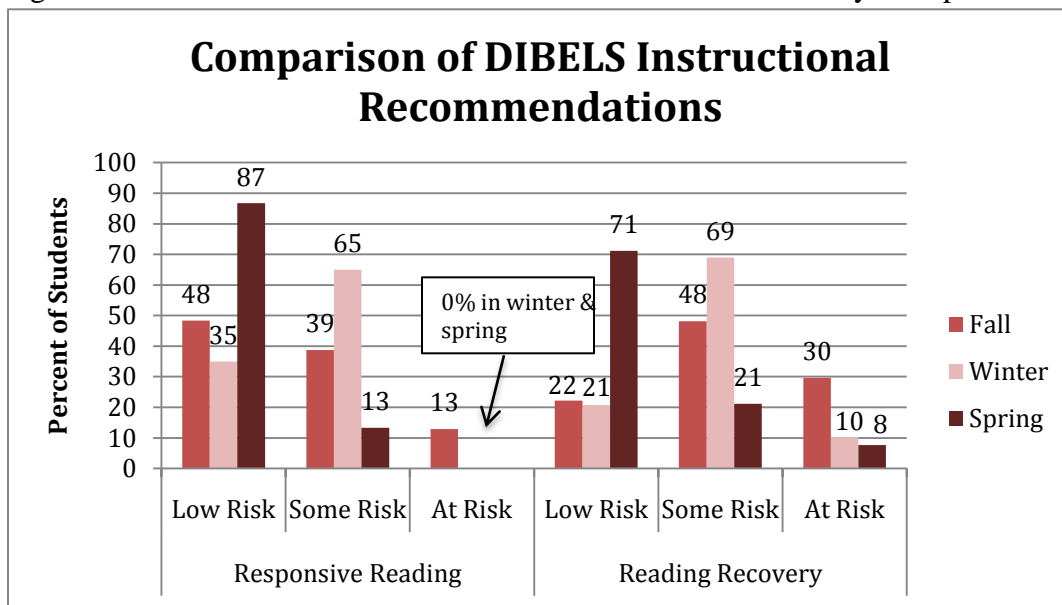
Figure 19. First Grade DIBELS Oral Reading Fluency Scores By Group for the Same Passages



Note. Blue line indicates benchmark goal at the end of year assessment.

Further examination of the DIBELS data included determining the percentage of students considered *Low Risk*, *Some Risk*, and *At Risk* based on DIBELS instructional recommendations from benchmark scores for the two groups across the year. As indicated in Figure 20, the Responsive Reading group ended the year with a higher percentage of students considered *Low Risk* (87%) and a lower percentage of *At Risk* students (0%) than the Reading Recovery group (Low risk = 71%, At risk = 8%). Of note is that there were no *Low Risk* students in the Responsive Reading group in the winter or spring. It should also be noted that the Responsive Reading group began the year with more *Low Risk* students and less *Some Risk* and *At Risk* students than the Reading Recovery group.

Figure 20. First Grade DIBELS Instructional Recommendations By Group



To summarize, both groups attained end of year DIBELS scores above benchmark goals. However, Responsive Reading students had higher spring PSF, NWF and ORF scores. When taking into account differences in student performance prior to the intervention, students in Responsive Reading had significantly higher ORF scores. An additional assessment of ORF that included the same passages across the year also found that the skills of the Responsive Reading students were significantly higher than the skills of the Reading Recovery students when controlling for differences in students' scores at the beginning of the year. NSF and PSF scores did not differ significantly between the groups. Further, at the end of the year, the Responsive Reading group had a higher percentage of *Low Risk* students and a lower percentage of *At Risk* students than did the Reading Recovery group. However, it should be noted that these differences in risk levels were also found at the beginning of the year.

Evaluation Question 4: Are there group differences in student performance as measured by the DIBELS during second grade?

To assess the continued effects of the intervention on student performance, the DIBELS scores of students were examined across the second grade year. Both groups demonstrated gains across the year and had scores above benchmark for the fall NWF. With ORF, the students who participated in Responsive Reading began the year above benchmark and the students who participated in Reading Recovery were slightly below benchmark at the beginning of the year. Both groups were above benchmark goals for the winter and spring ORF, although the Reading Recovery group was only slightly above benchmark. As seen in Table 10, the mean scores were higher for the Responsive Reading group for all measures at all assessment points.

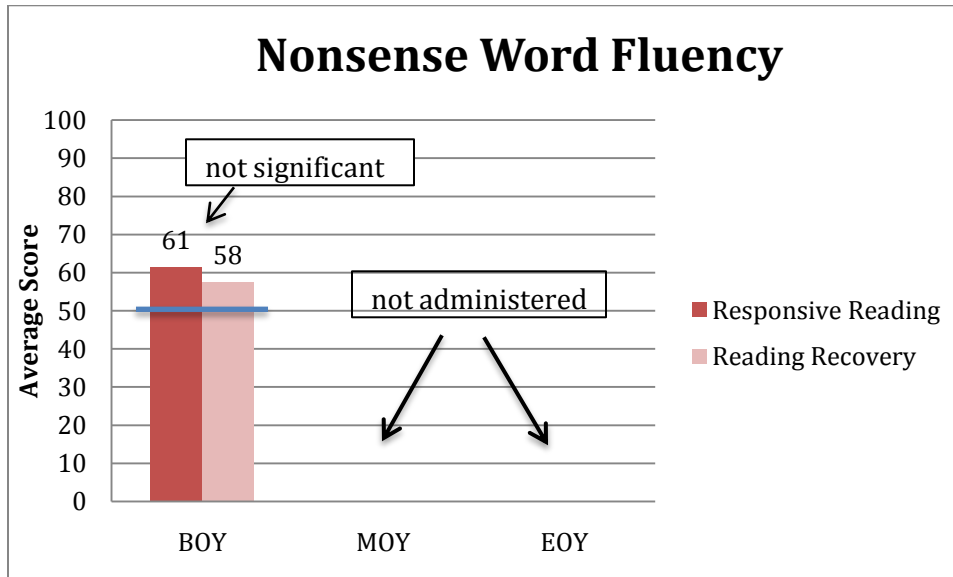
Table 10. Means and Standard Deviations of Second Grade DIBELS Scores By Group

Measure	Responsive Reading N=21		Reading Recovery N=47	
	Mean	Standard Deviation	Mean	Standard Deviation
<i>Fall DIBELS</i>				
Nonsense Word Fluency	61.43	23.30	57.62	22.12
Oral Reading Fluency	48.38	15.355	37.92	14.28
<i>Winter DIBELS</i>				
Oral Reading Fluency	83.00	20.133	70.76	23.56
<i>Spring DIBELS</i>				
Oral Reading Fluency	104.00	20.57	92.57	26.23

See Figures 21 and 22 for the mean scores, benchmark goals and any significant differences in scores between the groups for each measure. A MANOVA examined the differences between the groups. Preliminary analyses indicated that the first grade fall DIBELS scores did not significantly contribute to the second grade scores, and were therefore not needed as covariates in the analyses. The Pillai's Trace multivariate test of overall differences between groups was statistically significant ($p=.033$). Univariate tests of between subject effects found significant differences between groups on fall and

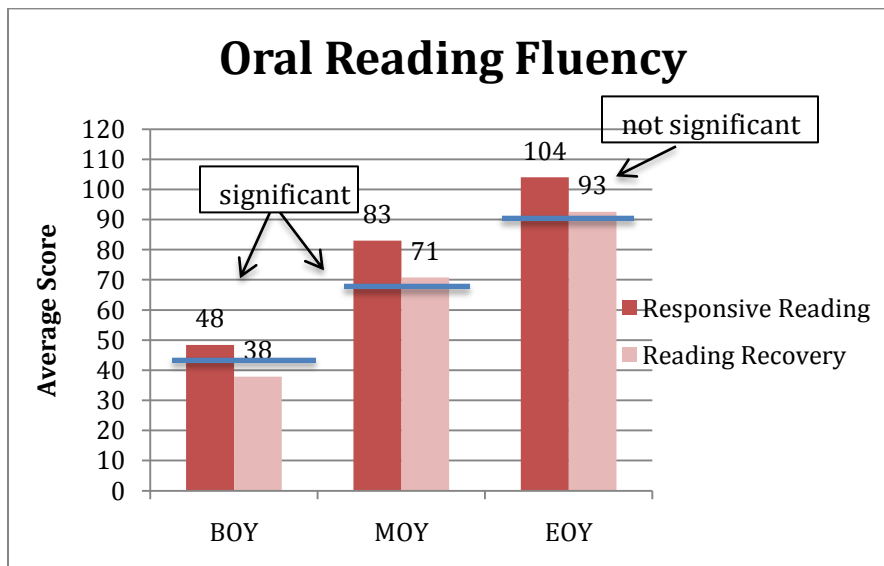
winter ORF scores. This indicates that there were significant differences between groups for the DIBELS performance on ORF in the beginning and middle of second grade.

Figure 21. Second Grade Nonsense Word Fluency Scores



Note. Blue lines indicate benchmark goals for each assessment point.

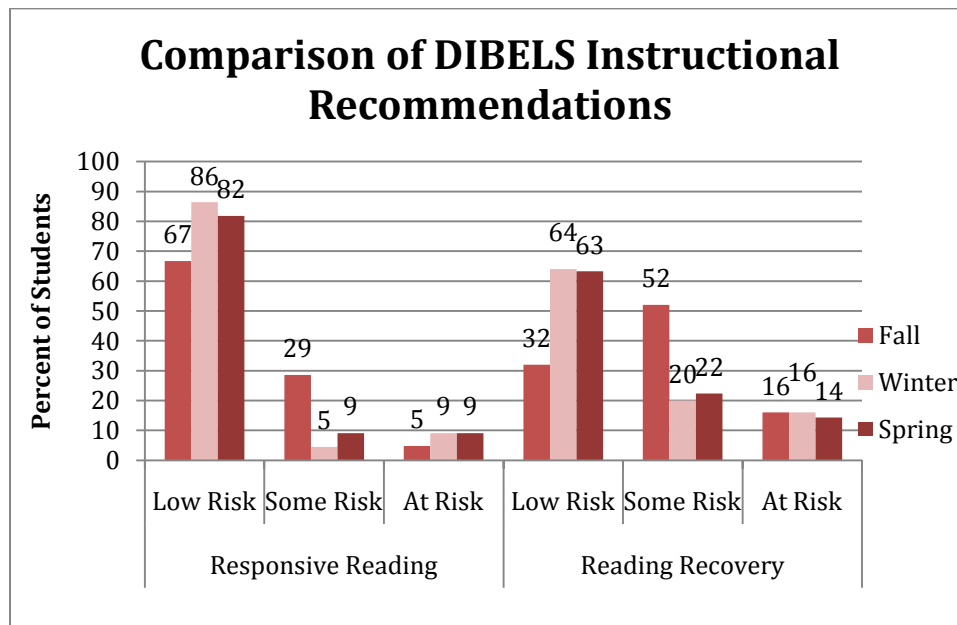
Figure 22. Second Grade Oral Reading Fluency Scores



Note. Blue lines indicate benchmark goals for each assessment point.

An examination of the DIBELS Instructional Recommendations found that the Responsive Reading group had a higher percentage of students enter second grade at *Low Risk* as well as have a higher percentage of students continue as *Low Risk* across the year (see Figure 23). This group also had a lower percentage of *At Risk* students at the beginning of second grade and throughout the year than found among the Reading Recovery group. The change in students' risk levels over the summer was particularly noteworthy. Although both groups displayed decreases in the percentage of *Low Risk* students and increases in the percentage of *At Risk* students from the end of first grade to the beginning of second grade, the students in the Responsive Reading group had a smaller decrease in *Low Risk* students and a smaller increase in *At Risk* students. With Responsive Reading, the students ended first grade with 87% of students at *Low Risk* and 0% of students *At Risk*, and began second grade with 67% of students at *Low Risk* and 5% of student *At Risk*. In comparison, Reading Recovery students ended first grade with 71% *Low Risk* students and 8% *At Risk* students, then began second grade with 32% *Low Risk* students and 16% *At Risk* students.

Figure 23. Second Grade DIBELS Instructional Recommendations By Group



In conclusion, both groups displayed gains in DIBELS scores across second grade. However, the performance of students in second grade on the ORF measure was significantly higher for the Responsive Reading students in the fall and winter. No significant differences between groups were found for the spring ORF score or for the fall NWF score. An examination of instructional recommendations found that the Responsive Reading group had a substantially higher percentage of *Low Risk* students at the beginning of second grade and across the year. In addition, the percentage of *At Risk* students was lower among the Responsive Reading students throughout the year. The

change in risk levels across the summer was particularly noteworthy, with the Responsive Reading students having a smaller decrease in *Low Risk* students and a smaller increase in *At Risk* students from the end of first grade to the beginning of second grade.

Conclusions

The purpose of this evaluation was to determine whether students who participated in the first grade Responsive Reading program gained significantly more literacy skills than students who participated in the Reading Recovery program. To that end, the evaluation examined four questions: (1) Do students demonstrate differences in selected literacy skills at the completion of the intervention? (2) Do students demonstrate differences in selected literacy skills at the end of the year? (3) Are there group differences in student achievement as measured by the DIBELS? (4) Are there group differences in student performance during second grade?

Overall, these findings suggested that while both intervention programs lead to gains over time, students in the Responsive Reading program gained significantly more literacy skills that helped them in the short-term (i.e., immediately following the intervention) as well as in the long-term (i.e., at the end of the year) during first grade. Students in Responsive Reading also had significantly higher end-of-year performance in Oral Reading Fluency as measured by DIBELS. Moreover, a higher percentage of students in the Responsive Reading intervention as compared to the Reading Recovery students had *Low Risk* instructional recommendations based on DIBELS performance in first grade, although the Responsive Reading students also began first grade with a similar pattern. A follow-up of students in second grade also found that the students who had participated in Responsive Reading had significantly higher scores for Oral Reading Fluency in both the beginning and middle of the year. In addition, there was a higher percentage of *Low Risk* students and a lower percentage of *At Risk* students throughout second grade among the students who had participated in Responsive Reading. The change in students' risk levels over the summer was particularly noteworthy, with the students from the Reading Recovery group demonstrating a greater increase in the percentage of *At Risk* students from the end of first grade until the beginning of second grade.

In conclusion, although both intervention programs resulted in substantial gains in low achieving first graders, the Responsive Reading program provided students with significantly more skills than the Reading Recovery program.