



**YEAR THREE REPORT: EVALUATION STUDY OF
THE WRITING ROAD TO READING**

Second Grade Study

Gary Bitter, Mary Aleta White

Arizona State University

September, 2009

**MARY LOU FULTON COLLEGE OF EDUCATION
TECHNOLOGY BASED LEARNING AND RESEARCH
ASU SkySong
1475 N. Scottsdale Rd., Suite 200 Scottsdale, AZ 85257-3538
Phone # (480) 884-1694 Fax # (480) 884-1888
<http://tblr.ed.asu.edu>**

Table of Contents

List of Tables	3
INTRODUCTION	4
RESEARCH DESIGN	4
METHOD	
Teacher Measures	4
Student Measures	5
Participants	5
PROGRAM IMPLEMENTATION, RESULTS, & DISCUSSION	
Implementation by Experimental Teachers	8
Student Performance Results	9
Connecting Teachers to Students	16
SUMMARY	16
REFERENCES	17

List of Tables

Tables		Page
Table 1	Schools Included in the Analysis	5
Table 2	Student Distribution by Student Level Variables	7
Table 3	Demographics of the Experimental and Control Groups (percentages)	7
Table 4	Second-Grade Students From Year 1 By School	8
Table 5	Number of schools administering subscales, Fall	9
Table 6	Comparative Mean Scores of Spalding and Control Second-Grade Students on the DIBELS (Fall 2008, Winter 2009, Spring, 2009)	10
Table 7	Comparative DIBELS Statistics of the Spalding and Control Second-Grade Students (Fall 2008, Winter and Spring, 2009)	11
Table 8	Interpretations of effect size	12
Table 9	Second Grade Students' Mean Post-Scores on DIBELS	13
Table 10	DIBELS Cut Scores for Bench Marks	13
Table 11	Comparative Mean Scores of Spalding Kindergarten and First-Grade Students, Year 1 and Year 2	14
Chart 1	Student NCE reading scores	15
Table 12	Demographics of Participating Teachers	16

YEAR THREE REPORT: EVALUATION STUDY OF THE WRITING ROAD TO READING

Gary Bitter, Mary Aleta White
Arizona State University

INTRODUCTION

The research project was initiated in 2006-2007 to validate the effectiveness of Spalding's *Writing Road to Reading* in teaching reading to children from varied backgrounds who attend different types of schools.

RESEARCH DESIGN

Arizona State University conducted the third of a four-year quasi-experimental study for Spalding in the 2008-2009 school year. This study involved 44 teachers and 996 students in the final study sample. In this third year, the study followed the 2006-2007 kindergarten students into second grade in the same 5 experimental and 6 control schools. The following research questions continue to guide the study design and methods:

1. Do children who participate in *Spalding's Writing Road to Reading* program demonstrate significant learning gains in reading skills?
2. How does the reading skill attainment of children participating in Spalding's *Writing Road to Reading* program compare to that of children participating in other core reading programs?
3. How well do teachers implement Spalding's *Writing Road to Reading* program in their varied classrooms?

METHOD

This section presents descriptions of the different study components including measures, participants, and program procedures.

Teacher Measures

The study evaluated teachers' implementation of Spalding's *Writing Road To Reading*. Researchers utilized a uniform quantitative instrument to measure how *The Writing Road to Reading* was being implemented in the experimental classrooms. In order to measure program implementation, researchers collected data through classroom observations using observation protocols. Four researchers, in teams of two or as a whole group, visited the teacher classrooms three times per year and observed individual teachers to ensure inter-observer agreement and reliability. The observation protocol was designed to measure constructs such as classroom management, adherence to program philosophy, and strategies for spelling, writing, and reading content. Classroom observations lasted for approximately 45 minutes to one hour and focused on whole group instruction.

Both experimental and control teachers completed a survey questionnaire that provided a variety of background information including degrees, certifications, endorsements, and professional development activities over the past ten years. Other items included length of time implementing reading programs, materials used, assessment practices, and the number of years teaching at the current grade level.

Student Measures

Researchers employed the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) as the primary measure to assess changes in students' reading skills during the 2008-2009 school year. Researchers selected the DIBELS assessment because it has broad visibility, acceptance in the field, and it demonstrates high technical merit. The Arizona State Department of Education has adopted DIBELS as the assessment for its Reading First program. Administration periods for DIBELS occurred at the beginning, middle, and end of the school year, and subtests are designed for administration across multiple years. The DIBELS measures used in this study were primarily Oral Reading Fluency, and Retell Fluency. Some, but not all, schools tested for Word Use Fluency. All participating students were first tested within the first three weeks of the 2008-2009 school year as required by the Arizona State Department of Education, again in December, 2008 or January, 2009, and finally in May or early June, 2009.

Participants

This study was conducted in 11 diverse Arizona schools with a total of 991 participating second-grade students at the first benchmark test, and increasing slightly to 996 total students by the year-end test. The experimental students can be further separated into two groups: the first group includes those students who were also in the kindergarten study (total of 351 at year-end), and the second group was all students in the grade level (540 at year end or 189 new experimental students by year-end). A quasi-experimental design was used to assign schools as a control or experimental school. Schools were matched on socioeconomic status of students, class size, race/ethnicity/gender of students, and geography. Table 1 lists the study schools, number of classes, and number of students at year-end.

Table 1: *Schools Included in the Analysis*

Group	Name of School	# Total 2nd graders	# of classrooms
Experimental	Alhambra	76	3
	Bret Tarver	128	5
	CTA-Liberty	141	5
	Gallego	90	4
	Valley Academy	105	4
Total	5	540	21
Control	#1 – M	103	5
	#2 – N	104	5
	#3 – O	84	4
	#4 – P	22	1
	#5 – Q	111	5
	#6 - R	32	3
Total	6	456	23

Classes in the experimental condition (21) used the Spalding curriculum an average of 90 minutes each day, while control classes (23) used their standard reading programs (described below on page 9). Twenty-one first grade teachers participated in the study as an experimental group, while 23 teachers were in the control group for a total of 44 teachers. Class size averaged 26 students in the experimental group and 20 in the control group. As an incentive for participation, experimental teachers received materials and training in *The Writing Road to Reading* program without charge. Control teachers each received \$200 gift certificates to a bookstore for classroom materials.

There is a slight decrease in the control sample size between Year 1 and Year 3. It is the result of a reorganization of one control school which had 4 classrooms averaging 23 students in previous years. The school now has a single second-grade classroom of only 22 students. The attrition is the result of a variety of environmental factors.

As a result of the school's diminishing enrollment, the District has decided to convert the control school site into a traditional school during the 2009-2010 school year. This includes instituting Spalding as the reading method. Project researchers met with school district staff and received approval to "substitute" another control school with very similar demographics in the fourth and final year of the study. The district's Research Director has recommended another, similar school with the following student demographics:

Control School #4 (Original)

Girls:	53%
F/R:	84%
ELL:	69%
% Minority:	81%

Control School #4 (Replacement)

Girls:	49%
F/R:	93%
ELL:	52%
% Minority:	93%

Additionally, a number of students from the original school are expected to have enrolled in the other two, original control schools within the district. Fortunately, the district is also able to provide DIBELS data for the replacement control school from the beginning of the program study (in addition to year 4) to be used as needed in the final research report.

Table 2 presents the student distribution information for the experimental and control groups at the start of each school's academic year. The school districts provided early data for 1,004 students: 527 in the experimental group and 477 students in the control group. Table 3 (below) provides the same data but uses percentages to demonstrate the student demographics.

Compared to Year 1 and Year 2 data, there has been a steady decline in the number of students identified as English Language Learners in both the experimental and in the control groups. In this past year, the difference between the two groups on this variable is less than 4%. On the other hand, the SES measure, free and reduced lunch has increased in control schools from the

first year, so that overall 60% of participating students in the control group qualified as low-income compared to 38% for the experimental group. This differential has been constant and has widened since Year 1. The overall percentage of minority students by group has been higher in the experimental group during Years 1 and 3.

Table 2: Student Distribution by Student level variables

		Experimental (n = 527)	Control (n = 477)	Overall (n = 1,004)
Gender				
	Female	258	240	498
	Male	269	237	506
Ethnicity				
	Asian	45	22	67
	Black	22	25	47
	Hispanic	256	258	514
	Native Am	15	12	27
	White	189	160	349
SES				
	F/R	127	238	365
Language Ability				
	ELL	126	160	286

Table 3: Demographics of the Experimental and Control Groups (percentages)

Group	% Girls	% ELL	% Hispanic	% Minorities	% F/RL
Experimental	48.7	24.6	51.4	66.8	38.1
Control	50.0	27.3	47.9	62.3	59.9

The declining ELL population may be the result of the implementation of two recent state laws. One recent Arizona law that became effective in January, 2008 gave the state the authority to suspend or revoke the business license of employers who intentionally or knowingly employ workers who are unauthorized to work in the United States. For the most part, in Arizona, this means Mexican immigrants. The Pew Hispanic Center estimates there are 500,000 illegal immigrants in Arizona, and they make up about 9 percent of the state's population. Illegal immigrants make up 10 percent to 12 percent of the work force, according to Pew and the Center for Immigration Studies. A large number of immigrants left the state once the law took effect. They withdrew a number of students, many classified as ELL. In addition, the state's procedure for educating ELL students has undergone a tremendous revision in recent years. A federal district judge in Arizona ruled in 2000 and a U.S. District Judge in 2007 ruled that the state was not adequately educating or funding the learning of ELL students.

As a result of those rulings, the Arizona English Language Learners (ELL) Task Force was established and charged with developing and adopting research based models of structured English immersion (SEI) programs. Arizona Revised Statutes (A.R.S.). §15-756.01, requires that

the SEI models include a minimum of four hours per day of English language development (ELD) for the first year in which a pupil is classified as an English Language Learner.

The Arizona ELL Task Force adopted the new SEI Program Models that was then implemented by all Arizona school districts beginning this year (08-09), requiring all ELL students to be placed in 4 hours of ELD instruction daily. A recent notice from the Arizona Department of Education Superintendent’s Office states that this will continue into the 2009-2010 school year.

Both of these changes have had a unique impact on the number of ELL students enrolled and on their educational process within the targeted schools.

Longitudinal Student Analysis

The study is now in the third of its four years. Table 4 lists the experimental study schools, the total number of study students at the final benchmark, and the kindergarten cohort. As described below, to better track their progress, additional analysis was conducted on the 351 second-grade students who were in the original kindergarten cohort reported in the 2006-2007 study.

Table 4: *Second grade students from year1, by school, at Spring 2009 benchmark*

Group	Name of School	Total 2 nd graders Spring, 2009	Total Kindergarten cohort (06-07)	# 2 nd graders from initial kindergarten cohort	Retention rate from Year 1
Experimental	Alhambra	76	89	58	65%
	Bret Tarver	128	123	60	49%
	CTA-Liberty	141	114	89	78%
	Gallego	90	95	65	68%
	Valley Academy	105	117	91	78%
Total	5	540	538	363	67%

PROGRAM IMPLEMENTATION, RESULTS, & DISCUSSION

This section presents the results of the classroom implementation study, the student assessment scores, and an analysis of teacher characteristics as they relate to student achievement.

Implementation by Experimental Teachers

The classroom observations were the primary measure for program implementation. The goal for observations is to see consistent Spalding instruction across grade levels and schools. Observers noticed an increase in consistency within and across the five schools by the fourth-quarter observations. This is attributed to consistent use of the *Teacher Guides*.

In terms of experimental teachers’ performance, the year-end researchers’ overall observation protocol results showed that at least 81% of program practices were satisfactorily implemented by experimental teachers with 11% of experimental teachers’ behaviors needing further

refinement. The final observation summary showed that in the area of program philosophy and spelling, most teachers were successfully adhering to *The Spalding Method*. In program philosophy, observers noted a number of teachers who did not encourage higher-level thinking, particularly in the reading example. Observers also saw a decline in the number of classrooms where children read a decodable book in unison. In all instances this was because of poor time management. The observation schedule places teachers on a shortened schedule (1 hour) that creates the time constraints.

In the Spelling component, observers noted a decline in the number of teachers who required students to evaluate their handwriting based on the day’s focus. In the Writing component, there was a similar decline in the number of teachers who modeled one additional objective from the *Teacher’s Guide*. Last, in the Reading component, observers noted some difficulty in coaching student’s use of the five mental actions.

According to control teacher questionnaires, all second-grade control schools used either Houghton or Harcourt reading program. These programs were evaluated by the Arizona Department of Education as core reading programs under Reading First. Control teachers received from 2 to 5 hours of inservice training on these publishers’ materials.

After the National Reading Panel Report in 2000, most publishers hastened to add the five reading research-based components. Since both the control groups and the experimental groups used detailed teacher guides evaluated by NCLB for the reading research components, they should produce similar results. One major difference between the teacher programs, however, is in the training. The teacher methodology in the experimental groups differed from that of the control groups because it requires more hours (45 compared to an average of 25).

Student Performance Results

The DIBELS tests administered to second-grade students in Fall 2008 were for Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), Word Use Fluency (WUF), Oral Reading Fluency (ORF) and Retell Fluency (RTF). The DIBELS tests administered to second-grade students in Winter and Spring 2009 were for Word Use Fluency (WUF), Oral Reading Fluency (ORF), and Retell Fluency (RTF). The chart below indicates the number of schools that administered the DIBELS subtests.

Table 5: Number of schools administering DIBELS subscales, Fall, 2008 – Spring, 2009

Group	PSF	NWF	WUF	ORF	RTF
Experimental	0	5	3	5	4
Control	1	6	4	6	4

Because of the incomplete data, PSF results were deleted from the Fall, 2008 analysis, and WUF results were deleted from the Winter and Spring analysis.

Table 6 displays the comparative performance of the Spalding and the control students on the DIBELS that were administered in the Fall of 2008, Winter and Spring of 2009.

Table 6: *Comparative Mean Scores of Spalding and Control Second-Grade Students on the DIBELS (Fall 2008, Winter 2009, Spring, 2009)*

		Experimental	Longitudinal Group	Control	Difference
Fall, 2008	NWF	84.88*	87.05	67.27	17.61
	WUF	44.50*	44.90	32.59	11.91
	ORF	73.63*	77.06	49.71	23.92
	RTF	28.59*	29.74	15.42	13.17
Winter, 2009	ORF	98.59*	101.65	66.33	32.26
	RTF	38.72	41.09	25.39	13.33
Spring, 2009	ORF	109.96**	111.17	87.48	22.48
	RTF	44.83**	45.82	33.47	11.36

*p<.05

**p<.01

Similar to last year, Spalding students had consistently higher *mean* values on all DIBELS areas, which provides preliminary evidence that Spalding has been more effective than other methods used in the control schools in teaching those reading skills. The students who have participated in the study since Year 1 have scores that are higher than or equal to the overall experimental students. Table 7 provides further descriptive statistics on the tested areas. Students in both groups improved in reading skills by the end of year two; however, in every category the experimental group students had higher mean scores than control group students by an average of 14.6 points.

Table 7: Comparative DIBELS Statistics of the Spalding and Control Second-Grade Students (Fall 2008, Winter and Spring, 2009)

	Fall, 2008				Winter 2009		Spring, 2009	
	NWF	WUF	ORF	RTF	ORF	RTF	ORF	RTF
Spalding, Experimental Schools								
N, Valid	527	285	527	388	535	397	512	375
Mean	84.88	44.50	73.63	28.59	98.59	38.72	109.96	44.83
Std. Devia	34.36	15.76	37.93	15.76	38.01	17.36	37.58	17.47
Skewness	.422	.572	.542	.465	.076	.548	-.054	.397
Kurtosis	-.807	1.285	-.245	-.579	-.149	.622	-.230	.039
Max	198	121	188	75	212	112	222	92

	Fall, 2008				Winter 2009		Spring, 2009	
	NWF	WUF	ORF	RTF	ORF	RTF	ORF	RTF
Spalding, Longitudinal Group								
N, Valid	355	212	355	266	361	272	363	274
Mean	87.05	44.90	77.06	29.74	101.65	41.09	111.17	45.82
Std. Devia	35.37	14.85	38.20	15.51	37.07	16.99	37.30	17.08
Skewness	.353	.214	.530	.446	.135	.620	-0.35	.397
Kurtosis	-.948	-.365	-.389	-.653	-.094	.809	-.189	-.098
Max	198	82	187	68	212	112	222	91
Control Schools								
N, Valid	442	259	463	260	432	255	439	244
Mean	67.27	32.59	49.71	15.42	66.33	25.39	87.48	33.47
Std. Devia	33.59	15.74	34.78	12.23	38.03	16.71	42.73	21.03
Skewness	.623	.230	1.014	1.186	.678	1.203	.387	1.403
Kurtosis	-.336	.500	.925	1.811	.296	2.130	.095	2.463
Max	142	89	205	65	208	94	251	113

In addition to measures of statistical significance, researchers frequently calculate and report measures of practical significance, known as the effect size. The effect size is a way to help educators decide whether a statistically significant difference between programs translates into a meaningful difference—one that would justify a program adoption for instance. There are different ways to measure effect sizes. One commonly used measure is called Cohen's d. The formula is below:

$$\text{Cohen's } d = \frac{[M_1 - M_2]}{S.D._{\text{control}}}$$

The equation for pooled SD is:

$$SD_{\text{pooled}} = \sqrt{\frac{(N_1 - 1)S.D._1^2 + (N_2 - 1)S.D._2^2}{N_1 + N_2 - 2}}$$

The formula measures effect sizes in standard deviation units, and the results are commonly interpreted as follows:

- Small effect sizes: $d = .2$ to $.5$
- Medium effect sizes: $d = .5$ to $.8$
- Large effect sizes: $d = .8$ and higher

Table 8 shows conversions of effect sizes to percentiles. Effect sizes can also be thought of as the average percentile standing of the average treated (or experimental) participant relative to the average untreated (or control) participant. An ES of 0.0 indicates that the mean of the treated group is at the 50th percentile of the untreated group. An ES of 0.8 indicates that the mean of the treated group is at the 79th percentile of the untreated group. An effect size of 1.7 indicates that the mean of the treated group is at the 95.5 percentile of the untreated group.

Table 8: *Interpretations of effect sizes*

Cohen's Standard	Effect Size	Percentile Standing
	2.0	97.7
	1.9	97.1
	1.8	96.4
	1.7	95.5
	1.6	94.5
	1.5	93.3
	1.4	91.9
	1.3	90
	1.2	88
	1.1	86
	1.0	84
	0.9	82
LARGE	0.8	79
	0.7	76
	0.6	73
MEDIUM	0.5	69
	0.4	66
	0.3	62
SMALL	0.2	58
	0.1	54
	0.0	50

Cohen's *d* using a pooled standard deviation was computed for DIBELS ORF scores at the end of each benchmark assessment. The effect size for Fall 2008 was .7; for Winter 2009 it was .8; and in Spring 2009 the effect size is .6. This means that the intervention has a positive, medium effect on student achievement and is a more effective reading program. Using the chart above, that would mean the average student in the Spalding sample, at the end of the year, would score higher than 73% of the control sample.

As shown in Tables 9 and 10 below, additional analyses of the extent to which experimental students experienced learning gains by the end of second grade as well as between the beginning and the middle of the school year show that they exceeded the DIBELS decision rules benchmarks for achievement at each testing period.

Table 9: *Second-grade Students' Mean Post-Scores on DIBELS*

		Mean Test Scores			
		Spalding Experimental Schools	Spalding Longitudinal Group	Control Schools	DIBELS End of Second-grade Low risk score
		(n = 512)	(n = 363)	(n = 439)	
Oral Reading Fluency (ORF)					
	Spring test	109.96	111.17	87.48	90

As can be seen, Spalding participants experienced significant gains in reading performance from the beginning, to the middle and end of the school year. Unfortunately, by the middle and ending benchmarks in second grade, the average control student is not meeting the DIBELS assessment for low-risk scores.

Table 10: *DIBELS cut scores for bench marks*

		ORF
Beginning		
	DIBELS Benchmark	44
	Experimental	73.63
	Spalding Longitudinal	77.06
	Control	49.71
Middle		
	DIBELS Benchmark	68
	Experimental	98.59
	Spalding Longitudinal	101.65
	Control	66.83

Table 10 (cont'd): *DIBELS cut scores for bench marks*

End		
	DIBELS Benchmark	90
	Experimental	109.96
	Spalding Longitudinal	111.17
	Control	87.48

In an effort to examine threats to the study’s internal validity (such as the Hawthorne effect), an additional analysis was conducted to compare the 2008-2009 kindergarten and first grade student achievement with that of the prior year cohort. During the year, the kindergarten and first-grade students and teachers were not participants in the study; however, data was collected on a sample of students. The analysis checked for inflated scores during the year of study as compared to a “normal” year.

Although kindergarten and first grade teachers and classrooms were not the grade level under study, student performance slightly exceeded the original cohort scores. The chart below provides mean scores on the final testing categories. The schools were identically matched.

Table 11: *Comparative Mean Scores of Selected Spalding Kindergarten and First-Grade Students on the year-end DIBELS (Spring, 2007 = Year 1; Spring, 2008 = Year2; Spring 2009 = Year3)*

		Year 1 ('07) Kindergarten	Year 2 ('08) Kindergarten	Year 3 ('09) Kindergarten	Difference Year 1-3
Year End Testing Results	Letter Name	48.43	49.35	53.16**	4.73
	Phoneme Segmentation	50.09	51.46	49.60	-0.49
	Nonsense Word	45.92	50.41	50.44*	4.52
	Word Use	37.51	41.46	39.05	1.54

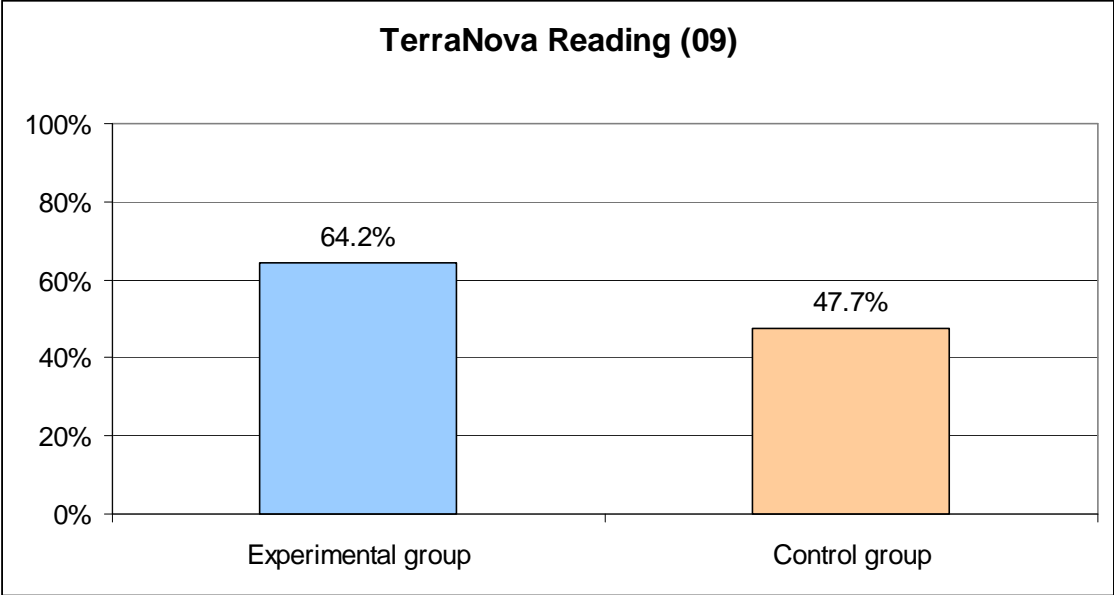
			Year 1 ('08) First Grade	Year 2 ('09) First Grade	Difference
Year End Testing Results	Phoneme Segmentation		51.30	56.47**	5.17
	Nonsense Word		86.32	93.32**	7.0
	Word Use		51.19	51.85	0.66
	Oral Reading Fluency		77.74	85.89**	8.15
	Retell Fluency		33.41	36.84	3.43

*p<.05

**p<.01

Another analysis of reading achievement was available in this year’s study because all second grade students are required to complete the state’s norm-referenced achievement test, TerraNova. The chart below represents a sample of the study students (three control and three experimental schools) and their average NCE score on the TerraNova reading portion. As would be expected from reviewing the DIBELS scores, the Spalding students’ NCE scores were significantly higher than the control students on the state test ($p < .01$).

Chart 1: *Student NCE reading scores from Spring, 09 AZ TerraNova exam*



Connecting Teachers to Student Data

As noted earlier, teacher survey data is available for both the experimental and control group. Demographic data from the surveys show that the average teacher age between the two groups is fairly close: 40 for the experimental group and 36 for the control group. The experimental group had slightly more teaching experience (9 versus 7). The following table displays additional demographics for the participating teachers.

Table 12: *Demographics of Participating 2nd-Grade Teachers*

Group	Average age	# Females	# Males	# Minority	# White	Avg # years teaching
Experimental	40	19	2	5	16	9
Control	36	21	3	5	18	7

Teacher experience was re-coded into five categories; namely, beginning or little (1-3 years of experience); some experience (4-6 years of teaching experience); experience (7-10 years), well experienced was 11-14 years, and highly experienced was above 15 years of teaching. The average ORF scores of students within the five categories vary widely. That is, additional years of teaching was not associated with higher levels of student assessment results. The same was true for the number of years using the Spalding method. Again, teachers were placed into five categories based on their years of using the Spalding method. Student ORF scores varied by category with no distinct pattern. Additional years of teaching Spalding were not associated with higher levels of student achievement. Effective teaching must rest on other principles rather than experience and the longevity of program usage. Based on the survey feedback data, compared to control group teachers, experimental teachers were required to devote a greater number of hours preparing to teach the reading method. Experimental teachers were also required to work on the integration of the language arts skills with consistent delivery of sequential, explicit instruction.

SUMMARY

According to the year three results, students who used *The Writing Road to Reading* continue to demonstrate statistically significant learning gains as measured by DIBELS. In addition, their scores were significantly higher than control group student scores again this year. Since both the control groups and the experimental groups used detailed teacher guides evaluated by NCLB for research-based reading components, theoretically, they should have produced similar results. This was not the case. These preliminary findings are strongly suggesting that use of *The Writing Road to Reading* curriculum is an effective method for enhancing performance on critical early literacy skills.

References

- Auckerman, R. C. 1984. *Approaches to beginning reading*. 2nd ed. New York: John Wiley and Sons.
- Ball, E. W., and B. A. Blachman. 1991. Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly* 26: 49-66.
- Beck, I. L., and C. Juel. 1995. The role of decoding in learning to read. *American Educator* 19: 8, 21-25.
- Biemiller, A. J. 1970. The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly* 6: 75-96.
- Byrne, B., and R. Fielding-Barnsley. 1993. Evaluation of a program to teach phonemic awareness to young children: A one year follow up. *Journal of Educational Psychology* 85: 104-111.
- Byrne, B., and R. Fielding-Barnsley. 1995. Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology* 87: 488-503.
- Cohen's D. Retrieved online, September 23, 2009. University of Colorado at Colorado Springs website, link: <http://web.uccs.edu/lbecker/Psy590/es.htm>.
- Collins, A., J., S. Brown, and S. E. Newman. 1989. Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick, ed. *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Cunningham, A. (1989). Phonemic awareness: The development of early reading competency. *Reading Research Quarterly, 24*(4), 474–72.
- Farnham-Diggory, S. 1987. From theory to practice in reading. Paper presented at the annual meeting of the Reading Reform Foundation, San Francisco, July.
- _____, S. 1990. *Schooling*. 124-131. Cambridge: Harvard University Press.
- _____. S. 1992. *Cognitive processes in education*. New York: HarperCollins.
- Fletcher, J. M. and G. R. Lyon. 1998. Reading: A research-based approach. In W. M. Evers, ed. *What's gone wrong in America's classrooms*. Stanford: Hoover Institution Press.
- Foorman, B., Francis, D., Fletcher, J., Schatschneider, C., and Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37–55.
- Foorman, B. R., D. J., Francis, S. E., Shaywitz, B. A. Shaywitz, and J. M. Fletcher. 1997. The case for early reading intervention. In B. A. Blachman, ed. *Foundations of reading acquisition and dyslexia: Implications for early intervention*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Good, R. H., Simmons, D. S., Kame'enui, E. J., Kaminski, R. A., & Wallin, J. (2002). *Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade* (Technical Report No. 11). Eugene, OR: University of Oregon.
- Good, R. H., Kaminski, R. A., Smith, S., Simmons, D. S., Kame'enui, E. J., & Wallin, J. (In press). *Reviewing outcomes: Using DIBELS to evaluate a school's core curriculum and system of additional intervention in kindergarten*. In S. R. Vaughn & K. L. Briggs (Eds.),

- Reading in the classroom: Systems for observing teaching and learning.* Baltimore: Paul H. Brookes.
- Goodman, K. S. and Y. Goodman. 1979. Learning to read is natural. In L. B. Resnick and P. A. Weaver, eds. *Theory and practice of early reading 1.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227.
- Kaminski, R. A., & Good, R. H. (1998). Assessing early literacy skills in a problem-solving model: Dynamic Indicators of Basic Early Literacy Skills. In M.R. Shinn (Ed.) *Advanced applications of curriculum-based measurement* (pp. 113-142). New York: Guilford Press.
- Laimon, D.E. (1994). *The effects of a home-based and center-based intervention on at-risk preschool children's early literacy skills.* Unpublished doctoral dissertation, University of Oregon.
- Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*(3), 263–284.
- Lyon, G. R. 2001. Statement before the House Committee on Education and the Workforce. Available from www.house.gov/ed_workforce/hearings/oeri5400/lyons.h+m.
- Moats, L. C. 2000. *Speech to print: Language essentials for teachers.* Baltimore: Paul H. Brooks Publishing Company.
- North Central Reading Educational Lab. *A closer look at the five essential components of effective reading instruction: A review of scientifically based reading research for*

teachers. Retrieved February 26, 2007, from NCREL Web Site:

<http://www.ncrel.org/rf/components.pdf>

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

North, M. E. 1991. The writing road to reading: From theory to practice. *Annals of Dyslexia* 42: 110-123.

_____. M. E. 1995. The effects of Spalding instruction on special education students. In C. W. McIntyre and J. P. Pickering, eds. *Clinical studies of multisensory structured language education*. Salem, OR: International Multisensory Structured Language Education Council.

Policy Maker's Primer on Education Research (2004). Retrieved online: September 23, 2009.

<http://www.ecs.org/html/educationIssues/Research/primer/understandingtutorial.asp>

Smith, F. 1971. *Understanding reading*. New York: Holt, Rinehart & Winston.

Wagner, R., Torgesen, J. K., & Rashotte, C. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities*, 27, 276-286.