EVALUATION OF THE WATERFORD EARLY READING PROGRAM IN KINDERGARTEN 2005-06

STEPHEN POWERS, PH.D. AND CONNIE PRICE-JOHNSON, M.A.
CREATIVE RESEARCH ASSOCIATES
2030 E. BROADWAY, SUITE 221
TUCSON, ARIZONA 85719
(520) – 884-8667
SVPOWERS@AOL.COM

JULY 2006

ACKNOWLEDGEMENTS

We would like to express our appreciation to the Tucson Unified School Districts's Office of Accountability and Research for their cooperation in retrieving relevant data from their database and for their expertise in merging and organizing the data for our analysis. This evaluation study would not be possible without their cooperation.

Stephen Powers, Ph.D. and Connie Price-Johnson, M.A.
Creative Research Associates, Inc.

ABSTRACT

Background: The Waterford Early Reading Program (WERP), a technology-based program for early elementary grades, was provided through Arizona all day kindergarten funds to kindergarten students in 15 Title I elementary schools in the Tucson Unified School District (TUSD) in the 2005-06 school year. The purpose of this study is to evaluate the reading achievement of kindergartners in the WERP schools and in a Comparison group of 15 schools in the same district.

The schools where the WERP was implemented are identified in this report as Schools A-L. The comparison schools are identified as Schools M-AA.

Research Design: This evaluation design was a comparison-group study (quasi-experimental design) involving a treatment (WERP) implemented in 15 Title I schools ranked with the highest percentages of students on free/reduced lunch. A Comparison group of 15 schools was selected from those with the next highest percentages of students on free/reduced lunch. The comparison schools did not receive the WERP.

Both matching techniques and statistical controls were used to make the groups similar in the analysis.

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency, Letter Naming Fluency, Word Use Fluency, Phoneme Segmentation Fluency, and Nonsense Word Fluency and the district's Core Curriculum Standard Assessment (CCSA) Reading Test were given as pretests and posttests during the school year. In addition, the amount of time that each kindergartner used the WERP computer software was extracted from the software and used in the analysis.

Statistical Analysis: Dependent samples t-tests were used to determine gains for the WERP and Comparison groups, and gain score analysis was used to compare these gains for the WERP and Comparison schools. Analysis of covariance was used to adjust the posttest means for differences on the pretest means of the students.

Data were disaggregated by school, gender, ethnicity, pretest achievement quartiles, primary home language, and English language learner (ELL) status in order to determine patterns of achievement among these groups.

Important Findings:

•The WERP kindergartners consistently outperformed the Comparison group kindergartners on all outcome measures. Comparison school kindergartners did make substantial and in some cases outstanding gains from pretest to posttest. However, when WERP kindergartners were compared with Comparison kindergartners, the WERP gains were substantially and significantly greater.

•Effect sizes of gains favored the WERP kindergartners, as well as effect sizes comparing the posttest achievement of the WERP kindergartners with the Comparison kindergartners.

•WERP gains were greater for males in the WERP program than for males in the Comparison group, and for females in the WERP than for females in the Comparison group.

•WERP gains were greater for Whites, Hispanics, African Americans, Native Americans, and Asians than for their counterparts in the Comparison group.

•WERP gains of White, African American, Hispanic, and Asian kindergartners were greater than the gains of White kindergartners in the Comparison group.

•WERP gains of kindergartners with a primary home language of English, Spanish, and other languages were greater than their counterparts in the Comparison group.

•WERP gains of kindergartners with a primary home language of Spanish were greater than the gains of English primary home language kindergartners in the Comparison group. That is, WERP Spanish home language students who were learning English reading skills outperformed the Comparison group English primary home language students.

•WERP gains of kindergartners in four different quartile levels of reading pretest achievement outperformed the Comparison students with the largest gains in the top (fourth) quartile.

•WERP English language learners outperformed Comparison group English language learners.

•WERP English language learners with emergent reading skills outperformed the non-English language learners (native speakers) in the Comparison group.

•Usage of the WERP software was found to be significantly correlated with the reading outcome measures and pretest to posttest gains in the outcome measures. It is an important findings that the greater the use of WERP content the greater the reading gains.

TABLE OF CONTENTS

TABLES

I. BACKGROUND AND PURPOSE

A. Background

The importance of early reading interventions has been argued by many researchers (Finn, 2001; NAEYC & IRA, 1998). Finn (2001) has noted the problems with an achievement gap especially among ethnic groups, and how this gap widens as the years pass. The value of technology in the early grades and its integration with instruction has been noted by many (NAEYC, 1996). Walberg (2001), a well-known evaluator, reported after reviewing the Waterford Early Reading Program (WERP) that it was "spectacularly effective for beginning readers who initially scored in the lower third of the group" (p. 11).

Relevant to the present study in the Tucson Unified School District (TUSD), the Waterford Institute (2002) had specified how the WERP addresses issues of the No Child Left Behind legislation in the major areas of emergent reading skills. The Waterford Institute (2002) also has specified in detail how the learning activities of the WERP address the skills assessed by the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), which is used in TUSD to assess reading in kindergarten.

The present evaluation focuses on the WERP, which was designed to teach children to read, write, and keyboard. It is research-based and uses technology integrated into learning activities. WERP was implemented in 15 schools of the Tucson Unified School District in the 2005-06 school year. It is from this year of WERP implementation that the data for the present study comes.

B. Purpose

The purpose of this study was to evaluate the effectiveness of the WERP in the kindergartens of TUSD Title I schools and to compare the pretest-to-posttest reading achievement of the kindergartners during the 2005-06 school year with that of 15 Comparison schools that did not receive the program.

II. METHODS

A. Study Setting

The TUSD school district is the largest school district in the Tucson area and the second largest in Arizona. It is a multiethnic school district with over 60,000 students, 3,700 teachers and over 200 administrators.

B. Study Population

The WERP was implemented in 15 Title I schools in TUSD with the highest rates of free/reduced lunch. A Comparison group of 15 schools was selected from those schools with the next highest free/reduced lunch rates.

WERP Schools. Schools with the WERP installed in kindergartens are identified in this report as Schools A-L.

Two additional schools in which the WERP was implemented were not selected for the study because it was impossible to extract the WERP usage data from their computers. Another school was originally slated to receive the WERP software but opted for the Waterford Early Math & Science program instead, so it also was not selected for the study.

Usage data showing the number of minutes each student used the WERP software was extracted from the WERP computers at the 12 remaining schools. Only kindergartners with sufficient exposure to the WERP program (1100 minutes or six months) were included in the analysis of most of the data. This excluded students from Schools I and L, bringing the number of schools to a total of 10 that were included in most of the analyses.

Comparison Schools. The Comparison group schools are identified in this report as Schools M-AA.

WERP software usage data was extracted for a total of 740 kindergarten students in the WERP schools, and 1480 kindergarten students participated in the Comparison schools. Only the 358 kindergartners with total usage of the WERP materials for a six-month period total or 1100 minutes were included in the analysis. After selecting only those students with both pretest and posttest, the number of students in the study was reduced to 334 students in the WERP schools and 1211 in the Comparison schools for a total of 1545. See Table 1.

Table 1. Kindergartners in the WERP Evaluation

Group	Total	1100 mins	Pre-Posttest
WERP	740	358	334
Comparison	1480	1480	1211
Total	2220	1838	1545

Note. WERP students were selected if they had used the WERP 1100 minutes (6 months) or more. Only students with both pretest and posttest were included in most analyses.

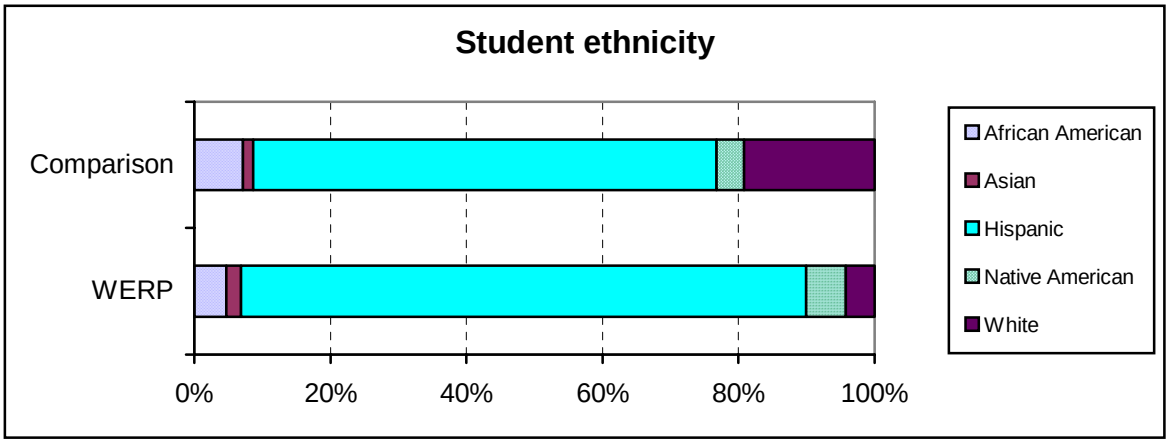
In the WERP group, 51% were males and 49% females. Of the Comparison group, 50% were male and 50% were female. The primary language of the WERP group was 48% English, 49% Spanish and 3% other languages. In the Comparison group, 68% used English as a primary language, 29% used Spanish, and 3% percent used another language. See Table 2 for the ethnic composition of the WERP and Comparison group schools.

Table 2. Ethnicity of Kindergartners in the WERP and Comparison Groups

Ethnicity	WERP		Comparison	
	N	%	N	%
African American	17	4.7	106	7.2
Asian	8	2.2	22	1.5
Hispanic	297	83.0	1008	68.1
Native American	21	5.9	60	4.1
White	15	4.2	284	19.2
Total	358	100.0	1480	100.0

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.

Figure 1. Ethnicity of Kindergartners in the WERP and Comparison Groups



C. Measurement of Outcomes

Usage minutes. The number of minutes of usage of the WERP by each student was collected directly from the computers by staff of Pearson Digital Learning, which markets the Waterford Institute's products. TUSD's Office of Accountability and Research matched these records with student test scores and eliminated personal identifiers before the records were analyzed in this study.

Only students with sufficient exposure to the WERP (i.e., 1100 minutes or six months) were used in the analysis. Table 3 shows the number of students who used Level 1, Level 2 or the Phonological Awareness component of the WERP for any amount of time, and the range of minutes a single student spent on that level.

Table 3. Kindergartners' Usage Minutes with the WERP

Usage	N	Minutes
Reading Level 1: total minutes in course	700	0 – 2175
Reading Level 2: total minutes in course	203	0 – 2585
Phonological Awareness: total minutes	704	0 – 962
Total of all usage minutes	725	0 - 4003

Dynamic Indicators of Basic Early Literacy Skills (DIBELS). The Dynamic Indicators of Basic Early Literacy Skills (DIBELS), developed by researchers and specialists in early childhood education at the University of Oregon, is a standardized assessment administered by TUSD to all kindergartners in the district three times a year and sent to the developers of the test for scoring. Scores are reported as raw scores and local percentiles. The DIBELS is composed of five subscales:

- Initial Sounds Fluency
- Letter Naming Fluency
- Word Use Fluency
- Phoneme Segmentation Fluency, and
- Nonsense Word Fluency scales.

Good and Kaminski (2002) reported psychometric research into the properties of the DIBELS. In summary, these authors report alternate-form and test-retest reliabilities and predictive and concurrent validities of the subscales to range from .36 to .91 with a median reliability of .66. The content of the subscales were carefully described and the constructs were also clearly described and related to the subscales so that one could conclude a high degree of content validity of these subscales. It was concluded that the DIBELS subscales were adequate for the present study.

The Waterford Institute (2002) provided a detailed analysis how the WERP activities were assessed by the DIBELS, as well as how the WERP addressed issues of the No Child Left Behind law.

For purposes of the present study, the average of the five DIBELS subscales was computed to provide an overall measure of the pretest and posttest reading achievement, or Total Reading Score, of the kindergarten students. The internal consistency (alpha) reliability of the test was .79. Only students who completed all 5 subscales were entered into the average (Total Reading Score).

Each of the DIBELS subscales was used to compute the percentile ranks within TUSD. These local percentile scores were used to obtain an estimate of the gains of students relative to other students within TUSD. The primary measure used to assess growth was the raw score of the DIBELS.

CCSA Reading Test. The Core Curriculum Standard Assessment (CCSA) Reading Test was developed by TUSD for district use. It parallels the criterion-based Arizona's Instrument to Measure Standards (AIMS) and is given in the grades where the AIMS is not. The CCSA places kindergartners in four levels of achievement (0, 1, 2, 3), which correspond to the AIMS levels of Falls Far Below, Approaches, Meets and Exceeds. Scores of 2 and 3 (Meets and Exceeds) are considered passing or mastery of the content. The CCSA was given in the fall and the spring by TUSD teachers, serving as a pretest and posttest along with the DIBELS.

Administration of Reading Measures in 2005-06.

Table 4. Administration of DIBELS and CCSA Reading 2005-06

Measure	Fall 2005	Winter	Spring 2006
DIBELS			
Initial Sounds Fluency	X	X	
Letter Naming Fluency	X	X	X
Word Use Fluency	X	X	X
Phoneme Segmentation Fluency		X	X
Nonsense Word Fluency		X	X
Total Reading Score	X		X
CCSA Reading			
Reading Performance	X		X

Other Student Characteristics. Other student characteristics considered in this analysis were:

- Student ethnicity
- Student gender
- Student primary home language, and
- Student English Language Learner (ELL) status.

D. Statistical Methods

Effect Estimates of the Intervention. Various methods were used to analyze the WERP and Comparison groups:

- Pretest-posttest gains of kindergartners in the WERP and comparison schools were analyzed using dependent samples t-tests. In addition, gain score analysis was used to determine if the gains of the WERP and Comparison groups were significantly different.

- Pretest–posttest gains of WERP and Comparison schools were estimated with effect sizes (ES). These analyses compared the pretest-posttest gains to the standard deviation of the pretest.

- Mean gain score analyses of the DIBELS Total Reading Score and Total Percentile for each of WERP and Comparison schools.

- Analyses of covariance (ANCOVA) and effect size analyses of WERP and Comparison schools. The ANCOVA adjust the posttest means on the reading test using the pretest as a covariate. After this, the posttest means of the WERP and Comparison schools were compared using effect sizes. Effect sizes are usually estimated comparing the treatment group (here the WERP) posttest mean with the Comparison group posttest mean. Effect sizes have been categorized (Cohen, 1977) as small (.20), medium (.50), and large (.80). These effect sizes indicate important differences due to a treatment effect.

Subgroup Analyses. In addition to the analyses of the WERP and Comparison groups as a whole, several analyses of subgroups were carried out:

- Three schools using the WERP and the Reading First programs during the 2005–06 school year were compared to Comparison schools with nearly the same pretest mean reading scores.

- Pretest to posttest gains of male and female kindergartners in the WERP and comparison schools were made.

- Pretest to posttest gains of African-American, Asian, Hispanic, Native American and White kindergartners in the WERP and Comparison schools were compared.

- Pretest to posttest gains of English, Spanish, and other primary home language kindergartners in the WERP and Comparison schools were compared.

- Pretest to posttest gains of kindergartners in four reading achievement quartiles of the WERP and Comparison groups were made in order to compare reading gains at different ability levels.

- Pretest to posttest gains of ELL kindergartners and non-ELL kindergartners in the WERP and Comparison groups were compared.

WERP Usage Effects. Correlational analyses between the total minutes of usage, reading achievement, and reading gains of the WERP students were completed to examine the relationship and effectiveness of the usage of WERP.

In addition, WERP students were categorized according to minutes using the WERP software. Their gains in reading were computed from pretest to posttest for each of the seven levels of usage.

III. RESULTS

A. Effect Estimates of the Intervention

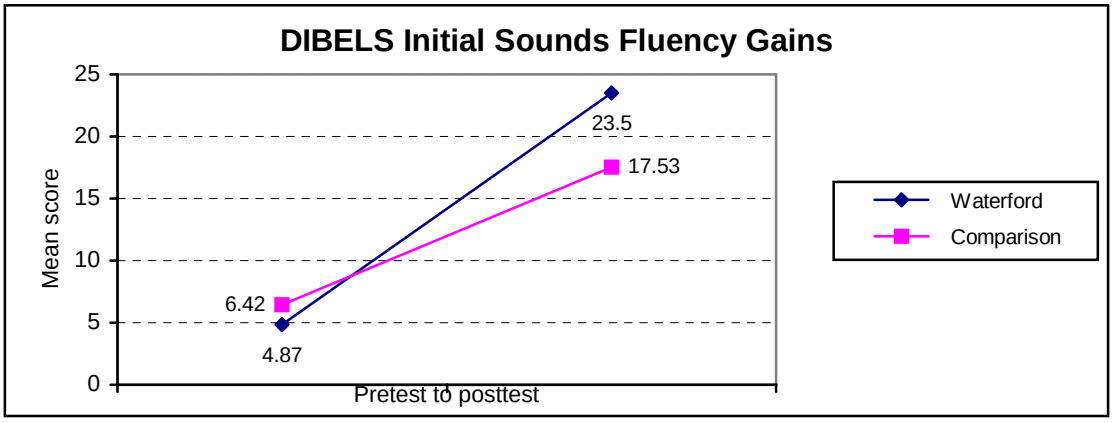
Table 5. WERP and Comparison Group Gains on All Outcome Measures

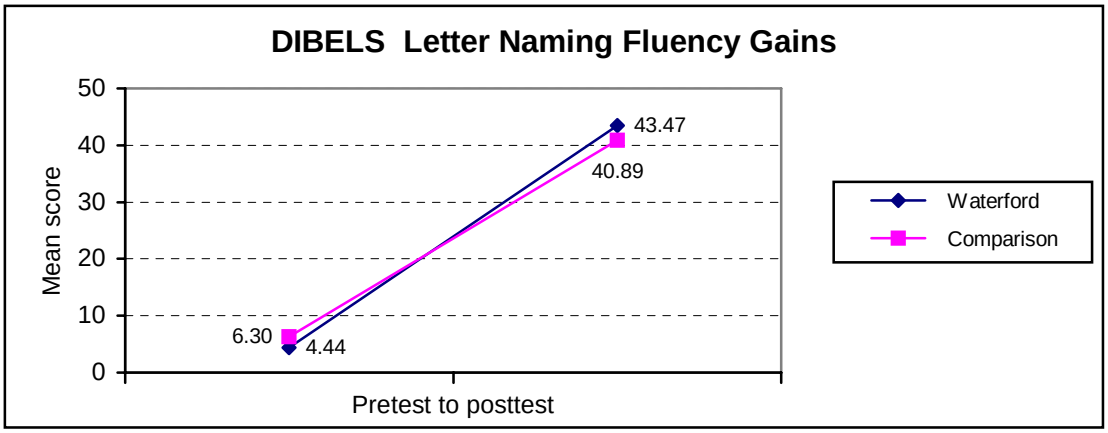
		Pretest		Posttest				
Measures	N	M	SD	M	SD	Gains	t	p
DIBELS: ISF								
WERP	334	4.87	5.71	23.50	13.88	18.63	24.87	.000
Comparison	1218	6.42	6.82	17.53	12.24	11.11	31.86	.000
WERP vs. Comparison						7.52***		
DIBELS: LNF								
WERP	334	4.44	8.17	43.47	16.32	39.03	44.92	.000
Comparison	1155	6.30	10.15	40.89	16.36	34.59	76.84	.000
WERP vs. Comparison						4.44***		
DIBELS: WUF								
WERP	325	3.57	7.15	32.93	20.53	29.36	26.85	.000
Comparison	998	4.94	10.40	32.34	20.87	27.40	40.55	.000
WERP vs. Comparison						1.96[a]		
DIBELS: PSF								
WERP	355	21.05	15.93	46.43	15.05	25.38	30.72	.000
Comparison	1219	17.10	15.87	39.34	18.69	22.24	46.70	.000
WERP vs. Comparison						3.14**		
DIBELS: NWF								
WERP	355	18.26	14.61	38.67	20.59	20.41	24.81	.000
Comparison	1217	14.66	15.02	31.46	20.21	16.80	38.21	.000
WERP vs. Comparison						3.61***		
DIBELS: Total Reading								
WERP	334	10.62	6.98	33.59	12.42	22.97	48.80	.000
Comparison	1211	10.22	8.29	29.32	13.33	19.10	71.46	.000
WERP vs. Comparison						3.87***		
TUSD: CCSA Reading								
WERP	311	1.09	0.49	2.68	0.64	1.59	38.00	.000
Comparison	1263	1.07	0.59	2.41	1.02	1.34	46.44	.000
WERP vs. Comparison						0.25***		

Note. ISF = Initial Sounds Fluency, LNF = Letter Naming Fluency, WUF = Word Use Fluency, PSF = Phoneme Segmentation Fluency, NWF = Nonsense Word Fluency. WERP students selected with 1100 minutes (6 months) or more usage of the Waterford Early Reading Program.

*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains. [a] p = .142

Figure 2. WERP and Comparison Group Gains on All Outcome Measures





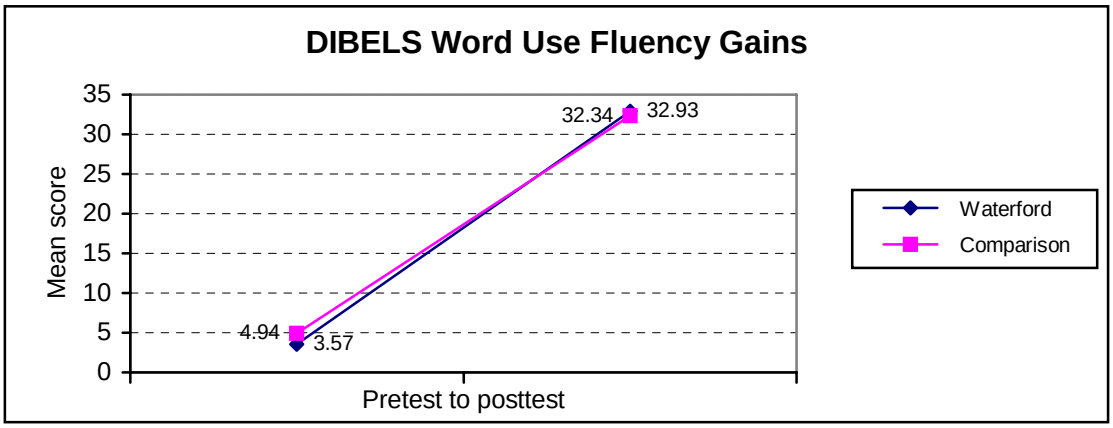
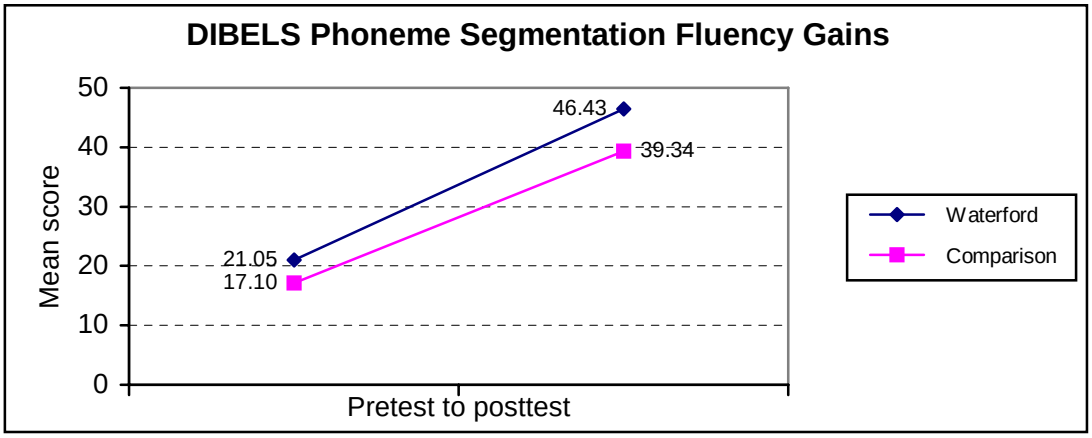
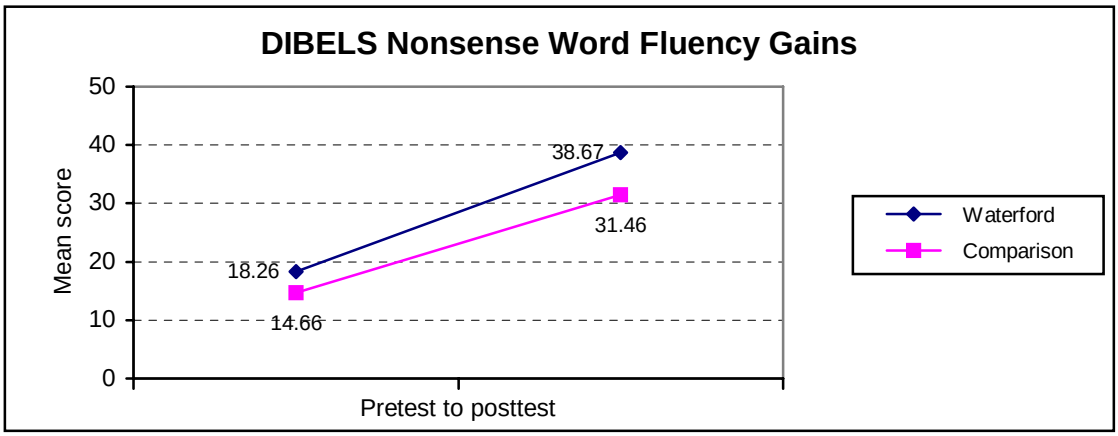


Figure 2. WERP and Comparison Group Gains on All Outcome Measures (continued)





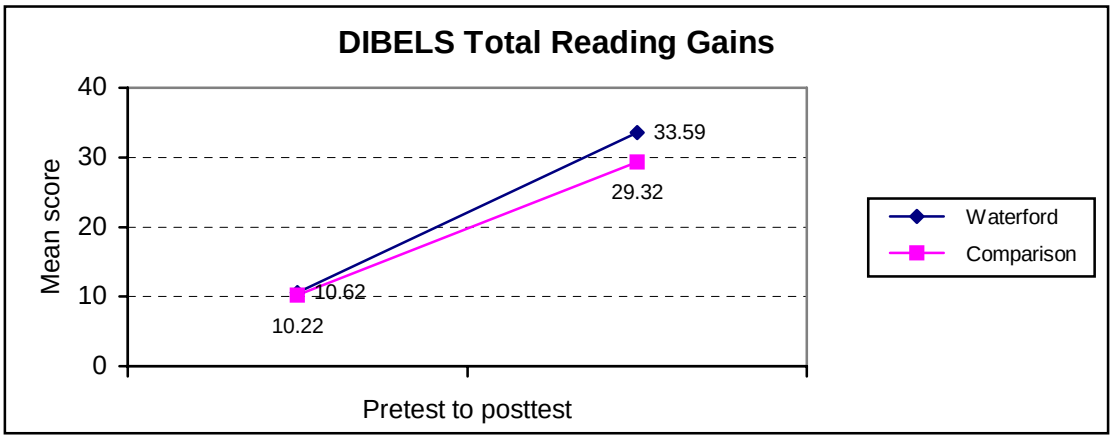
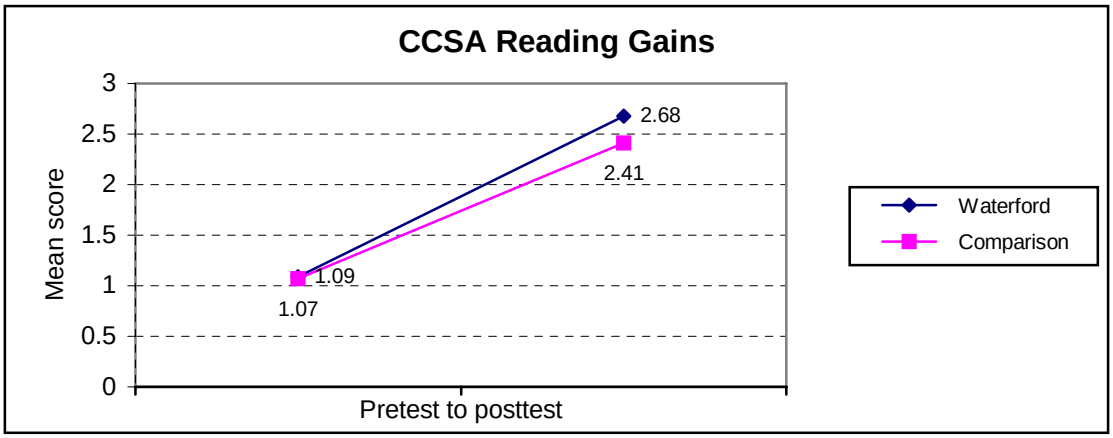


Figure 2. WERP and Comparison Group Gains on All Outcome Measures (continued)



Findings. On all of the outcome measures both the WERP and Comparison groups made highly significant gains ($p < .001$) from pretest to posttest. These significant gains are indicated by the "p" (i.e. probability) column of the table with the number .000. Generally any probability number less than .05 indicates statistical significance. Therefore it can be seen on this table that the pretest to posttest gains far exceeded the .05 level of significance.

On all of the outcome measures the WERP group gains were greater than the Comparison group gains. The largest gain was on the Initial Sounds Fluency where the WERP gain of 18.63 exceeded the Comparison group gain of 11.11 by 7.52.

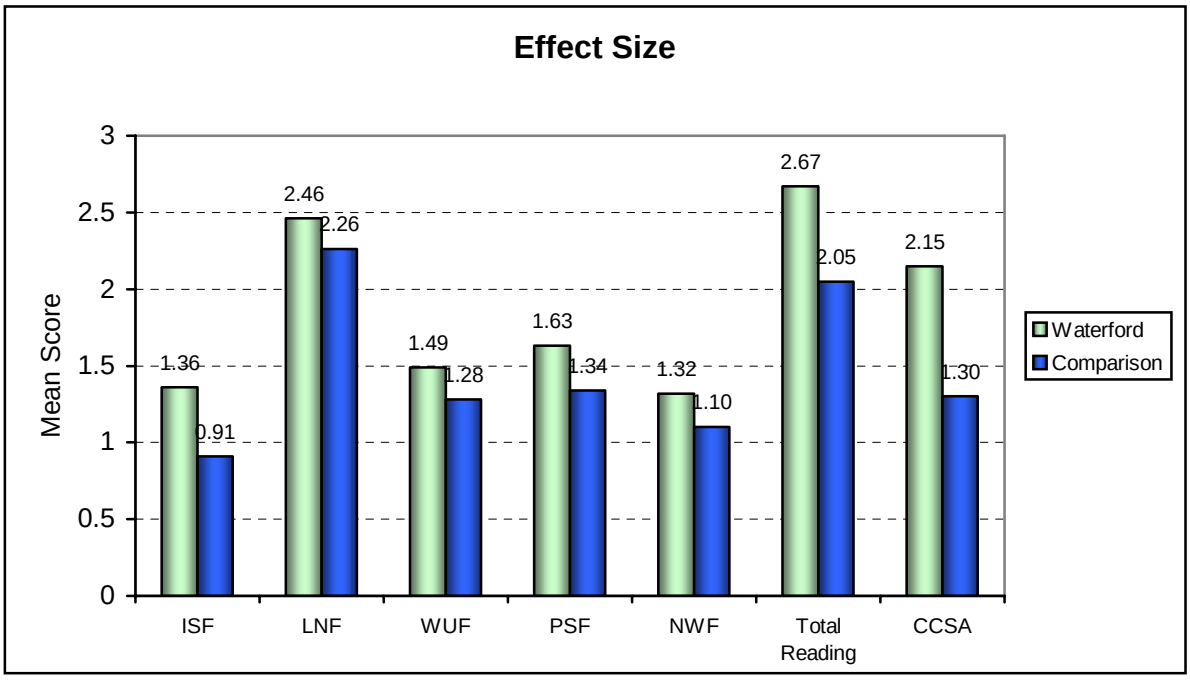
A pattern of gains emerges showing that the gains of WERP students, especially on Initial Sounds Fluency, Letter Naming Fluency, Phoneme Segmentation Fluency, Nonsense Word Fluency and the Total Reading Score, were clearly greater than the Comparison group gains. WERP students also made significantly greater gains on the CCSA than did the Comparison students.

Table 6. Effect Size (ES) of WERP and Comparison Schools on All Outcome Measures

Measures	Gains N	Gains M	PreSD	ES	Ratio
DIBELS: ISF					
WERP	334	18.63	13.70	1.36	1.49
Comparison	1218	11.11	12.17	.91	
WERP vs Comparison				.45	
DIBELS: LNF					
WERP	334	39.03	15.88	2.46	1.09
Comparison	1155	34.59	15.30	2.26	
WERP vs Comparison				.20	
DIBELS: WUF					
WERP	325	29.36	19.72	1.49	1.16
Comparison	998	27.40	21.34	1.28	
WERP vs Comparison				.21	
DIBELS: PSF					
WERP	355	25.38	15.57	1.63	1.22
Comparison	1219	22.24	16.63	1.34	
WERP vs Comparison				.29	
DIBELS: NWF					
WERP	355	20.41	15.50	1.32	1.20
Comparison	1217	16.80	15.33	1.10	
WERP vs Comparison				.22	
DIBELS: Total Reading					
WERP	334	22.97	8.60	2.67	1.30
Comparison		19.10	9.30	2.05	
WERP vs Comparison				.62	
TUSD: CCSA Reading					
WERP	311	1.59	.74	2.15	1.65
Comparison	1263	1.34	1.03	1.30	
WERP vs Comparison				.85	

Note. ISF = Initial Sounds Fluency, LNF = Letter Naming Fluency, WUF = Word Use Fluency, PSF = Phoneme Segmentation Fluency, NWF = Nonsense Word Fluency. N = Number, M = Mean Gain, PreSD = Pretest Standard Deviation, ES = Effect Size of Gain, Ratio = Ratio of WERP ES to Comparison ES. WERP students selected with 1100 minutes (6 months) or more usage of the Waterford Early Reading Program. The pretest-posttest effect size in the mean gain divided by the standard deviation (Walberg 2001).

Figure 3. Effect Size (ES) of WERP and Comparison Schools on All Outcome Measures



Findings. All of the effect size estimates indicate substantial effectiveness of the gains of the WERP over the gains of the Comparison group.

The differences between the effect sizes of the WERP students over the Comparison group students ranged from a high of .85 with the TUSD CCSA Reading test to a low of .20 for Letter Naming Fluency. The lowest effect size of .20 is within the range of substantial effectiveness of a treatment program over a comparison program (Cohen, 1977).

Table 7. Local Percentiles of WERP and Comparison Schools on DIBELS Total Reading Score

		Pretest		Posttest		
	N	M	SD	M	SD	Gain
WERP	334	46.88	16.63	56.00	20.12	9.12
Comparison	1210	46.45	18.84	47.99	21.60	1.54
WERP vs. Comparison						7.58

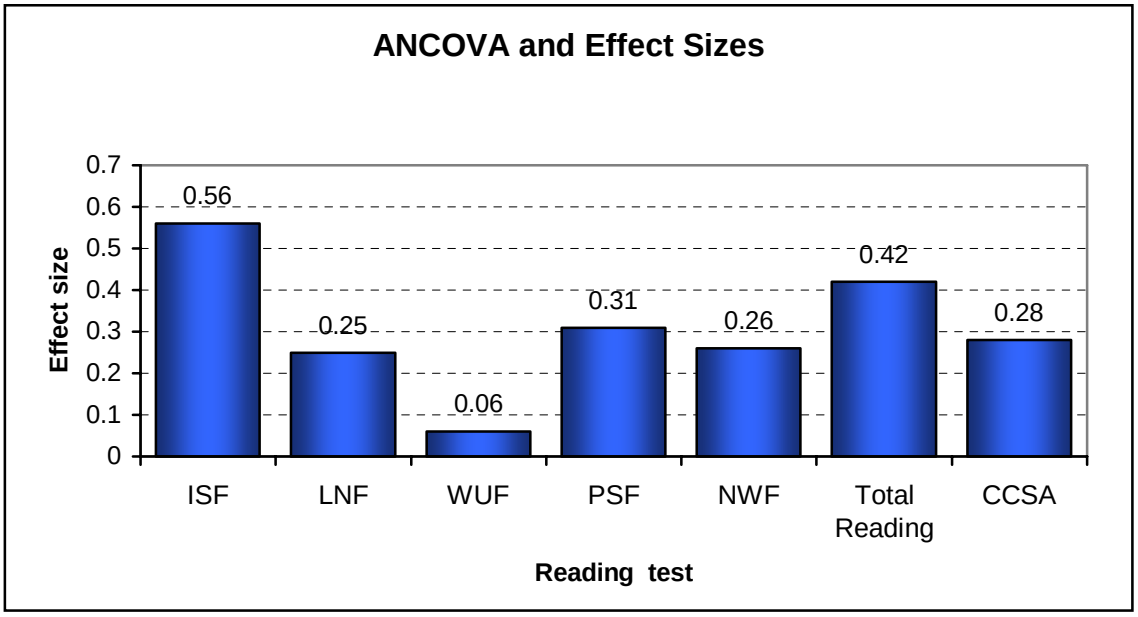
Findings. The University of Oregon calculates local percentiles for students taking the DIBELS. This provides a measure of how well students perform relative to other students in TUSD. Overall the WERP group's gain (9.12) exceeded the Comparison group's gain (1.54) by 7.58 percentile points, a highly relevant gain. In addition, WERP students' posttest reading (56.00) exceeded the TUSD average (50th percentile).

Table 8. ANCOVA and Effect Sizes of WERP and Comparison Groups on All Outcome Measures

		Covariate		AdjPosttest				
Measures	N	M	SD	M	SD	ES	F	p
DIBELS: ISF								
WERP	334	4.87	5.71	24.14	13.88	.56	81.57	.000
Comparison	1218	6.42	6.82	17.35	12.24			
DIBELS: LNF								
WERP	334	4.44	8.17	44.41	16.32	.25	16.33	.000
Comparison	1155	6.30	10.15	40.61	16.36			
DIBELS: WUF								
WERP	325	3.57	7.15	33.41	20.53	.06	.89	.345
Comparison	998	4.94	10.40	32.18	20.87			
DIBELS: PSF								
WERP	355	21.05	15.93	44.58	15.05	.31	26.22	.000
Comparison	1219	17.10	15.87	39.88	18.69			
DIBELS: NWF								
WERP	355	18.26	14.61	37.06	20.59	.26	1.16	.282
Comparison	1217	14.66	15.02	31.94	20.21			
DIBELS: Total Reading								
WERP	334	10.62	6.98	33.22	12.42	.42	46.16	.000
Comparison	1211	10.22	8.29	29.43	13.33			
TUSD: CCSA Reading								
WERP	311	1.09	0.49	2.67	0.64	.28	20.04	.000
Comparison	1263	1.07	0.59	2.41	1.02			

Note. ISF = Initial Sounds Fluency, LNF = Letter Naming Fluency, WUF = Word Use Fluency, PSF = Phoneme Segmentation Fluency, NWF = Nonsense Word Fluency. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program. The effect size is the adjusted mean posttest difference divided by the square root of the ANCOVA mean squared residual.

Figure 4. ANCOVA and Effect Sizes of WERP and Comparison Groups on All Outcome Measures



Findings. After the WERP and Comparison students' posttest means were adjusted for differences on their pretest status, there were still highly significant differences between the adjusted posttest means ($p < .001$) for Initial Sounds Fluency, Letter Naming Fluency, Phoneme Segmentation, Total Reading Score, and the CCSA Reading Assessment.

The effect sizes show the effectiveness of the WERP over the Comparison group. Six of the seven outcome measures had significant and substantial effect sizes ranging from .20 to .56.

B. Intervention Effects on Subgroups

This section examines the outcome measures when the student population is disaggregated by program, by gender, by ethnicity, by primary home language and by ELL status.

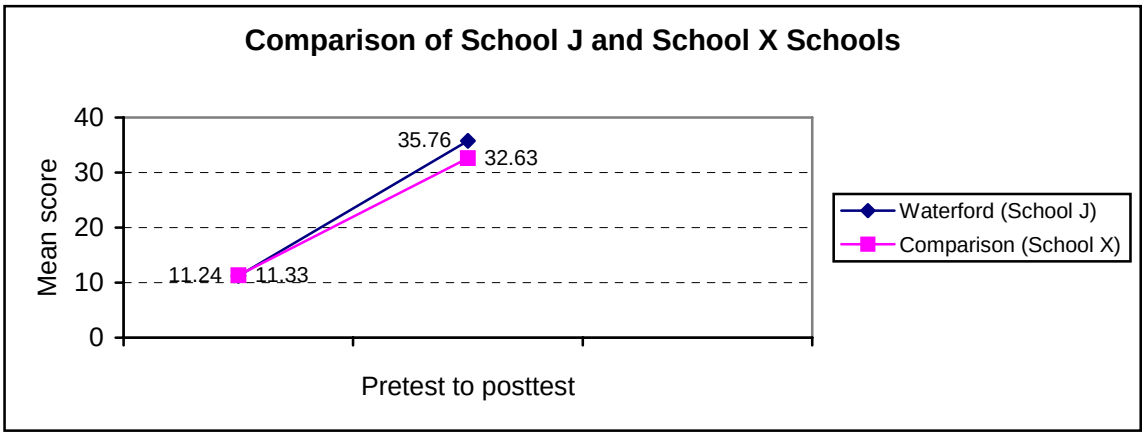
The following pairs of schools were examined because these WERP schools had a high level of implementation of the program, were also Reading First schools, and were closely matched on the DIBELS pretest Total Reading Score with a Comparison school. Thus School J (WERP) was matched with School X (Comparison), School K (WERP) was matched with School V (Comparison), and School H (WERP) was matched with School M (Comparison).

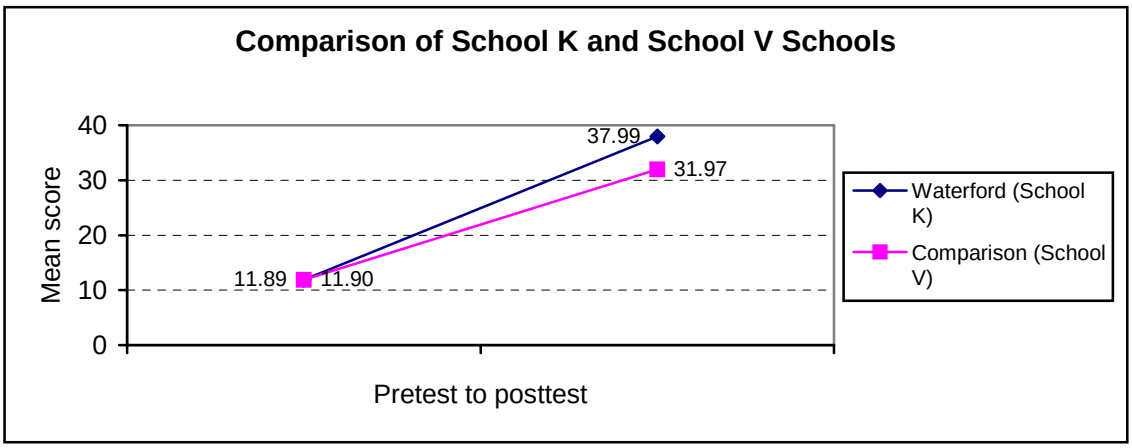
Table 9. WERP + Reading First and Comparison Schools Total Reading Pretest-Posttest Means and Gains

Schools	Pretest	Posttest	Gain
WERP+ Reading First: School J	11.24	35.76	24.52
Comparison: School X	11.33	32.63	21.31
WERP vs Comparison	-.09	3.13	3.21
WERP+ Reading First: School K	11.89	37.99	26.10
Comparison: School V	11.90	31.97	20.07
WERP vs Comparison	-.01	6.02	6.03
WERP+ Reading First: School H	12.86	32.98	20.12
Comparison: School M	13.06	29.56	16.50
WERP vs Comparison	-.20	3.42	3.62

Note. WERP and Comparison schools were matched by DIBELS Total Reading Score Pretest Mean.

Figure 5. WERP + Reading First and Comparison Schools Matched on DIBELS Reading Pretest Mean





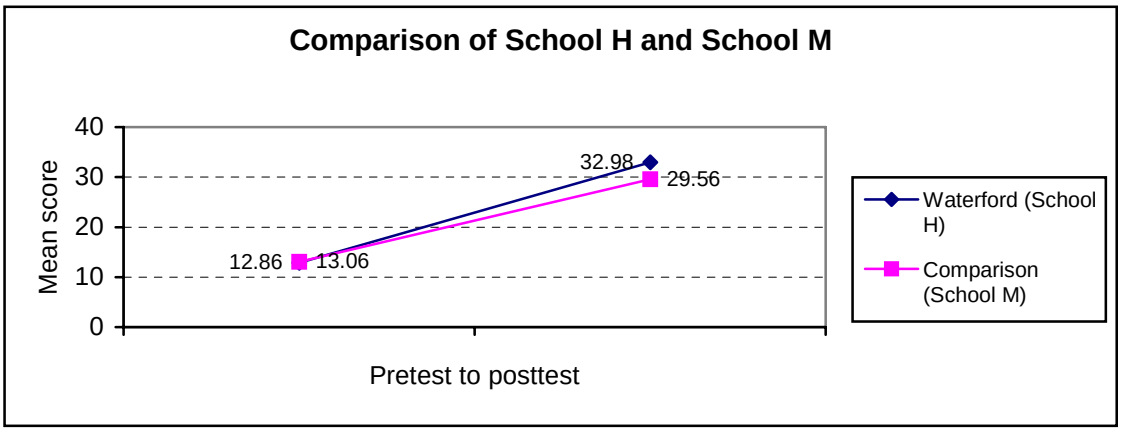
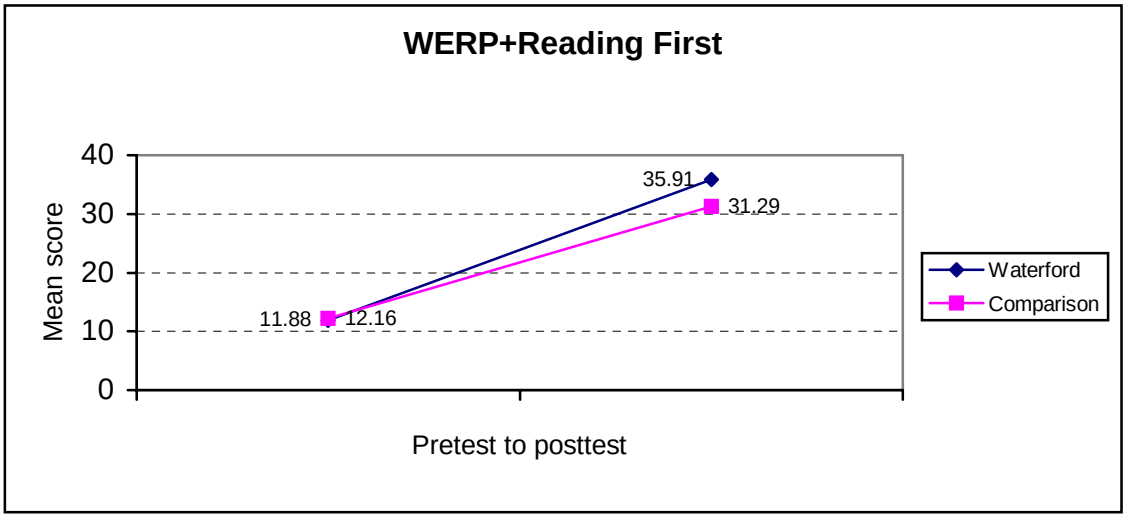


Table 10. WERP + Reading First (Schools J, K, H) and Comparison School (Schools X, V, M) Gains on All Outcome Measures

		Pretest		Posttest				
Measures	N	M	SD	M	SD	Gain	t	p
DIBELS: ISF								
WERP	214	5.10	5.40	25.24	13.81	20.14	21.62	.000
Comparison	260	6.50	7.24	15.00	11.42	8.50	11.19	.000
WERP vs. Comparison						11.64***		
DIBELS: LNF								
WERP	214	4.88	8.99	45.95	15.81	41.07	38.81	.000
Comparison	239	5.04	9.67	41.79	16.55	36.75	36.34	.000
WERP vs. Comparison						4.32**		
DIBELS: WUF								
WERP	211	3.41	7.55	32.72	20.63	29.31	20.90	.000
Comparison	84	9.46	13.16	39.51	20.71	30.05	13.94	.000
WERP vs. Comparison						-0.74		
DIBELS: PSF								
WERP	230	23.96	16.16	50.53	11.11	26.57	25.81	.000
Comparison	255	19.37	16.32	44.93	18.58	25.56	24.12	.000
WERP vs. Comparison						1.01		
DIBELS: NWF								
WERP	230	20.52	14.02	43.94	20.12	23.42	22.33	.000
Comparison	253	17.17	13.96	34.91	19.01	17.74	18.85	.000
WERP vs. Comparison						5.68***		
DIBELS: Total Reading								
WERP	214	11.88	7.06	35.91	11.46	24.03	44.11	.000
Comparison	258	12.16	9.31	31.29	13.47	19.13	30.99	.000
WERP vs. Comparison						4.90***		
TUSD: CCSA Reading								
WERP	189	1.06	.50	2.72	.58	1.66	31.75	.000
Comparison	265	1.08	.60	2.36	1.07	1.28	19.84	.000
WERP vs. Comparison						0.38***		

Note. ISF = Initial Sounds Fluency, LNF = Letter Naming Fluency, WUF = Word Use Fluency, PSF = Phoneme Segmentation Fluency, NWF = Nonsense Word Fluency. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program. *p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Figure 6. WERP + Reading First and Comparison Groups Gains on Total Reading Score



Findings. Three schools with the WERP and the Reading First program outperformed Comparison schools with which they were matched, both all together as a group and school by school. These schools were School J vs. School X; School K vs. School V; and School H vs. School M.

Analysis of the DIBELS subscales indicated that students receiving Reading First and the WERP performed significantly better than students in the Comparison schools in the DIBELS Initial Word Fluency, Letter Naming Fluency, Nonsense Word Fluency, and Total Reading Score and in the CCSA Reading Assessment. In the DIBELS Total Reading Score the difference was 4.90 points, statistically significant at the .001 level. See Table 10 and Figure 6.

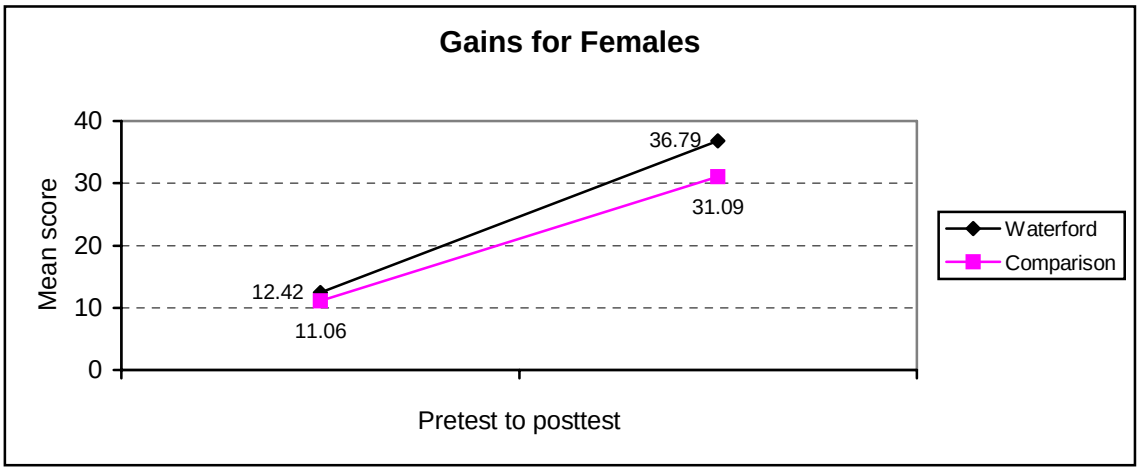
Table 11. Males and Females in WERP and Comparison Schools on DIBELS Total Reading Score

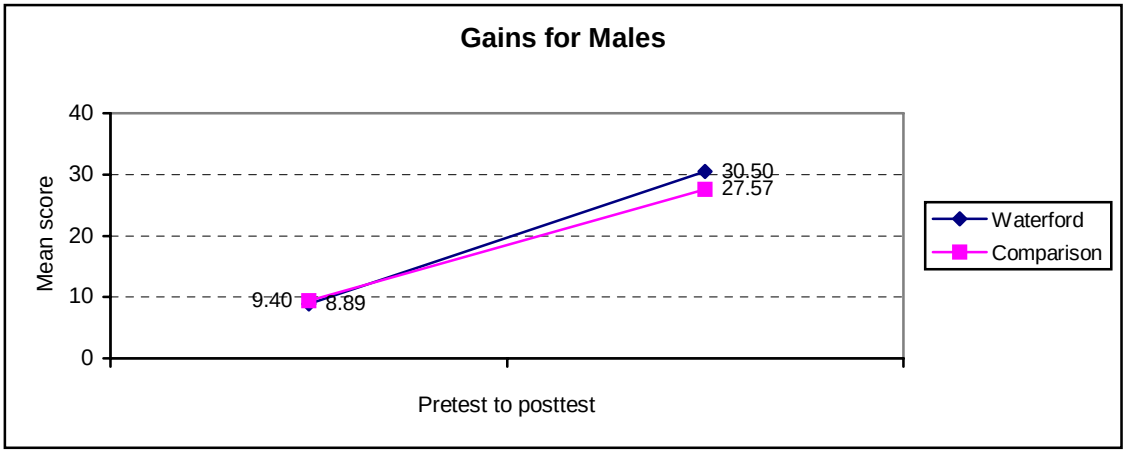
		Pretest		Posttest				
Group	N	M	SD	M	SD	Gain	t	p
WERP								
Female	164	12.42	7.17	36.79	11.67	24.37	38.07	.000
Male	170	8.89	6.33	30.50	12.38	21.61	32.09	.000
Female vs Male						2.76**		
Comparison								
Female	603	11.06	8.35	31.09	13.12	20.03	54.59	.000
Male	608	9.40	8.17	27.57	13.32	18.17	47.18	.000
Female vs Male						1.86***		

		Pretest		Posttest				
Group	N	M	SD	M	SD	Gain	t	p
Male								
WERP	170	8.89	6.33	30.50	12.38	21.61	32.09	.000
Comparison	608	9.40	8.17	27.57	13.32	18.17	47.18	.000
WERP vs. Comparison						3.44***		
Female								
WERP	164	12.42	7.17	36.79	11.67	24.37	38.07	.000
Comparison	603	11.06	8.35	31.09	13.12	20.03	54.59	.000
WERP vs. Comparison						4.34***		

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.
*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Figure 7. Males and Females in WERP and Comparison Schools on DIBELS Total Reading Score





Findings. Females significantly outperformed males on the DIBELS Total Reading Score in both WERP and Comparison schools. See Table 11.

Males in the WERP schools outperformed males in Comparison schools on the DIBELS Total Reading Score, and WERP females outperformed Comparison females. See Table 11 and Figure 7.

Table 12. Ethnic Groups in WERP and Comparison Schools on DIBELS Total Reading Score

		Pretest		Posttest				
Group	N	M	SD	M	SD	Gain	t	p
WERP								
White	13	14.48	7.61	40.71	9.67	26.23	10.68	.000
African American	16	6.16	4.97	26.35	9.75	20.19	10.15	.000
Hispanic	279	10.65	6.91	33.84	12.48	23.19	45.45	.000
Native American	18	9.63	5.86	29.27	9.86	19.63	10.69	.000
Asian	8	14.55	9.53	37.45	16.60	22.90	5.32	.000
Comparison								
White	219	12.43	9.17	32.25	13.91	19.82	31.97	.000
African American	81	10.24	9.09	28.07	14.95	17.83	15.67	.000
Hispanic	850	9.59	7.92	28.65	12.95	19.06	60.33	.000
Native American	46	10.69	7.94	29.70	12.20	19.01	13.93	.000
Asian	15	12.31	7.56	30.47	15.95	18.16	6.18	.000

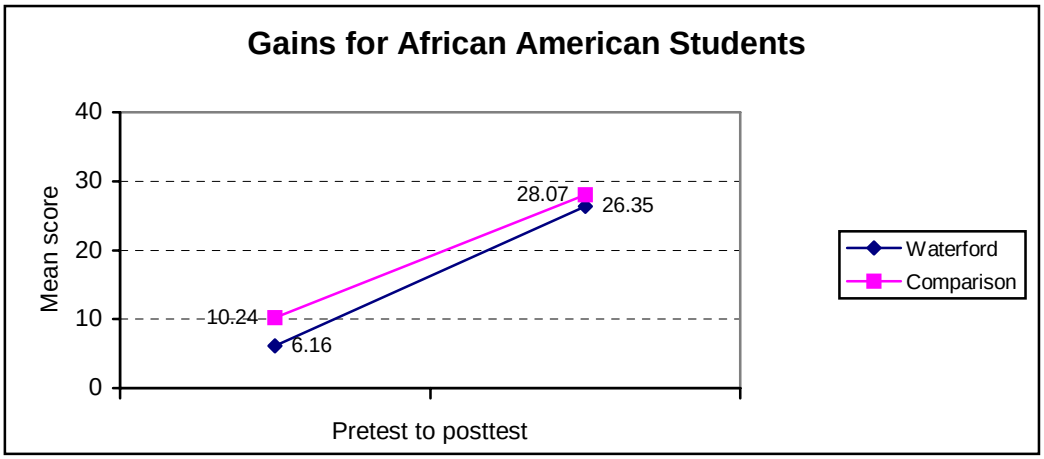
Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program. * $p < .05$, * < .01, *** $p < .001$ from independent t tests to compare gains.

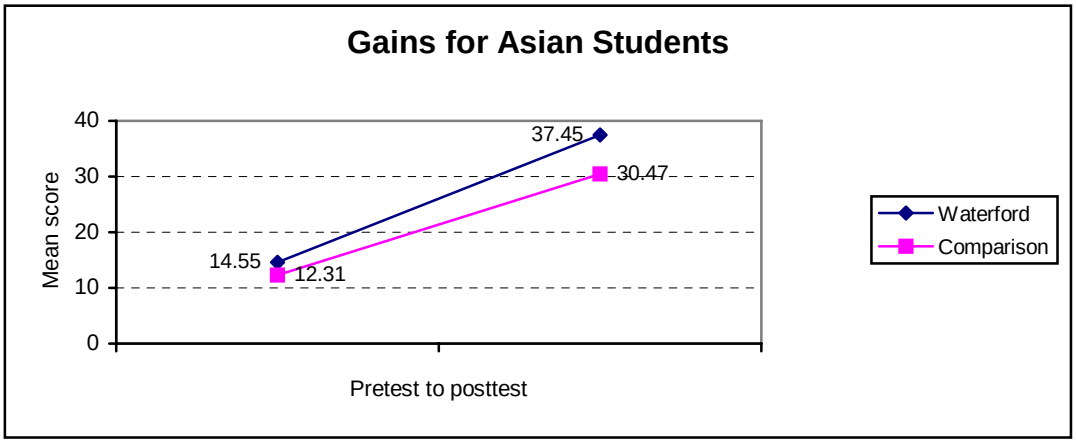
Table 12. Ethnic Groups in WERP and Comparison Schools on DIBELS Total Reading Score (continued)

		Pretest		Posttest				
Group	N	M	SD	M	SD	Gain	t	p
White								
WERP	13	14.48	7.61	40.71	9.67	26.23	10.68	.000
Comparison	219	12.43	9.17	32.25	13.91	19.82	31.97	.000
WERP vs. Comparison						6.41*		
African American								
WERP	16	6.16	4.97	26.35	9.75	20.19	10.15	.000
Comparison	81	10.24	9.09	28.07	14.95	17.83	15.67	.000
WERP vs. Comparison						2.36		
Hispanic								
WERP	279	10.65	6.91	33.84	12.48	23.19	45.45	.000
Comparison	850	9.59	7.92	28.65	12.95	19.06	60.33	.000
WERP vs. Comparison						4.13***		
Native American								
WERP	18	9.63	5.86	29.27	9.86	19.63	10.69	.000
Comparison	46	10.69	7.94	29.70	12.20	19.01	13.93	.000
WERP vs. Comparison						0.62		
Asian								
WERP	8	14.55	9.53	37.45	16.60	22.90	5.32	.000
Comparison	15	12.31	7.56	30.47	15.95	18.16	6.18	.000
WERP vs. Comparison						4.74		

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.
*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Figure 8. Ethnic Groups in WERP and Comparison Schools on DIBELS Total Reading Score





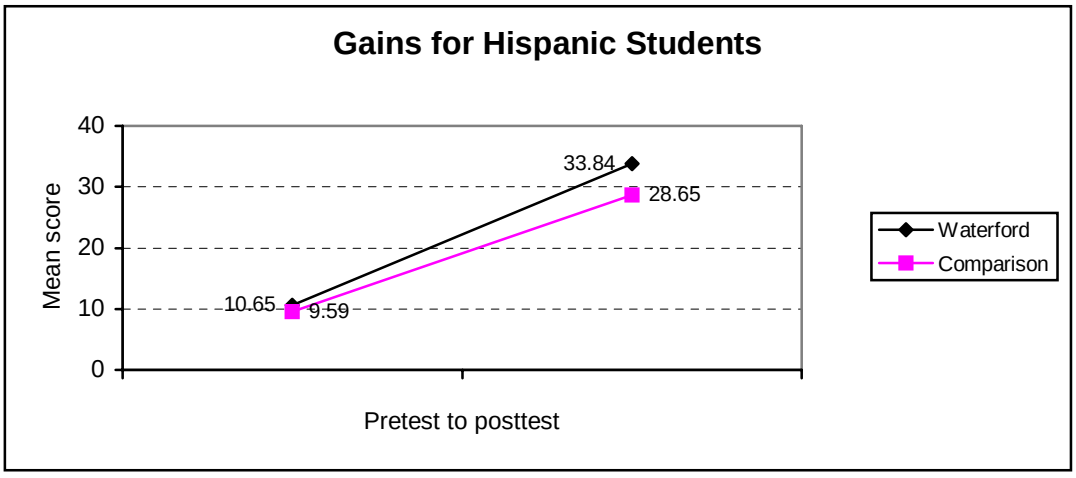
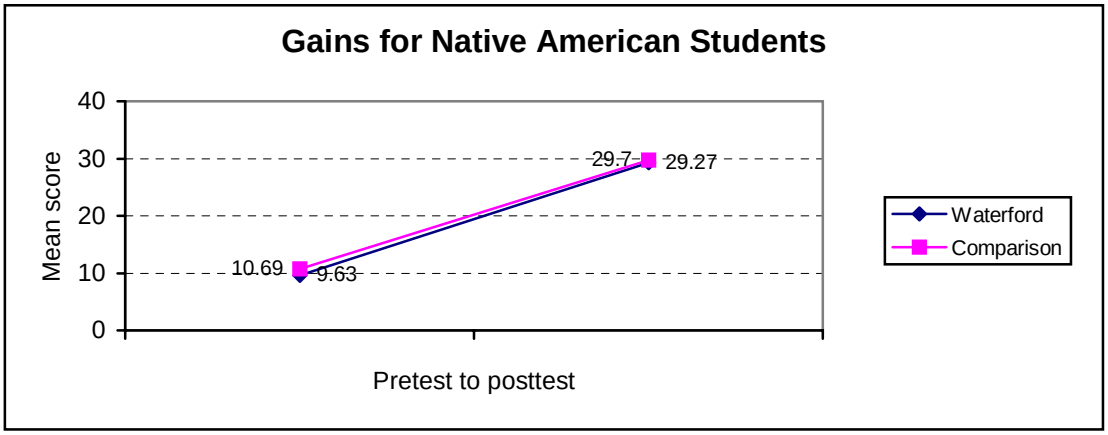


Figure 8. Ethnic Groups in WERP and Comparison Schools on DIBELS Total Reading Score (continued)



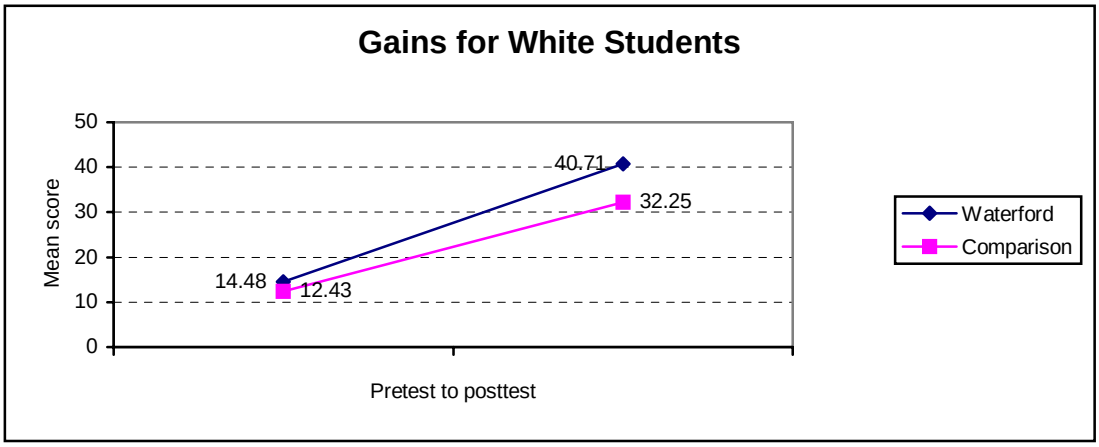
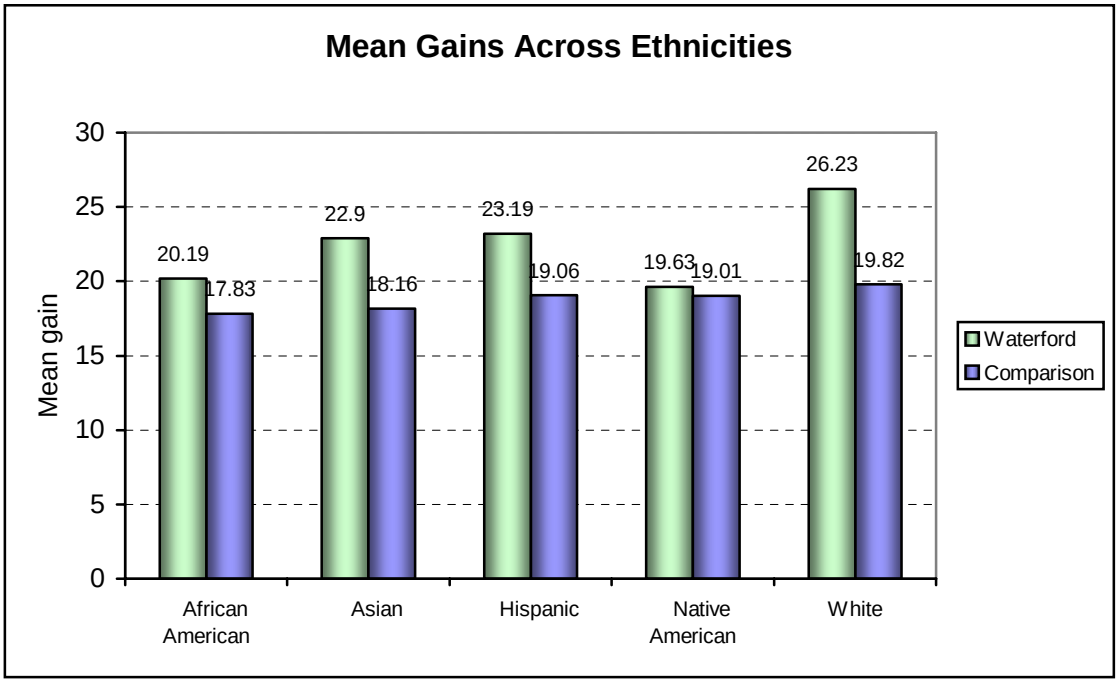


Figure 9. Mean Gains by Ethnicity in WERP and Comparison Schools on DIBELS Total Reading Score



Findings. All ethnic groups, whether in the WERP schools or in the Comparison schools, made important gains from pretest to posttest on the DIBELS Total Reading Score. See Table 12.

A comparison of gains shows that all ethnic groups receiving the WERP made greater gains than their counterparts in the Comparison group

A surprising finding was that Hispanic (23.19), Asian (22.90), and African American (20.19) students in the WERP schools made greater gains pretest to posttest on the Total Reading Score than did the White (19.82) students not receiving WERP.

The greatest gain pretest to posttest (26.23 points) was made by the White students in the WERP schools. These students also showed the greatest gain relative to their counterparts in the Comparison schools for a statistically significant difference of 6.41 points.

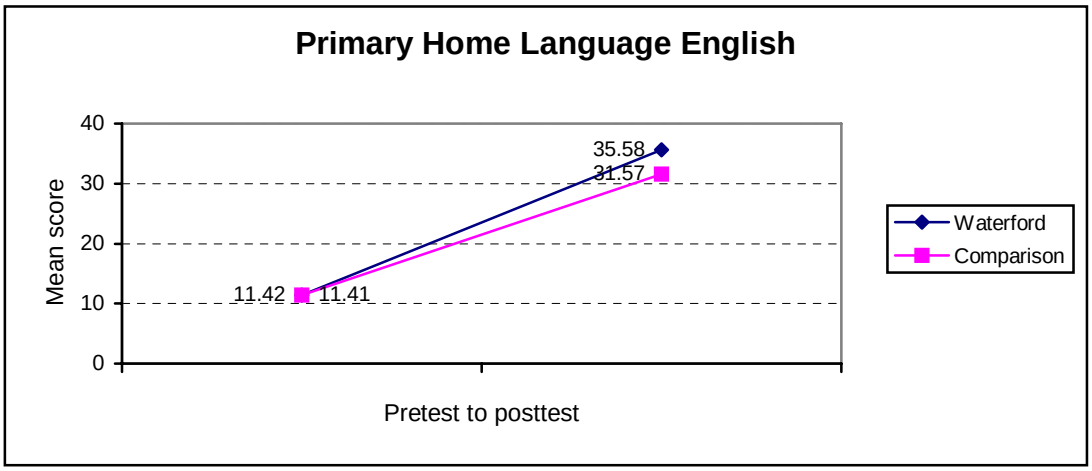
Table 13. Primary Home Languages in WERP and Comparison Schools on DIBELS Total Reading Score

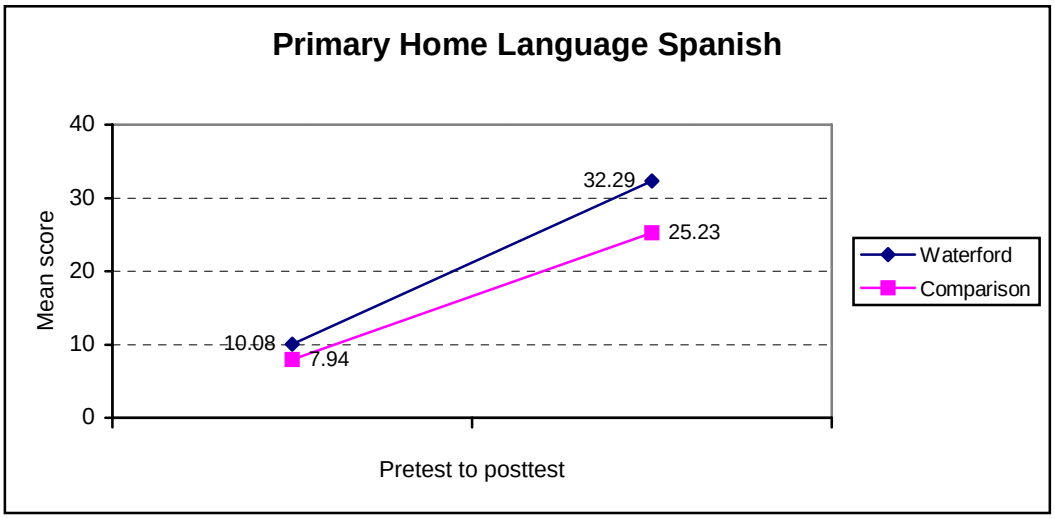
		Pretest		Posttest				
Group	N	M	SD	M	SD	Gain	t	p
WERP								
English	160	11.41	6.69	35.58	11.60	24.18	36.68	.000
Spanish	163	10.08	7.13	32.29	12.84	22.21	32.52	.000
Other	11	7.11	7.48	23.82	10.62	16.71	7.79	.000
Comparison								
English	823	11.42	8.77	31.57	13.17	20.15	62.79	.000
Spanish	362	7.94	6.53	25.23	12.17	17.29	36.49	.000
Other	26	4.15	4.88	15.30	12.31	11.15	6.43	.000

		Pretest		Posttest				
Group	N	M	SD	M	SD	Gain	t	p
English								
WERP	160	11.41	6.69	35.58	11.60	24.18	36.68	.000
Comparison	823	11.42	8.77	31.57	13.17	20.15	62.79	.000
WERP vs. Comparison						4.03***		
Spanish								
WERP	163	10.08	7.13	32.29	12.84	22.21	32.52	.000
Comparison	362	7.94	6.53	25.23	12.17	17.29	36.49	.000
WERP vs. Comparison						4.92***		
Other								
WERP	11	7.11	7.48	23.82	10.62	16.71	7.79	.000
Comparison	26	4.15	4.88	15.30	12.31	11.15	6.43	.000
WERP vs. Comparison						5.56		

Note. Other languages are Af-Mayma, Amharic, Arabic, Cantonese, Persian, Filipino, French, Laotian, Marshallese, Portuguese, Russian, Somali, and Vietnamese. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.
*p < .05, ** p < . 01, *** p < .001 from independent t test comparing gains.

Figure 10. Primary Home Languages in WERP and Comparison Schools on DIBELS Total Reading Score





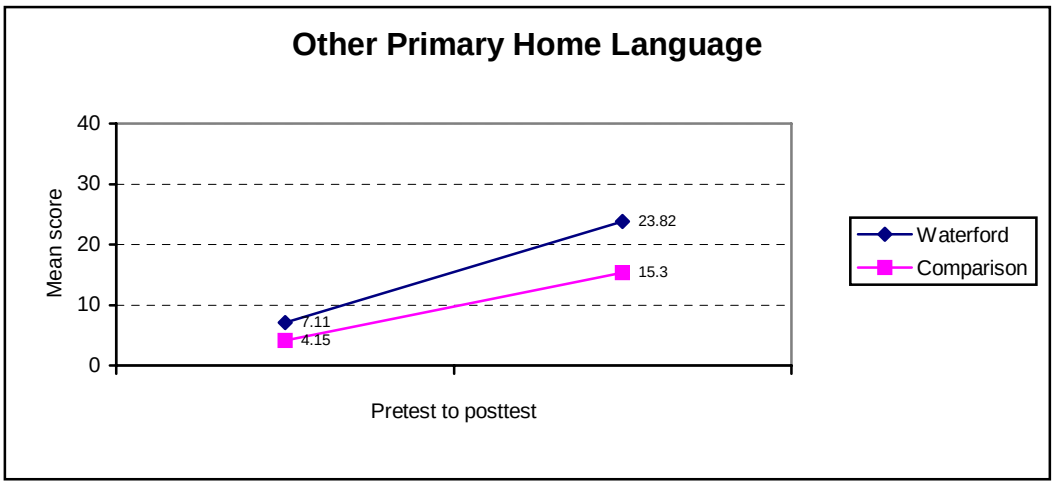
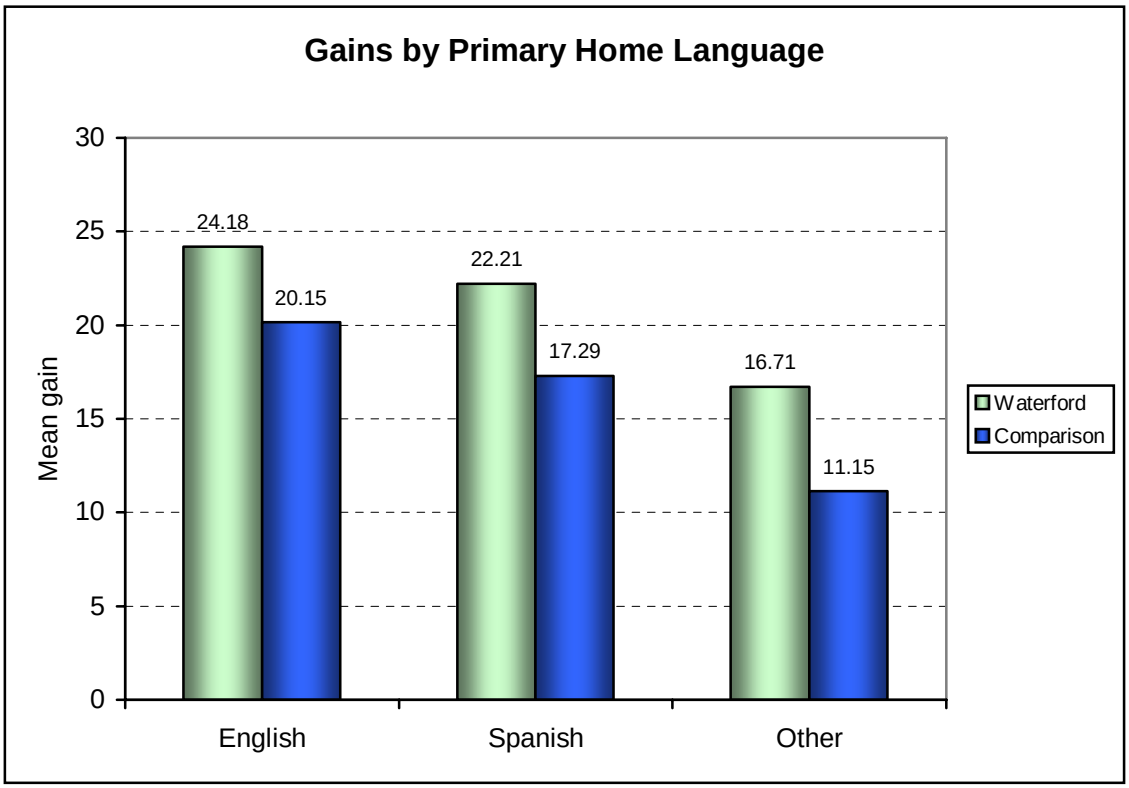


Figure 11. Gains by Primary Home Languages in WERP and Comparison Schools



Findings. Whether their primary home language was English, Spanish or another language, WERP students outperformed their counterparts in the Comparison group on the DIBELS Total Reading Score. This difference was statistically significant for the English and the Spanish home language groups.

It is interesting to note that WERP students with Spanish (22.21) as their primary home language significantly outperformed in gains the Comparison group students who spoke English as their primary home language (20.15).

The greatest gain in pretest to posttest scores was by the English-speaking WERP students, who gained 24.18 points.

The WERP group with the greatest gain (5.56 points) relative to the Comparison group was that of students who spoke a primary home language other than English or Spanish. This diverse group includes refugee children who often have a history of upheavals, trauma and no prior school experience.

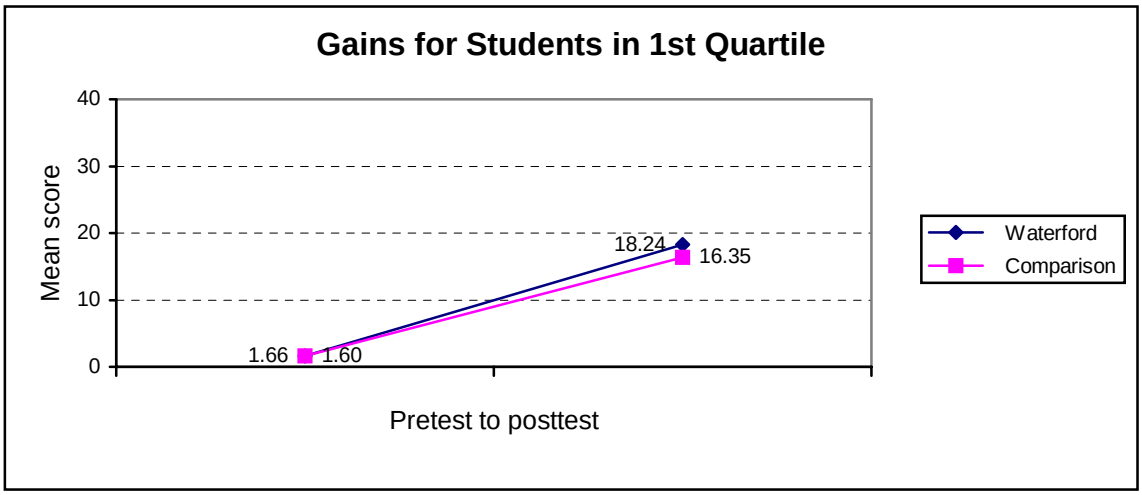
Table 14. WERP and Comparison Schools Four Achievement Quartiles of the DIBELS Total Reading Score

Group	N	Means Pretest	Posttest	Gain	t	p
1st Quartile						
WERP	60	1.60	18.24	16.64	13.64	.000
Comparison	304	1.66	16.35	14.69	27.62	.000
WERP vs Comparison				1.95		
2nd Quartile						
WERP	75	6.00	29.70	23.70	27.42	.000
Comparison	316	6.12	26.12	20.00	41.49	.000
WERP vs Comparison				3.70**		
3rd Quartile						
WERP	113	11.63	35.77	24.14	34.23	.000
Comparison	278	11.36	32.59	21.23	40.91	.000
WERP vs Comparison				2.91**		
4th Quartile						
WERP	86	19.62	44.82	25.20	28.08	.000
Comparison	313	21.68	42.26	20.58	39.40	.000
				4.62***		

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program. The four achievement quartiles (25%) based on all students' rankings on the DIBELS Total Pretest Score.

*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Figure 12. WERP and Comparison Schools by Achievement Quartile



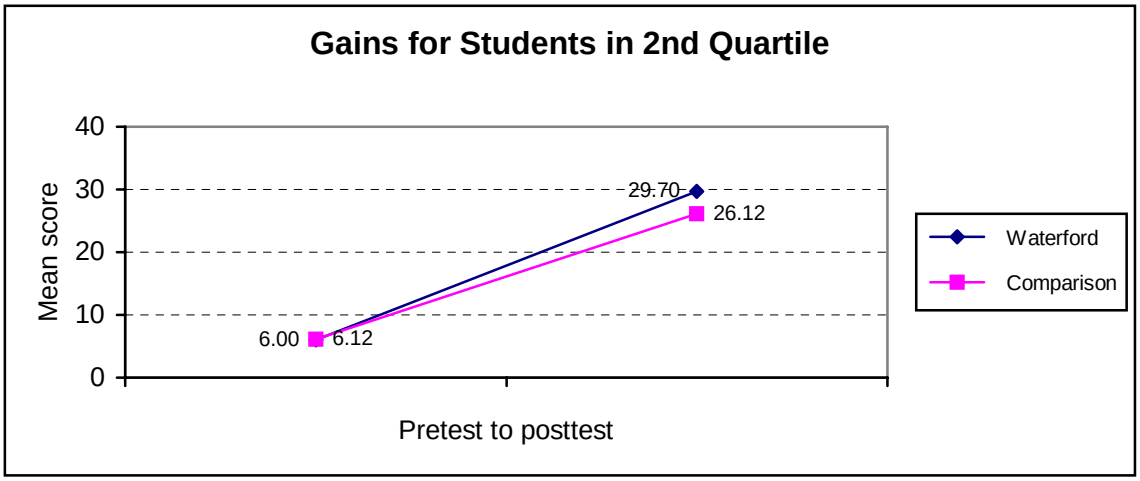
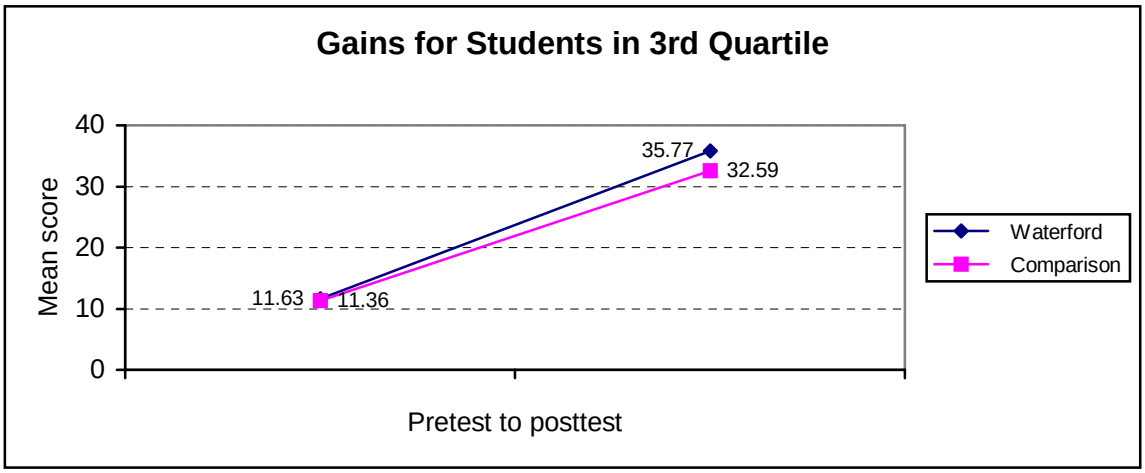
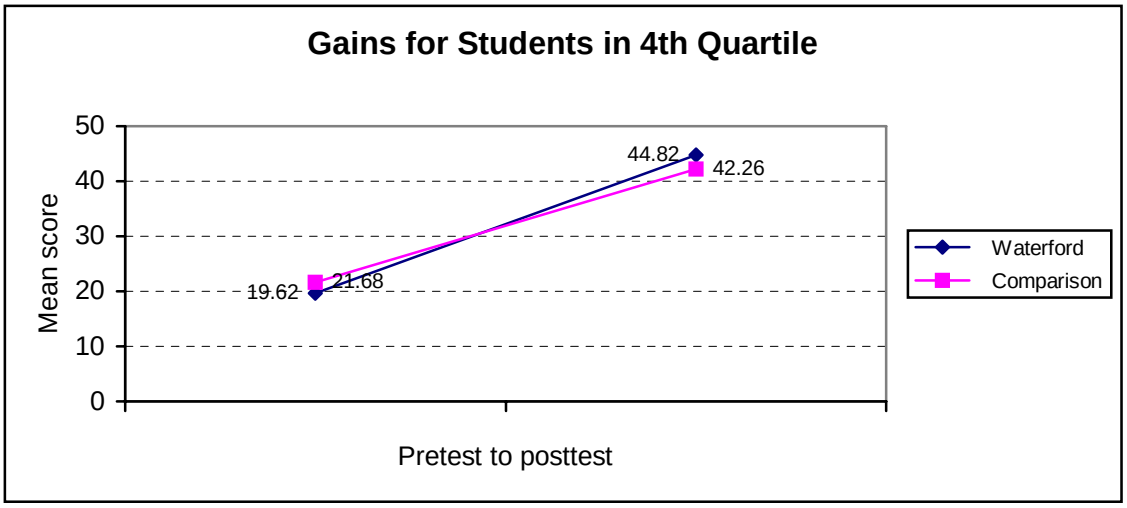


Figure 12. WERP and Comparison Schools by Achievement Quartile (continued)





Findings. Kindergartners were grouped based on the DIBELS Total Reading Score pretest into four quartiles. Students in each quartile at the WERP schools scored higher on the posttest than students in the same quartile at the Comparison schools.

When the pretest–posttest gains were compared, WERP kindergartners in the 2nd, 3rd, and 4th quartile made significantly greater gains than their counterparts in the 2nd, 3rd and 4th quartiles. The greatest gain (25.20 points) was made by the 4th quartile students in the WERP schools. This result suggests that the WERP provides content to allow children who come with more preliteracy experience to make more rapid gains.

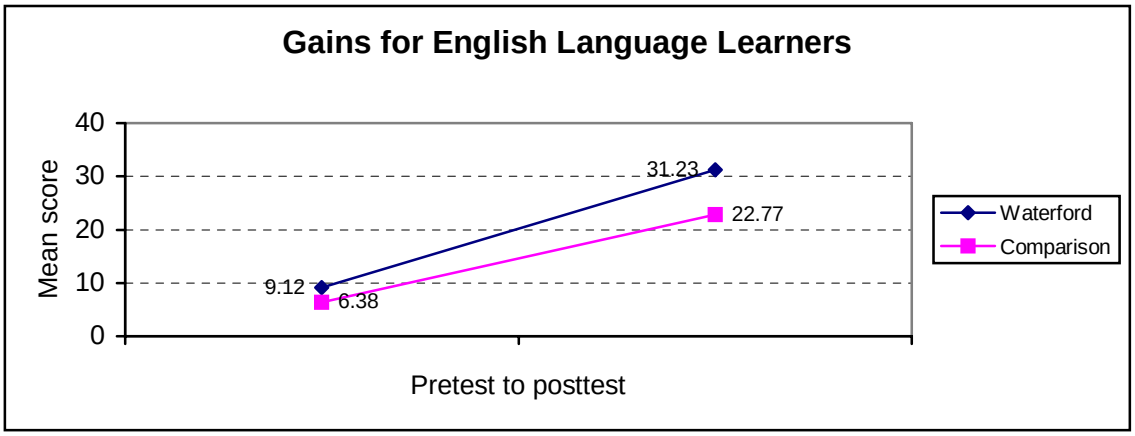
Table 15. ELL Students and Non-ELL (English Speakers) in WERP and Comparison Groups on DIBELS Total Reading Score

		Pretest		Posttest				
Measures	N	M	SD	M	SD	Gain	t	p
ELL students								
WERP	164	9.12	6.65	31.23	12.95	22.11	32.37	.000
Comparison	329	6.38	5.66	22.77	12.04	16.39	32.43	.000
WERP vs. Comparison						5.72***		
Non-ELL students								
WERP	170	12.07	7.00	35.86	11.49	23.79	36.94	.000
Comparison	882	11.66	8.66	31.77	12.97	20.11	65.26	.000
WERP vs. Comparison						3.68***		

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.

*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Figure 13. ELL Students and Non-ELL (English Speakers) in WERP and Comparison Groups



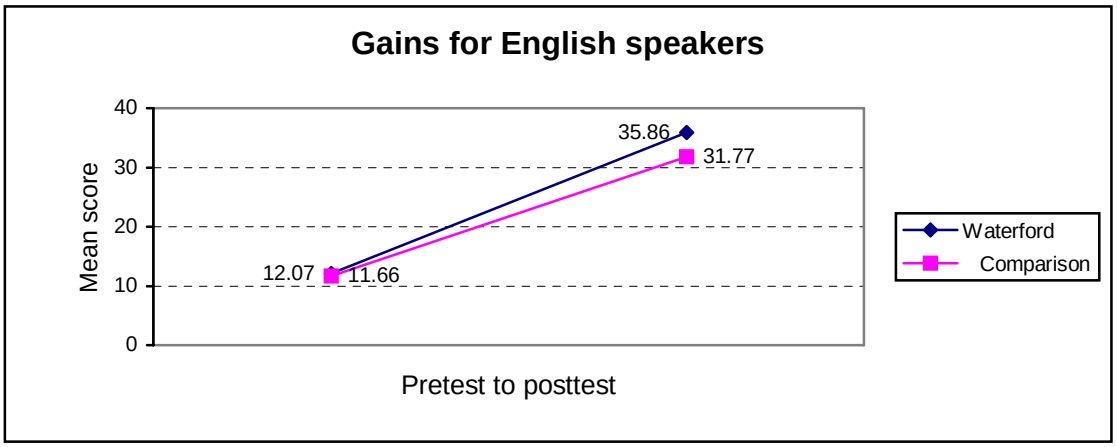
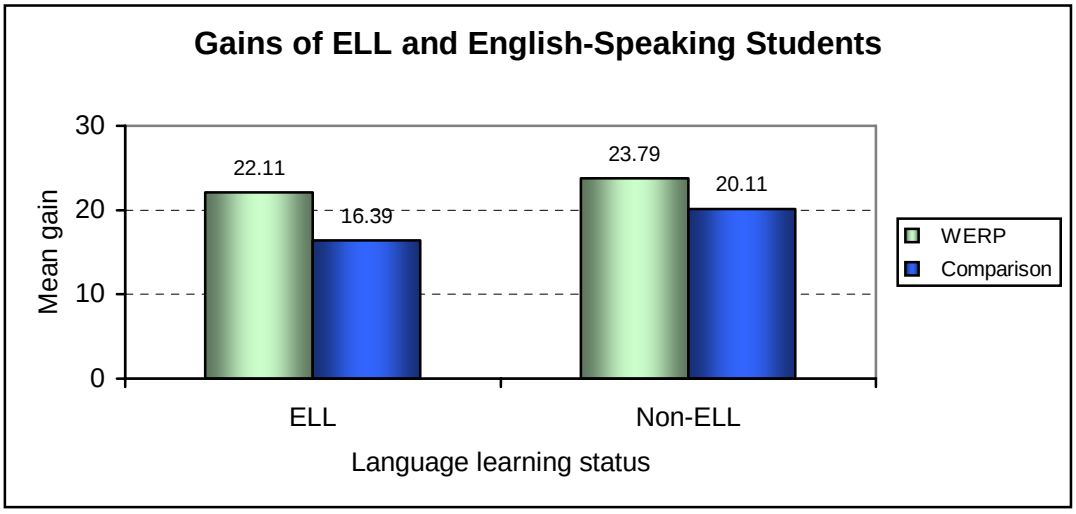


Figure 14. Gains by Language Learning Status in WERP and Comparison Groups



Findings. Both ELL students and non-ELL students (English speakers) made significantly greater gains pretest to posttest than their counterparts in the Comparison group.

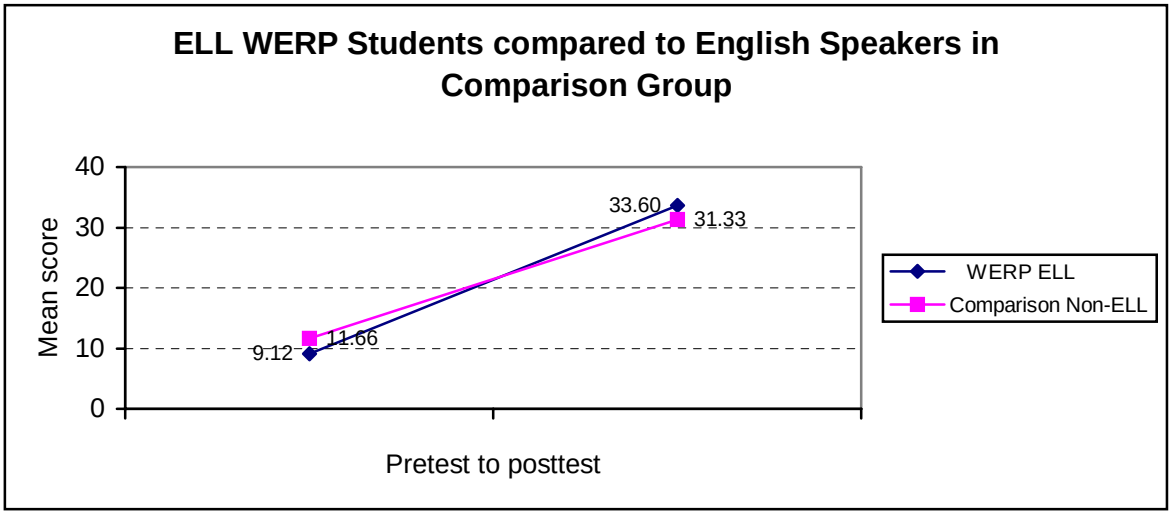
Table 16. ANCOVA of WERP ELL and Comparison Non-ELL (English Speakers) on DIBELS Total Reading Score

		Pretest		AdjPosttest				
Group	N	M	SD	M	SD	Gain	F	p
DIBELS: Total Reading								
WERP ELL	164	9.12	6.65	33.60	12.95	24.48	8.62	.003
Comp. English Speak.	882	11.66	8.66	31.33	12.97	19.67		
WERP vs Comparison						4.81***		

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.

*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Figure 15. ANCOVA of WERP ELL and Comparison Non-ELL (English Proficient Speakers)



Findings. WERP ELL students' gain (24.48) was compared to the Comparison non-ELL (English speakers) students' gain (19.67) using ANCOVA in order to adjust for initial differences. The WERP ELL students statistically outperformed the English-speaking students in the Comparison group in gains. See Table 16.

C. WERP Usage Effects

Table 17. Correlations of Usage of WERP, Reading Achievement and Reading Gains

Measures	Usage			
	Total	Level 1	Level 2	PA
Initial Sounds Fluency	.24**	.15**	.23**	.18**
Letter Naming Fluency	.16**	.03	.40**	.12**
Word Use Fluency	.07*	.06	.02	.01
Phoneme Segmentation Fluency	.19**	.11**	.24**	.19**
Nonsense Word Fluency	.33**	.17**	.50**	.26**
Total DIBELS	.27**	.15**	.36**	.19**
CCSA Reading	.12**	.15**	.02	.10*
Gains Initial Sounds Fluency	.20**	-.02	.18*	.05
Gains Letter Naming Fluency	.25**	-.01	.23**	.18**
Gains Word Use Fluency	.04	.03	.09	-.13
Gains Phoneme Segmentation Fluency	.08	.13*	-.08	.12*
Gains Nonsense Word Fluency	.24**	.03	.32**	.15**
Gains Total DIBELS	.11*	-.01	.06	.00
Gains CCSA Reading	.22**	.07	.11	.17**

Note. PA = Phonological Awareness. Usage Level 1 = Reading Level 1 total minutes in the course; Usage Level 2 = Reading Level 2 total minutes in the course; Total = the total usage minutes of Level 1, Level 2, and Phonological Awareness.

Findings. The total usage in minutes of the WERP software was significantly and positively correlated with posttest reading measures (Initial Sound Fluency, Letter Naming Fluency, Word Use Fluency, Phoneme Segmentation Fluency, Nonsense Word Fluency, the Total DIBELS Reading, and CCSA Reading). The median of the significant correlations was .20. This meant that students who spent more time using the WERP software tended to have higher posttest reading achievement.

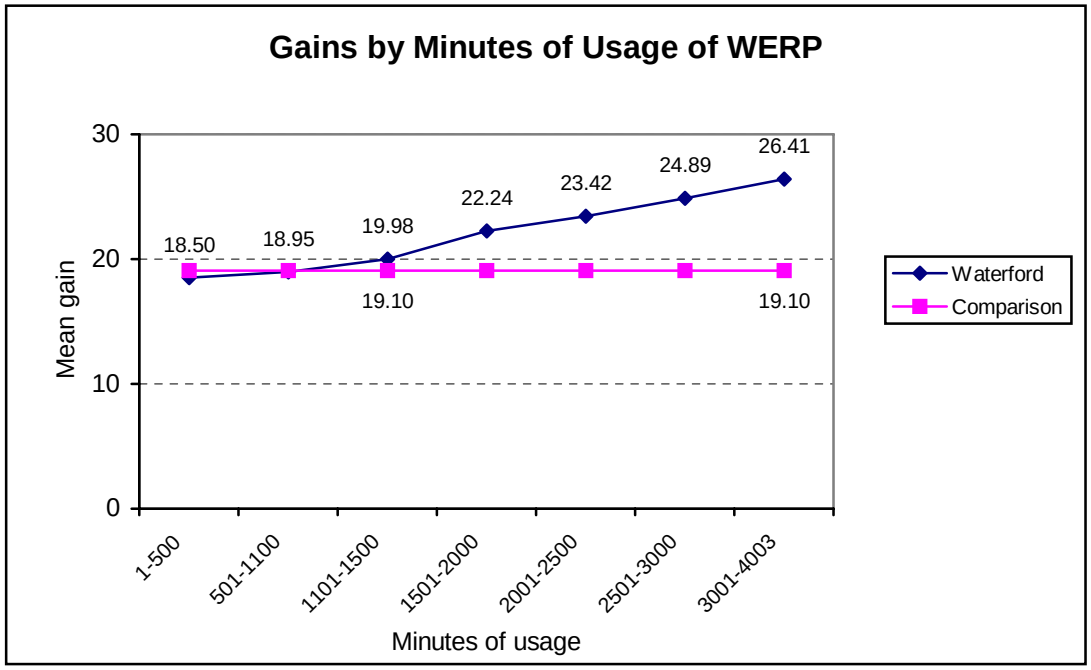
There was also a correlation between the total usage in minutes of the WERP software and the gains of students in Initial Sounds Fluency, Letter Naming Fluency, Nonsense Word Fluency, DIBELS Reading, and CCSA Reading. The median of these significant correlations was .20. These correlations suggest that those who spent more time using the WERP materials tended have greater gains in these reading areas.

Table 18. WERP and Comparison Group Gains on DIBELS Total Reading Score

		Pretest		Posttest		
Usage Groups	N	M	SD	M	SD	Gains
WERP 1-500 mins	153	9.72	7.14	28.22	12.86	18.50
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						-.60
WERP 501 –1100 mins	139	8.94	8.28	27.89	11.23	18.95
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						-.15
WERP 1101 – 1500 mins	76	8.67	5.99	28.65	11.46	19.98
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						.88
WERP 1501 – 2000 mins	84	10.05	6.40	32.29	12.19	22.24
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						3.14**
WERP 2001 – 2500 mins	56	11.76	7.94	35.18	12.95	23.42
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						4.32**
WERP 2501 – 3000 mins	94	11.35	6.96	36.24	11.58	24.89
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						5.79***
WERP 3001 – 4003 mins	24	13.27	8.16	39.68	12.92	26.41
Comparison	1211	10.22	8.29	29.32	13.33	19.10
WERP vs Comparison						7.31***

Note. *p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains. All pretest – posttest comparisons with paired – samples t test were highly significant p < .001. Spearman's correlation between usage and reading gains was r_s = .93, p < .001 indicating the more usage the greater the reading gains.

Figure 16. Gains by Level of WERP Usage Compared to Comparison Group



Findings. The reading achievement of the WERP students was grouped by seven increasing levels of usage. WERP student gains at each level were compared with the comparison group gains (19.10). WERP gains increased relative to the Comparison group beginning with 1 – 500 minutes of usage with the significant gains starting with 1500 minutes of usage. WERP student gains over the Comparison group increased by 3.14, 4.32, 5.79, and 7.31 as the usage of WERP increased. In addition, this suggests that the greater the usage of WERP content the more significant gains are made.

IV. SUMMARY AND DISCUSSION

A. Summary and Discussion

A summary of the major findings of this evaluation follows. A full presentation of the findings has been presented in tables and charts of this report. The findings are unusual in their consistency across ethnic, language, gender, and achievement groups in the manner with which they present gains favoring the WERP.

The present study contributes to the research on the effectiveness of the WERP by disaggregating the results by gender, language and ethnicity subgroups and by considering the effects of varying dosages (minutes of usage) of the WERP.

B. Significant Findings

- The WERP kindergartners consistently outperformed the Comparison group kindergartners on all reading outcome measures. Comparison school kindergartners did make substantial and in some cases outstanding gains from pretest to posttest. However, when WERP kindergartners were compared with Comparison kindergartners, the WERP gains were substantially and significantly greater.

- Effect sizes of reading pretest to posttest gains favored the WERP kindergartners, as well as effect sizes comparing the DIBELS reading posttest achievement of the WERP kindergartners with the Comparison kindergartners.

- WERP reading gains were greater for males in the WERP program than were the reading gains for males in the Comparison group, and for females in the WERP than for females in the Comparison group.

- WERP reading gains were greater for Whites, Hispanics, African Americans, Native Americans, and Asians than for their counterparts in the Comparison group.

- WERP reading gains of White, African American, Hispanic, and Asian kindergartners were greater than the reading gains of White kindergartners in the Comparison group.

- WERP reading gains kindergartners with a primary home language of English, Spanish, and other languages were greater than the reading gains of their counterparts in the Comparison group.

- WERP reading gains of kindergartners with a primary home language of Spanish were greater than the reading gains of English primary home language kindergartners in the Comparison group. That is, WERP Spanish primary home language students who were learning English reading skills outperformed the gains of the Comparison group English primary home language students.

- WERP reading gains of kindergartners in four different quartile levels of reading achievement outperformed the reading gains of Comparison students with the largest gains in the top (fourth) quartile.

- WERP English language learners (ELL) reading gains were greater than the reading gains of the Comparison ELL group.

- WERP ELL student reading gains were greater than the reading gains of the non-English language learners (native speakers) in the Comparison group.

- Usage of the WERP software was found to be significantly correlated with the reading outcome measures and pretest to posttest gains in the outcome measures. This suggests that the more the student experiences the WERP content, the greater the reading gains.

- Findings from the average reading score gains by minutes of usage analyses indicate the WERP group quickly closed the gap with the Comparison group and significantly outperformed the Comparison group starting with 1501 minutes of usage.

APPENDICES

Table 19. WERP and Comparison Schools on the DIBELS Total Reading Score

WERP Schools		Pretest		Posttest		Gain
	N	M	SD	M	SD	
School A	13	10.02	5.29	33.71	8.27	23.69
School B	43	8.13	6.15	29.41	14.64	21.29
School C	13	10.03	4.72	30.51	7.50	20.48
School D	41	8.22	6.64	30.69	12.72	22.47
School E	1	0.00	.00	14.00	.00	14.00
School F	3	16.87	2.14	33.87	4.70	17.00
School G	6	1.29	.71	10.07	8.02	8.78
School H	52	12.86	6.32	32.98	8.96	20.12
School J	83	11.24	6.82	35.76	11.35	24.52
School K	79	11.89	7.76	37.99	12.68	26.10
Total	334	10.62	6.98	33.59	12.42	22.97
Comparison Schools		**Pretest**		**Posttest**		**Gain**
	N	M	SD	M	SD	
School M	92	13.06	9.34	29.56	12.67	16.50
School N	90	12.27	8.45	34.33	13.40	22.06
School O	63	12.08	8.33	32.86	10.86	20.78
School P	95	8.10	8.00	28.65	11.80	20.55
School Q	47	10.51	6.26	38.30	11.94	27.79
School R	58	10.18	7.67	30.74	12.87	20.57
School S	102	8.14	6.74	22.56	12.03	14.42
School T	112	13.30	8.90	32.54	13.51	19.24
School U	77	7.14	6.49	19.67	12.64	12.53
School V	97	11.90	9.63	31.97	12.45	20.07
School W	49	8.51	7.44	30.74	11.85	22.34
School X	69	11.33	8.84	32.63	15.65	21.31
School Y	114	9.89	8.07	29.20	12.26	19.31
School Z	51	11.18	7.52	31.51	12.43	20.33
School AA	95	5.70	5.99	21.83	10.79	16.12
Total	1211	10.22	8.30	29.32	13.33	19.10

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.

Table 20. WERP and Comparison Schools on DIBELS Total Reading Percentiles

WERP Schools		Pretest		Posttest		Gain
	N	M	SD	M	SD	
School A	13	49.94	11.90	55.06	16.05	5.12
School B	43	41.43	17.37	47.07	23.08	5.63
School C	13	48.25	12.88	47.63	13.09	-0.62
School D	41	41.07	14.93	49.78	21.14	8.70
School E	1	15.80	.00	17.60	.00	1.80
School F	3	62.13	1.45	60.60	11.68	-1.53
School G	6	20.50	4.41	17.80	14.99	-2.70
School H	52	54.58	14.88	54.38	15.84	-0.20
School J	83	46.67	16.39	62.45	16.79	15.78
School K	79	49.10	16.50	63.11	19.03	14.01
Total	334	46.88	16.63	56.00	20.12	9.12
Comparison Schools		**Pretest**		**Posttest**		**Gain**
	N	M	SD	M	SD	
School M	92	54.85	19.87	48.66	21.45	-6.19
School N	90	51.50	18.61	58.16	22.15	6.66
School O	63	51.82	19.77	56.80	18.14	4.97
School P	95	42.79	17.23	44.66	19.89	1.87
School Q	47	46.79	15.61	61.13	15.88	14.34
School R	58	47.35	18.48	49.49	19.53	2.13
School S	102	42.76	16.17	36.76	20.08	-6.01
School T	112	52.28	17.74	52.47	19.56	0.19
School U	77	40.83	17.58	32.42	20.40	-8.41
School V	97	44.58	19.99	54.28	21.53	9.70
School W	49	44.48	19.81	49.74	19.55	5.26
School X	68	45.29	22.02	49.48	22.93	4.19
School Y	114	46.92	18.16	48.33	20.59	1.42
School Z	51	48.33	16.39	51.27	21.24	2.95
School AA	95	36.75	16.30	36.62	17.91	-0.13
Total	1210	46.45	18.84	47.99	21.60	1.54

Note. WERP students selected with 1100 minutes (6 months) or more usage of Waterford Early Reading Program. The pretest-posttest effect size is the mean gain divided by the standard deviation (Walberg 2001).
*p < .05, ** p < . 01, *** p < .001 from independent t tests comparing gains.

Table 21. Comparison of all WERP and Comparison Schools Students on DIBELS Total Reading Score

WERP Schools		Pretest		Posttest		Gain
	N	M	SD	M	SD	
School A	40	8.25	5.85	29.29	10.47	21.04
School B	68	7.56	6.08	27.17	13.14	19.61
School C	37	7.52	4.76	28.57	8.20	21.05
School D	49	9.00	6.89	31.81	13.00	22.81
School E	60	11.19	7.04	32.28	10.79	21.09
School F	35	9.03	6.87	26.18	11.06	17.15
School G	44	8.58	8.97	25.34	14.18	16.76
School H	66	14.88	8.83	35.34	10.14	20.46
School I	18	8.72	6.44	25.00	10.31	16.28
School J	86	11.25	6.83	35.55	11.33	24.30
School K	80	11.94	7.72	38.26	12.84	26.32
School L	53	7.91	5.60	24.19	12.42	16.28
Total	636	10.08	7.34	31.11	12.58	21.03
Comparison Schools		Pretest		Posttest		Gain
	N	M	SD	M	SD	
School M	92	13.06	9.34	29.56	12.67	16.50
School N	90	12.27	8.45	34.33	13.40	22.06
School O	63	12.08	8.33	32.86	10.86	20.78
School P	95	8.10	8.00	28.65	11.80	20.55
School Q	47	10.51	6.26	38.30	11.94	27.79
School R	58	10.18	7.67	30.74	12.87	20.56
School S	102	8.14	6.74	22.56	12.03	14.42
School T	112	13.30	8.90	32.54	13.51	19.24
School U	77	7.14	6.49	19.67	12.64	12.53
School V	97	11.90	9.63	31.97	12.45	20.07
School W	49	8.51	7.44	30.74	11.85	22.23
School X	69	11.33	8.84	32.63	15.65	21.30
School Y	114	9.89	8.07	29.20	12.26	19.31
School Z	51	11.18	7.52	31.51	12.43	20.33
School AA	95	5.70	5.99	21.83	10.79	16.13
Total	1211	10.22	8.30	29.32	13.33	19.10

Note. All WERP and Comparison students included without any selection.

Table 22. Rank Order of Pretest Means on the DIBELS Total Reading Score

Group	School	Pretest	Posttest	Gain
WERP	School E	0.00	14.00	14.00
WERP	School G	1.29	10.07	8.78
Comparison	School AA	5.70	21.83	16.12
Comparison	School U	7.14	19.67	12.53
Comparison	School P	8.10	28.65	20.55
WERP	School B	8.13	29.41	21.29
Comparison	School S	8.14	22.56	14.42
WERP	School D	8.22	30.69	22.47
Comparison	School W	8.51	30.74	22.34
Comparison	School Y	9.89	29.20	19.31
WERP	School A	10.02	33.71	23.69
WERP	School C	10.03	30.51	20.48
Comparison	School R	10.18	30.74	20.57
Comparison	School Q	10.51	38.30	27.79
Comparison	School Z	11.18	31.51	20.33
WERP	School J	11.24	35.76	24.52
Comparison	School X	11.33	32.63	21.31
WERP	School K	11.89	37.99	26.10
Comparison	School V	11.90	31.97	20.07
Comparison	School O	12.08	32.86	20.78
Comparison	School N	12.27	34.33	22.06
WERP	School H	12.86	32.98	20.12
Comparison	School M	13.06	29.56	16.50
Comparison	School T	13.30	32.54	19.24
WERP:	School F	16.87	33.87	17.00

Note. WERP students selected with 1100 minutes (6 months) or more usage of WERP Reading Program.

REFERENCES

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Finn, C. E., Rotherham, A. J., & Hokanson, C. R., Jr. (Eds.) (2001). *Rethinking special education for a new century.* Retrieved from http://www.edexcellence.net/ library/special.ed/special.ed.ch12.pdf.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills: Administration and scoring guide.* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement, University of Oregon.

National Association for the Education of Young Children. (1996, April). *Technology and young children – Ages 3 through 8.*[Position statement]. Retrieved from http//www.naeyc.org/resources/position_statements/pstech98.htm

National Association for the Education of Young Children and International Reading Association. (1998). *Learning to read and write: Developmentally appropriate practice for young children.* [Joint position statement]. Retrieved from http://www.naeyc.org/resources/position_statements/psread2.htm.

TUSDStats: *Useful things to know about interpreting AIMS/CCSA data*. (n.d.) Tucson Unified School District. Retrieved July 17, 2006 from http://www.tusd.k12.az.us

Walberg, H. J. (2001) *Final Evaluation of the Reading Initiative.* Albertson Foundation Report

Waterford Institute. (2002). Waterford Early Reading Program and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). *Waterford Teacher Manual*.

Waterford Institute. (2002). Waterford Early Reading Program and No Child Left Behind. *Waterford Teacher Manual.* Waterford Institute.

ABOUT THE AUTHORS

Stephen Powers, Ph.D., was a public school teacher for 12 years, a program evaluator with the Tucson Unified School District for 18 years and president of Creative Research Associates for 12 years. He was president of the Arizona Evaluation Network and the Arizona Educational Research Organization. As an adjunct professor in the Department of Educational Psychology of the University of Arizona and Northern Arizona University he taught research methods and statistics for over 20 years. He has published over 80 articles in research journals on topics of testing, minority assessment, personality, and student achievement.

Connie Price-Johnson, M.A., has been a research evaluator for four years with extensive experience in the evaluation of educational programs. She has expertise in second language learning and social marketing, and recently completed a national study of school districts with exemplary communication programs.