

Evaluation Report

Differences in Early Literacy Quick Assessment (ELQA) Scores for Students of Teachers Trained in Energizing Readiness (ER) and Students of Teachers with Demographically Similar Classrooms Not Trained in ER

Research conducted by:
Betty Harris, M.A.
Lucy Trautman, M.Ed.

The University of Oklahoma
College of Continuing Education
eteam

Educational Training, Evaluation, Assessment, and Measurement
1639 Cross Center Drive, Suite 308
Norman, OK 73019-5050
<http://eteam.ou.edu/>
September 30, 2011

Executive Summary

This report presents the findings from the Energizing Readiness (ER) Early Literacy Quick Assessment (ELQA) study conducted by The University of Oklahoma Department of Educational Training, Evaluation, Assessment and Measurement (E-TEAM), Division of Public and Community Services, College of Continuing Education at the request of Payne Education Center (PEC). Energizing Readiness (ER) is a Pre-kindergarten curriculum developed by Sue Kirk. This study will compare student ELQA scores from classrooms of teachers who have been trained in the ER curriculum to student scores from classrooms of teachers who have not been trained in the ER curriculum to determine if students from ER classrooms score higher on ELQA subtests.

PEC offered the use of the Early Literacy Quick Assessment during the 2010-2011 school year to teachers training in ER. The Early Literacy Quick Assessment (ELQA) is a software assessment tool that facilitates formative progress monitoring of early literacy skills (Picture Naming (Expressive Vocabulary, Receptive Vocabulary, Print Concepts, Rhyming, and Uppercase Letters). After the end of the school year, data from ER-trained teachers' ELQA assessments were compared to data from Oklahoma teachers with demographically similar classrooms who were not trained in ER.

ER lessons are organized around rhymes and the alphabet, include multiple activities designed to increase student vocabulary and, to a lesser degree, include activities to promote the development of print concepts (as measured by the ELQA). Across the five ELQA subtests, the largest difference between the two groups was for Rhyming, followed by Uppercase Letters and Picture Naming (Expressive Vocabulary). The average student in ER classrooms scored higher than 58-62% of comparison students on these three ELQA assessments. For both the Rhyming and Uppercase Letters assessments, the spread of student scores was more homogenous toward the end of the school year which indicates that there were fewer students who scored low on those assessments in ER Classrooms. This trend of decreased variability in student scores was also evident for Picture Naming; however, it did not reach statistical significance. ER students scored significantly higher at all five time periods on Rhyming. The statistically significant difference at Time 1 suggests that by the time the assessment was administered, students in ER classrooms had received enough rhyming instruction to move them ahead of students whose teachers were not trained in ER.

Scores on Receptive Vocabulary were in the expected direction with ER students scoring higher than the Non-ER students and approached statistical

significance when controlling for Time 1 scores. The ER group scored significantly higher at all five time periods on Receptive Vocabulary. Scores for Print Concepts were not in the expected direction using the adjusted ANCOVA means which control for Time 1 scores; however, using the unadjusted means, ER students scored significantly higher at Time 2 and trended toward scoring higher at Time 4. By Time 5, there were no statistically significant differences between the two groups on Print Concepts.

Recommendations include:

- Placing more emphasis on a broader array of print concepts in the ER curriculum and providing teachers with more documentation of developmentally appropriate multisensory activities to reinforce learning print concepts.
- Providing more detailed instructions in the curriculum for methods teachers can use to increase student vocabulary.
- Reinforcing the importance of both structured vocabulary instruction and taking advantage of opportunities to introduce and reinforce new vocabulary during ER training.
- Collecting data to determine if ER students score higher on other early literacy skills not assessed by the ELQA (e.g. blending, segmenting, beginning sounds, ending sounds, sound substitution, etc).
- Conducting a larger study that includes collecting data on both student early literacy scores over time and on their teachers' education, teaching methods, curricula and behavior to further establish the relationship between the ER curriculum/training and student scores and teacher behavior.

Table of Contents

Executive Summary	i
Background and Purpose.....	1
ER Curriculum	1
Methods.....	3
Recruitment	3
The Early Literacy Quick Assessment	4
Study Samples	5
Results.....	7
Picture Naming (Expressive Vocabulary).....	8
Receptive Vocabulary	11
Print Concepts.....	13
Rhyming	16
Uppercase Letters	19
Discussion	23
References	25
Appendix A. Box Plots.....	26

Table of Tables

Table 1. ER Lesson Components and Activities.....	1
Table 2. Student Demographics, Initial Sample	6
Table 3. Student Demographics after non-ER Sample was Adjusted	6
Table 4. Number and Percentage of Students Tested by Group	7
Table 5. Average Test Dates and Estimated Number of Days into the School Year Students were Tested by Group	7
Table 6. Picture Naming: Average Number Correct, Standard Deviation and p Values by Group and Time Period	9
Table 7. Picture Naming Score Categories by Group and Time Period.....	10
Table 8. Receptive Vocabulary: Average Number Correct, Standard Deviation and p Values by Group and Time Period	11
Table 9. Receptive Vocabulary Score Categories by Group and Time Period....	13
Table 10. Print Concepts: Average Number Correct, Standard Deviation and p Values by Group and Time Period	14
Table 11. Print Concepts Score Categories by Group and Time Period	15
Table 12. Rhyming: Average Number Correct, Standard Deviation and p Values by Group and Time Period	17
Table 13. Rhyming Score Categories by Group and Time Period.....	18
Table 14. Uppercase Letters: Average Number Correct, Standard Deviation and p Values by Group and Time Period.....	20
Table 15. Uppercase Letters Score Categories by Group and Time Period.....	21
Table 16. Adjusted Means, Effect Sizes and p values by ELQA Subtest	23

Table of Figures

Figure 1. Picture Naming: Average Number Correct by Group and Time period .	9
Figure 2. Picture Naming Score Categories by Group and Time Period.....	10
Figure 3. Receptive Vocabulary: Average Number Correct by Group and Time period.....	12
Figure 4. Receptive Vocabulary Score Categories by Group and Time Period...	13
Figure 5. Print Concepts: Average Number Correct by Group and Time period.	14
Figure 6. Print Concepts Score Categories by Group and Time Period.....	16
Figure 7. Rhyming: Average Number Correct by Group and Time period	17
Figure 8. Rhyming Score Categories by Group and Time Period.....	19
Figure 9. Uppercase Letters: Average Number Correct by Group and Time period.....	20
Figure 10. Uppercase Letters Score Categories by Group and Time Period.....	22
Figure 11. Box Plots: Picture Naming Scores by Group and Time Period	27
Figure 12. Box Plots: Receptive Vocabulary Scores by Group and Time Period...	28
Figure 13. Box Plots: Print Concepts Scores by Group and Time Period.....	29
Figure 14. Box Plots: Rhyming Scores by Group and Time Period.....	30
Figure 15. Box Plots: Uppercase Letters Scores by Group and Time Period.....	31

Differences in Early Literacy Quick Assessment (ELQA) Scores for Students of Teachers Trained in Energizing Readiness (ER) and Students of Teachers with Demographically Similar Classrooms Not Trained in ER

Background and Purpose

This report presents the findings from the Energizing Readiness (ER) Early Literacy Quick Assessment (ELQA) study conducted by The University of Oklahoma Department of Educational Training, Evaluation, Assessment and Measurement (E-TEAM), Division of Public and Community Services, College of Continuing Education at the request of Payne Education Center (PEC). Energizing Readiness (ER) is a Pre-kindergarten curriculum developed by Sue Kirk. This study will compare student ELQA scores from classrooms of teachers who have been trained in the ER curriculum to student scores from classrooms of teachers who have not been trained in the ER curriculum to determine if students from ER classrooms score higher on ELQA subtests.

ER Curriculum

The ER curriculum has nine lesson components and 32 activities within those components. There is an introductory lesson, one lesson for each letter (A-Z), one lesson for each number (0-10) and a final number lesson that covers the numbers learned. Across the 39 ER lessons, there are 1,082 activities. The majority of activities are repeated across all 39 ER lessons; however, not all activities occur in each lesson (see Table 1). ER uses a multisensory developmentally appropriate approach that provides multiple opportunities for students to experience the information to be learned in various ways (e.g. visual, auditory, kinesthetic, tactile, etc.) and includes opportunities for discovery learning.

Table 1. ER Lesson Components and Activities

Component/Activity	Description	In Lesson(s)	# Activities
Overall			1082
<i>Alphabet/Number Engine</i>			113
Touch	Teacher asks students to discover the target alphabet letter by manipulating a three dimensional letter and describing how it feels.	A-Z, 0-10	37
Name	Students are asked to guess the name of the target alphabet letter and verify if their guesses were correct.	A-Z, 0-10	37
Match	Teacher has students place new letter on Alpha-Mat in correct position. Letters discovered in previous lessons are poured out from letter bank. Students name each letter in order while placing on mat and when returning to container.	A-Z, 0-10, "1-10"	38
Count		"1-10"	1
<i>Copy Cat Tracks</i>			152
Trace It	Teacher models while students discover strokes for forming target letter by tracing in the air using large muscles.	Intro, A-Z, 0-10	38
Talk It Through	Teacher has students tell strokes while tracing target letter in the air two more times.	Intro, A-Z, 0-10	38

Component/Activity	Description	In Lesson(s)	# Activities
Skywrite	Teacher has students skywrite the target letter in the air using large muscles, and then skywrite the letter two more times.	Intro, A-Z, 0-10	38
Tell a Buddy	Teacher has students working in pairs, one student skywrites while the other talks through various strokes for forming letters. Students switch roles, then stand side by side and together say and skywrite strokes.	Intro, A-Z, 0-10	38
Rhyme Time			157
Rhyme of the Week	Teacher introduces ideas in rhyme of the week using guessing box, rhyme box, feely box, visual folder, puppets, cloze procedure, dress up clothes, visual props, auditory clues, guest, and/or student leader.	Intro, A-Z, 0-10, "1-10"	39
Hear it Say it	Teacher models saying the whole rhyme of the week to the class. Students join in saying the rhyme. Teacher may use word strips in pocket chart to point to each word with number pointer if desired.	Intro, A-Z, 0-10, "1-10"	39
Vocabulary Stretchers	Teacher introduces key vocabulary words identified from rhyme-of-the-week and may provide repeated exposures.	Intro, A-Z, 0-10, "1-10"	39
Historical Note	Teacher shares interesting history about rhyme-of-the-week with students when appropriate.	Intro, A-Z, 0-10, "1-10"	39
Discover	Teacher says, " <i>In our rhyme today we counted. What do we use to count? (Numbers, Teacher may point to a number line.) Now look at these (pointing to an alphabet). What are they? (Yes, letters.) We will be making great discoveries with our letters or alphabet this year. Does anyone know why we have letters?</i> " Teacher continues discussion to help students understand importance of letters, and how the letters will unlock the code of reading for them. Teacher may read book <i>Alphabet Tree</i> by Leo Lionni.	Intro	1
On-Line With The Rhyme			195
Who	Teacher has students identify the subject of the rhyme as teacher adds the "who" piece to the rhyme-o-gram.	Intro, A-Z, 0-10, "1-10"	39
What	Teacher has students tell what the "who" did or what happened in the rhyme as teacher adds the "what" piece to the rhyme-o-gram.	Intro, A-Z, 0-10, "1-10"	39
When	Teacher has students identify when the story rhyme is happening and add the "when" piece to the rhyme-o-gram.	Intro, A-Z, 0-10, "1-10"	39
Where	Teacher has students identify where the story rhyme is happening and add the "where" piece to the rhyme-o-gram.	Intro, A-Z, 0-10, "1-10"	39
Recap	After multiple practices with reciting the story rhyme, teacher has students retell the rhyme in a summary sentence using the rhyme-o-gram pieces.	Intro, A-Z, 0-10, "1-10"	39
Sound The Whistle			114
Rhyme Chime	Teacher asks, " <i>What rhymes with "word"?(_word_) Let's build a rhyming _word_ today. See how many words you can name that rhyme with _word_. We will add a piece of puzzle each time. Let's see if we can complete our _word_.</i> " Possible list:	Intro, A-Z, 0-10	38
Ears On	Teacher says, <i>Let's give a _action_ for (lesson letter). _action_! _action_! _action_! Listen, echo. (lesson letter) If you hear (lesson letter) in these words _action_, if you don't hear (lesson letter) stand still. Listen, echo.</i>	Intro, A-Z, 0-10	38
Sound Play	Various activities: (e.g. Listen and echo, blending, segmenting, rhyme completion, rhyme generation, beginning sounds, ending sounds, sound substitution, leave off sound, compound words, one to one match, sequentially leave off words at end of sentence.	Intro, A-Z, 0-10	38

Component/Activity	Description	In Lesson(s)	# Activities
<i>Fuel The Train</i>			117
Get Ready	The teacher asks open-ended questions relating rhyme content to student life experience to promote oral sharing and discussion.	Intro, A-Z, 0-10, "1-10"	39
Get Set	Students ask and answer questions using TELL (Think, Edit-express, Launch, and Lead) Procedure. Teacher scaffolds student responses using visual clues in TELL Cards to help guide and encourage use of expressive language.	Intro, A-Z, 0-10, "1-10"	39
Go	Teacher provides model sentence starter related to the rhyme of the week for students to personalize with their own ideas.	Intro, A-Z, 0-10, "1-10"	39
<i>Creation Station</i>			78
Picture It	Students illustrate the rhyme of the week the way they visualize or picture it in their mind.	Intro, A-Z, 0-10, "1-10"	39
Sensory Sensations	Teacher uses multisensory, hands on activities with students to reinforce rhyme concepts, letter shape, and letter sound.	Intro, A-Z, 0-10, "1-10"	39
<i>The Choosin Depot</i>			117
Circle Up	Teacher has students practice singing sentences with words that begin with the lesson letter while socializing and interacting with peers (e.g. rhyme sung to <i>Farmer in the Dell</i>).	Intro, A-Z, 0-10, "1-10"	39
Twist Your Tongue	Teacher has students repeat tongue twisters with words starting with the lesson letter.	A-Z	26
More Fun	Various activities: Act out rhyme, make curds and whey, write wishes, draw picture of home, game with shoes, game based on <i>Doggie Doggie</i> , graph favorite teas, graph favorite ways to eat eggs, graph favorite pickles, graph favorite season, make pie graph, color mixing with play dough, dress up, eat through alphabet, rhyme generation and students illustrate to make book, students illustrate when I grow up and make into class booklet, make hats, make pies, play <i>Eensy, Weensy Spider, Where's your E?</i> , practice fire drill, practice fire safety rules, sequence the rhyme, pretend market, share good dream, students change the time and think of new rhymes, one shoe off one shoe on activities, think of words that start like your name and that describe you.	Intro, A-Z, 0-10, "1-10"	39
Number Hunt	Walk around the room, school, and playground. Have students hunt for items that are the Lesson number.	0-10, "1-10"	12
Book and Print Awareness	Teacher has students use cut-outs of engine and caboose to label front and back of book.	Intro	1
<i>Story Book Caboose</i>			39
Book, Author, ISBN Number	Teacher reads books out loud related to lesson rhyme concepts or target letter for listening and extensions.	Intro, A-Z, 0-10, "1-10"	39

Methods

Recruitment

An e-mail was sent to all ER trained teachers on September 16, 2010 offering them free use of the ELOA for the 2010-2011 school year. Teachers were also given information about the free EQLA trial at ER trainings held during summer 2010. During the school year, all ELOA users received an e-mail reminding them of the approaching end of each testing period.

The Early Literacy Quick Assessment

The Early Literacy Quick Assessment (ELQA) is a software assessment tool that facilitates formative progress monitoring of early literacy skills. Children's early literacy skills are assessed at the beginning of the school year and again at intervals throughout the school year so that teachers are able to 1) identify children at risk for not meeting early literacy benchmarks, and 2) differentiate instruction to prepare young children to enter kindergarten with the necessary language and literacy skills to benefit from formal school instruction. ELQA consists of the following assessments.

Alphabet Knowledge

- **Uppercase Letters:** Students are shown 26 uppercase letters and are asked to name each letter.
- **Lowercase Letters:** If students correctly name 16 or more uppercase letters, they are shown 26 lowercase letters and are asked to name each letter.
- **Letter Sounds:** If students correctly name 9 or more lowercase letters, they are shown 26 uppercase letters and asked for the corresponding letter sounds.

Print Concepts

This assessment consists of ten items used to assess the child's knowledge of aspects of reading. The teacher uses an engaging early level text and asks the child questions about the book (e.g. story title, where to begin reading, which direction to read, return sweep, word by word pointing, meaning of period, etc.). An image that displays letters, numbers, and words is used to determine if the child can point to a letter, a word and the first letter in a word.

Phonological Awareness

- **Rhyme Recognition:** Students are asked to decide whether each of five pairs of words rhymes (e.g., cat and hat).
- **Rhyme Generation:** Students are asked to generate a word that rhymes with five different words spoken by the teacher.

Vocabulary

- **Receptive Vocabulary:** Students are shown a picture or sets of pictures and asked to point to the picture that represents the concept presented by the teacher. Concepts include: Body Parts, Directional Terms, Shapes, Ordinal Terms, Transportation, Community, Nature, Geography, Emotions, and Clothing.
- **Expressive Vocabulary:** Students are shown 25 pictures and are asked to name the object shown in each picture. Concepts include: Animals, Body Parts, Clothing, Food, Household Objects, Nature, Common Objects, People, Tools, Toys, and Transportation.

ELQA Reports

The ELQA software generates individual student and classroom-level reports. The student report shows student scores for each subtest over time. The classroom report illustrates class progression over time and provides recommended flexible groupings according to the ELQA scores. The group reports are used to facilitate differentiated instruction. The individual reports can be used to help facilitate transition from one classroom to another and may be used to aid communication during parent-teacher conferences.

Study Samples

After the completion of the school year, data were compiled and ER pre-K classrooms were identified by crosswalking the list of ER trained teachers to the ELQA database. Data was pulled from all other Oklahoma ELQA Pre-K classrooms to form the initial comparison group. Students were included if they were tested at all five time periods during the 2010-2011 school year on any ELQA subtest. The initial sample consisted of 14 ER teachers with 190 students in 15 classrooms in eight schools and 16 non-ER teachers with 223 students in 16 classrooms in five schools.

Initial student demographics are presented in Table 2. There were some differences between the two samples initially; however, there was only one demographic variable that was significantly different between the two groups ($p < .05$). The proportion of students who were Hispanic differed significantly between the two groups with the non-ER group having a much higher proportion of Hispanic students. Four other characteristics approached statistical significance: Age in years (due to the low variability of age as measured in years), the proportion of students who were African American, the proportion of students who were white, and the proportion of students whose language at home was not English or Spanish. The race/ethnicity and language spoken at home data should be interpreted with caution as both had a considerable amount of missing data (race 44%-46% missing, language (14%-45% missing). Thirty-one students were removed from the non-ER group to equalize the groups in terms of age, gender, race/ethnicity and language spoken at home (see Table 3). After adjustment there were no statistically significant differences in demographics between the two groups.

Table 2. Student Demographics, Initial Sample

Demographics	Mean			SD		N		<i>p</i>
	ER	Non-ER	Difference	ER	Non-ER	ER	Non-ER	
Age (Yrs)	4.1	4.1	0.1	0.4	0.3	196	221	0.08
Age (Months)	54.9	54.6	0.3	3.8	4.0	196	221	0.45
Age Missing						0%	1%	
Gender								
Male	48%	54%	-6%	0.5	0.5	196	223	0.24
Race/Ethnicity								
Hispanic	4%	11%	-7%	0.2	0.3	105	125	0.03
African American	9%	15%	-7%	0.3	0.4	105	125	0.12
White	54%	44%	10%	0.5	0.5	105	125	0.12
Other Race	33%	30%	4%	0.5	0.5	105	125	0.55
Race Missing						46%	44%	
Language								
English	93%	93%	-1%	0.3	0.2	168	123	0.83
Spanish	4%	6%	-2%	0.2	0.2	168	123	0.39
Other	4%	1%	3%	0.2	0.1	168	123	0.10
Language Missing						14%	45%	

Table 3. Student Demographics after non-ER Sample was Adjusted

Demographics	Mean			SD		N		<i>p</i>
	ER	Non-ER	Difference	ER	Non-ER	ER	Non-ER	
Age (Yrs)	4.1	4.1	0.0	0.4	0.3	196	190	0.21
Age (Months)	54.9	54.7	0.2	3.8	4.0	196	190	0.65
Age Missing						0%	1%	
Gender								
Male	48%	52%	-3%	0.5	0.5	196	192	0.54
Race/Ethnicity								
Hispanic	4%	6%	-2%	0.2	0.2	105	102	0.49
African American	9%	7%	2%	0.3	0.3	105	102	0.65
White	54%	54%	0%	0.5	0.5	105	102	0.96
Other Race	33%	33%	0%	0.5	0.5	105	102	1.00
Race Missing						46%	47%	
Language								
English	93%	97%	-4%	0.3	0.2	168	94	0.14
Spanish	4%	2%	1%	0.2	0.1	168	94	0.52
Other	4%	1%	3%	0.2	0.1	168	94	0.16
Language Missing						14%	51%	

Table 4 displays the number and percentage of students tested by group. Across the five ER subtests, ER teachers tested 96.4% of students and non-ER teachers tested 97.8% of students included in this study. The highest testing rate was for Picture Naming (100%) followed by Rhyming (99.0%) and Uppercase Letters (99.0%), Print Concepts (98.5%) and Receptive Vocabulary (98.2%).

Table 4. Number and Percentage of Students Tested by Group

	Number Tested		Percent Tested		Difference
	ER	Non-ER	ER	Non-ER	
Total Students	196	192			
Students Tested					
Picture Naming	196	192	100.0%	100.0%	0.0%
Print concepts	190	192	96.9%	100.0%	-3.1%
Receptive Vocabulary	194	187	99.0%	97.4%	1.6%
Rhyming	193	191	98.5%	99.5%	-1.0%
Uppercase Letters	172	177	87.8%	92.2%	-4.4%

August, 15 2010 was used as an overall start date to estimate the average number of days in students were in school when tested (see Table 5). ER and Non-ER students had been in school for approximately one month prior to time period one testing. On average, ER students were tested one day later in time periods 1 and 2, two days later in time periods 3 and 4 and nine days later in time period 5 than were non-ER students.

Table 5. Average Test Dates and Estimated Number of Days into the School Year Students were Tested by Group

Group	Time Period				
	1	2	3	4	5
Estimated School Days (from 8/15/2010)					
Non-ER	40	96	160	220	264
ER	41	97	158	222	273
Average	40	96	159	221	268
Average Test Date					
Non-ER	9/23/2010	11/18/2010	1/21/2011	3/22/2011	5/05/2011
ER	9/24/2010	11/19/2010	1/19/2011	3/23/2011	5/14/2011

Results

This report compares the ER and non-ER student groups in terms of their average student scores and the variability of student scores on each ELQA subtest over five time periods during the 2010-2011 school year. The statistical method used is an analysis of covariance (ANCOVA). The ANCOVA compares group scores at time period 5 while controlling for student scores at time period 1 testing. This method adjusts for any differences in individual student scores at time period 1 testing. The two groups are considered statistically different at time period 5 testing if the *p* value is less than or equal to 0.05. The groups are also compared on student scores at each time period using an independent samples t-test. The t-test determines if the differences between the two groups are large enough to not have occurred by chance. One tailed statistical tests are used to test the hypothesis that ER students should score higher than non-ER students. If the training in the ER curriculum works to increase scores on the ELQA, one would expect average student scores for the ER group to be higher than the non-ER group after the ER teaching methods have been implemented

in the classroom long enough to have an effect on student scores. Effect sizes and Cohen's U_3 are commonly used measures that indicate the size of differences between groups. They provide a standardized means of looking at the size of group differences across studies. Effect sizes (Hedge's g) and Cohens U_3 index were calculated using formulas from the What Works Clearinghouse Procedures Guide (2008). Hedges's g was calculated using covariate adjusted mean difference divided by unadjusted pooled within-group SD. The What Works Clearinghouse Procedures Guide (2008) considers effects that are not statistically significant but have an effect size of at least 0.25 "substantively important".

Picture Naming (Expressive Vocabulary)

Average scores and standard deviations by group are presented in Table 6. Figure 1 shows average scores over time by group. At time period 1, ER students correctly named on average almost 21 of 25 pictures. By time period 5 testing, the ER students correctly named a little over 23 pictures. ER students gained on average 2.7 points between time 1 and time 5 testing and non-ER students gained 2.4 points.

Differences in Average Scores

- When controlling for student scores at time period 1, the ER group scored significantly higher at time period 5 on their Picture Naming scores than did the non-ER group (Adjusted means: 23.4 vs. 23.1, $F(1, 385) = 3.7, p = 0.04$). The effect size for this difference is 0.20 which corresponds to a Cohen's U_3 index of 58%. The average student in ER classrooms scored higher than 58% of the non-ER students on their Picture Naming scores (scored eight percentile points higher than an average non-ER group student).
- The ER group scored significantly higher in Picture Naming at time period 4 ($t(386) = -1.7, p = 0.04$) and at time period 5 ($t(386) = 1.8, p = 0.04$). The ER group also trended toward scoring higher in Picture Naming at time period 2 ($t(382) = 1.5, p = 0.06$).

Differences in Variability

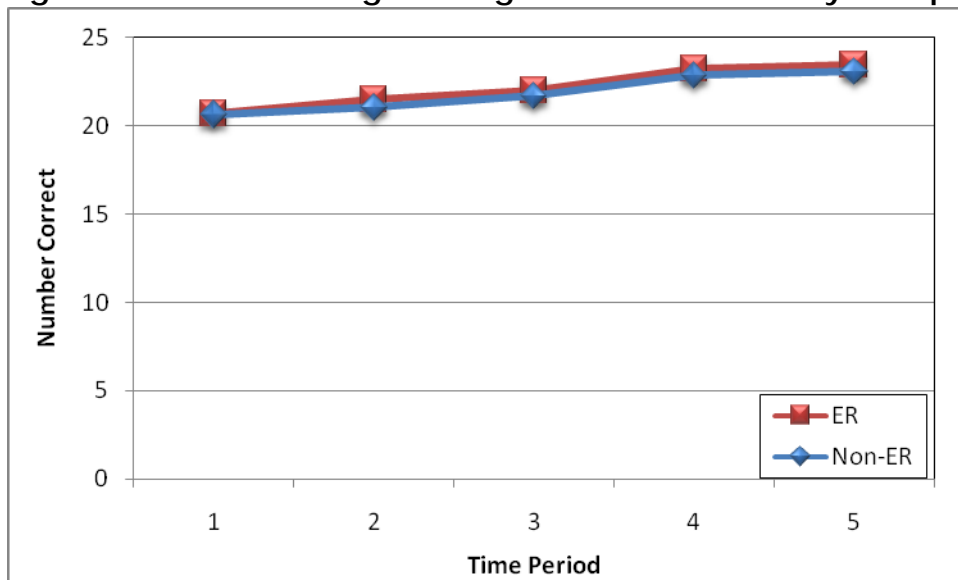
- The variability in scores for ER group was significantly smaller than for the non-ER group at time period 2 ($p = .02$) and approached statistical significance at time period 5 ($p = .09$). This means there was less spread in student scores for the ER group than for the non-ER group at time period 2 and a trend toward less spread at time period 5. See Appendix A for box plots of subtest scores by group and time period.

Table 6. Picture Naming: Average Number Correct, Standard Deviation and p Values by Group and Time Period

Group	Time Period					Gain
	1	2	3	4	5	
Average Number Correct						
Non-ER	20.6	21.0	21.7	22.8	23.1	2.4
ER	20.7	21.5	22.0	23.2	23.4	2.7
Difference	0.1	0.5	0.3	0.4	0.4	0.3
<i>p</i>	0.38	0.06	0.11	0.04	0.04	
Standard Deviation						
Non-ER	3.1	3.3	2.7	2.1	2.2	-1.0
ER	3.5	3.0	2.8	2.1	2.0	-1.5
Difference	0.4	-0.3	0.2	-0.1	-0.1	-0.5
<i>p</i>	0.17	0.02	0.31	0.50	0.09	

*ER N = 196, non-ER N =192

Figure 1. Picture Naming: Average Number Correct by Group and Time period



Scores were grouped into three categories based on the number of pictures the student was able to name correctly: 0-15 correct, 16-20 correct, and 21-25 correct. This allows one to look at how students scored at time period 1 and see where those students ended up at time period 5 (see Table 7 and Figure 2).

Of the students who scored below 16 items correct at time period 1 testing,

- in the ER group 50% moved up to scoring between 21 and 25 items correct at time period 5 testing, compared to 36% in the non-ER group;
- in the ER group 43% moved up to scoring between 16 and 20 items correct at time period 5 testing, compared to 64% in the non-ER group; and
- in the ER group 7% remained in the below 16 items correct group at time period 5 testing, compared to 0% in the non-ER group.

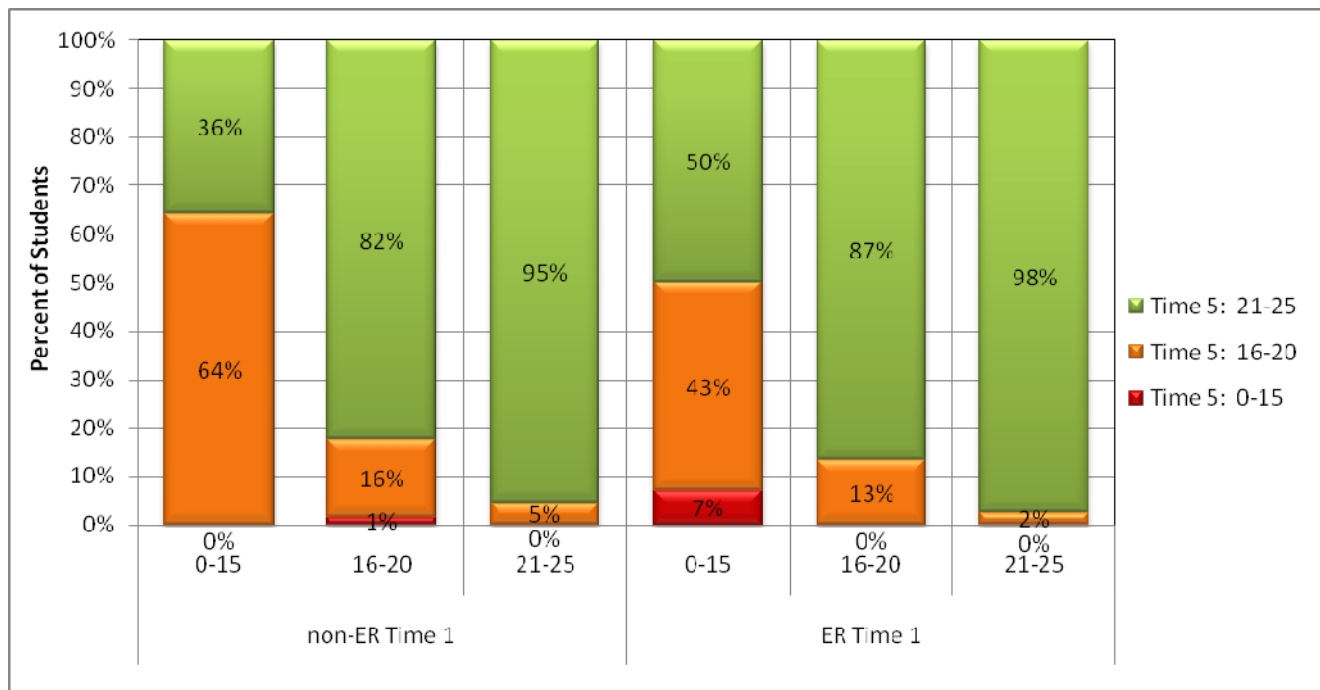
- Of the students who scored between 16 and 20 items correct at time period 1 testing,
- in the ER group 87% moved up to scoring between 21 and 25 items correct at time period 5 testing, compared to 82% in the non-ER group;
 - in the ER group 13% stayed the same scoring between 16 and 20 items correct at time period 5 testing, compared to 16% in the non-ER group; and
 - in the ER group 0% moved down to the below 16 items correct group at time period 5 testing, compared to 1% in the non-ER group.

- Of the students who scored between 21 and 25 items correct at time period 1 testing,
- in the ER group 98% stayed the same scoring between 21 and 25 items correct at time period 5 testing, compared to 95% in the non-ER group;
 - in the ER group 2% moved down scoring between 16 and 20 items correct at time period 5 testing, compared to 5% in the non-ER group; and
 - in the ER group and the non-ER group 0% moved down to the below 16 items correct group at time period 5 testing.

Table 7. Picture Naming Score Categories by Group and Time Period

Group	Time 1 Score Category	Time 5 Score Category						
		Number of Students				Percent of Students		
		0-15	16-20	21-25	N	0-15	16-20	21-25
non-ER	0-15	0	9	5	14	0%	64%	36%
	16-20	1	11	56	68	1%	16%	82%
	21-25	0	5	105	110	0%	5%	95%
ER	0-15	1	6	7	14	7%	43%	50%
	16-20	0	8	52	60	0%	13%	87%
	21-25	0	3	119	122	0%	2%	98%

Figure 2. Picture Naming Score Categories by Group and Time Period



Receptive Vocabulary

Average scores and standard deviations by group are presented in Table 8. At time period 1, ER students correctly identified on average 7.5 pictures out of 10 and non-ER students correctly identified on average 7.0 pictures. At time period 5, ER students correctly identified on average 8.8 pictures resulting in a gain of 1.3 points and non-ER students correctly identified on average 8.5 pictures resulting in a gain of 1.5 points.

Differences in Average Scores

- When controlling for student scores at time period 1, the ER group trended toward higher scores at time 5 on their Receptive Vocabulary scores than did the non-ER group (Adjusted means: 8.8 vs. 8.6, $F(1, 378) = 2, p = 0.08$). The effect size for this difference is 0.15 which corresponds to a Cohen's d index of 0.15. The average student in ER classrooms scored higher than 56% of the non-ER students on their Receptive Vocabulary scores (scored six percentile points higher than an average non-ER group student).
- The ER group scored significantly higher in receptive vocabulary at all five time periods:
 - time period 1 - $t(379) = 2.8, p < 0.01$
 - time period 2 - $t(379) = 3.0, p < 0.01$
 - time period 3 - $t(379) = 2.7, p < 0.01$
 - time period 4 - $t(379) = 3.4, p < 0.01$
 - time period 5 - $t(379) = 2.4, p = 0.01$

Differences in Variability

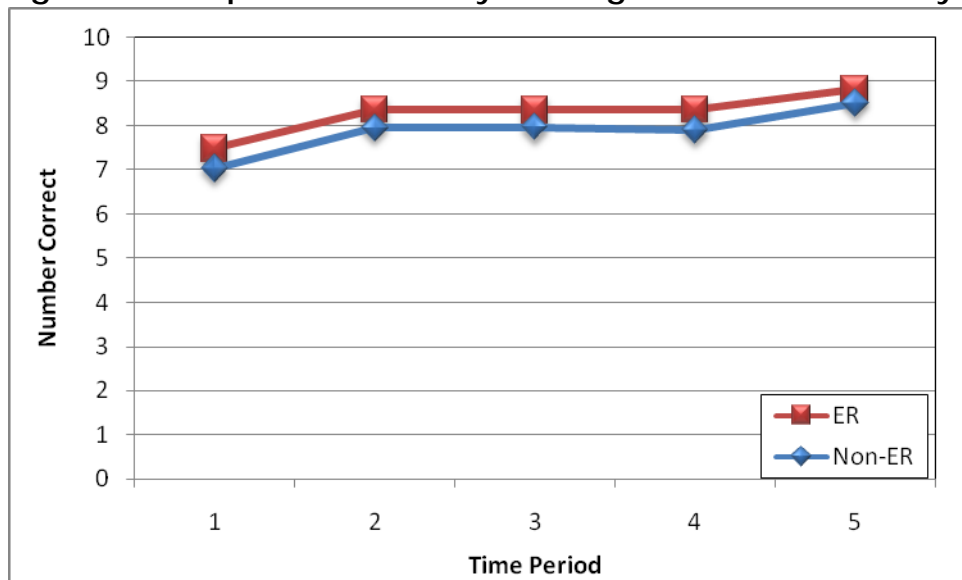
- There were no significant differences in variability of scores between the groups. This means there were no differences in the spread of student scores between the groups.

Table 8. Receptive Vocabulary: Average Number Correct, Standard Deviation and p Values by Group and Time Period

Group	Time Period					Gain
	1	2	3	4	5	
Average Number Correct						
Non-ER	7.0	7.9	8.0	7.9	8.5	1.5
ER	7.5	8.4	8.4	8.4	8.8	1.3
Difference	0.5	0.4	0.4	0.5	0.3	-0.1
p	0.00	0.00	0.00	0.00	0.01	
Standard Deviation						
Non-ER	1.7	1.5	1.4	1.4	1.4	-0.3
ER	1.6	1.3	1.4	1.3	1.3	-0.3
Difference	-0.1	-0.1	0.0	0.0	-0.1	0.0
p	0.49	0.12	0.13	0.28	0.26	

*ER N = 194, non-ER N = 187

Figure 3. Receptive Vocabulary: Average Number Correct by Group and Time period



Scores were grouped into three categories based on the number of pictures the student was able to identify correctly: 2-4 correct, 5-7 correct, and 8-10 correct. This allows one to look at how students scored at time period 1 and see where those students ended up at time period 5 (see Table 9 and Figure 4).

Of the students who scored below 4 items correct at time period 1 testing,

- in the ER group 44% moved up to scoring between 8 and 10 items correct at time period 5 testing, compared to 56% in the non-ER group;
- in the ER group 56% moved up to scoring between 5 and 7 items correct at time period 5 testing, compared to 44% in the non-ER group; and
- in the ER group and the non-ER group 0% remained in the 2-4 items correct group at time period 5 testing.

Of the students who scored between 5 and 7 items correct at time period 1 testing,

- in the ER group 80% moved up to scoring between 8 and 10 items correct at time period 5 testing, compared to 77% in the non-ER group;
- in the ER group 20% stayed the same scoring between 5 and 7 items correct at time period 5 testing, compared to 21% in the non-ER group; and
- in the ER group 0% moved down to the 2-4 items correct group at time period 5 testing, compared to 2% in the non-ER group.

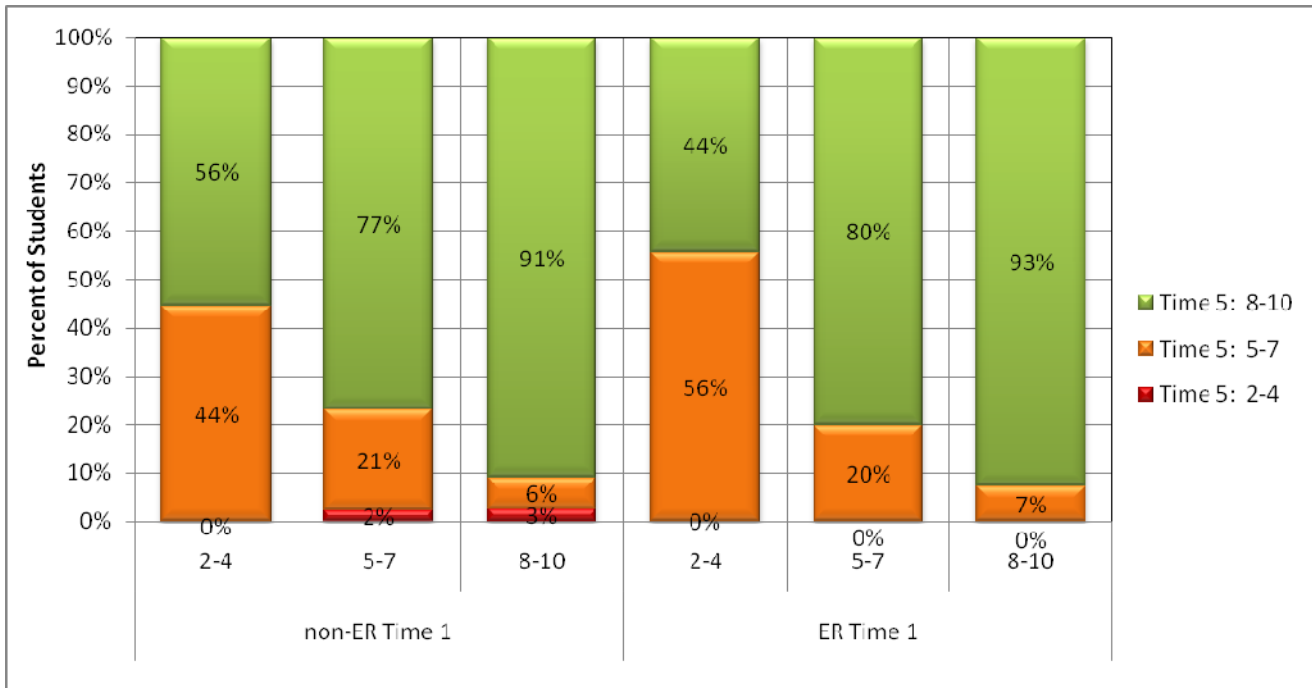
Of the students who scored between 8 and 10 items correct at time period 1 testing,

- in the ER group 93% stayed the same scoring between 8 and 10 items correct at time period 5 testing, compared to 91% in the non-ER group;
- in the ER group 7% moved down scoring between 5 and 7 items correct at time period 5 testing, compared to 6% in the non-ER group; and
- in the ER group 0% moved down to the 2-4 items correct group at time period 5 testing, compared to 3% in the non-ER group.

Table 9. Receptive Vocabulary Score Categories by Group and Time Period

Group	Time 1 Score Category	Time 5 Score Category						
		Number of Students				Percent of Students		
		2-4	5-7	8-10	N	2-4	5-7	8-10
non-ER	2-4	0	8	10	18	0%	44%	56%
	5-7	2	19	70	91	2%	21%	77%
	8-10	2	5	71	78	3%	6%	91%
ER	2-4	0	5	4	9	0%	56%	44%
	5-7	0	15	60	75	0%	20%	80%
	8-10	0	8	102	110	0%	7%	93%

Figure 4. Receptive Vocabulary Score Categories by Group and Time Period



Print Concepts

Average scores and standard deviations by group are presented in Table 10. Figure 5 shows the average score by group over time. At time period 1, ER students answered on average 4.4 items correct out of 10. By time period 5 testing, the ER students answered 8.1 items correctly. On average the ER group gained 3.7 points and the non-ER group gained 4.3 points between time 1 and time 5 testing.

Differences in Average Scores

- When controlling for student scores at time period 1, the ER group trended toward lower scores at time period 5 on their Print Concepts scores than did the non-ER group (Adjusted means: 8.0 vs. 8.3, $F(1, 379) = 2.0, p = 0.08$). The effect size for this difference is -0.15 which corresponds to a Cohen's U_3 index of 44%. The average student in ER classrooms scored higher than 44% of the non-ER

students on their Print Concepts scores (scored six percentile points lower than an average non-ER group student).

- The ER group scored significantly higher in Print Concepts at time period 1 ($t(380) = 2.3, p = 0.01$) and trended toward scoring higher in Print Concepts at time period 4 ($t(363) = 1.5, p = 0.07$).

Differences in Variability

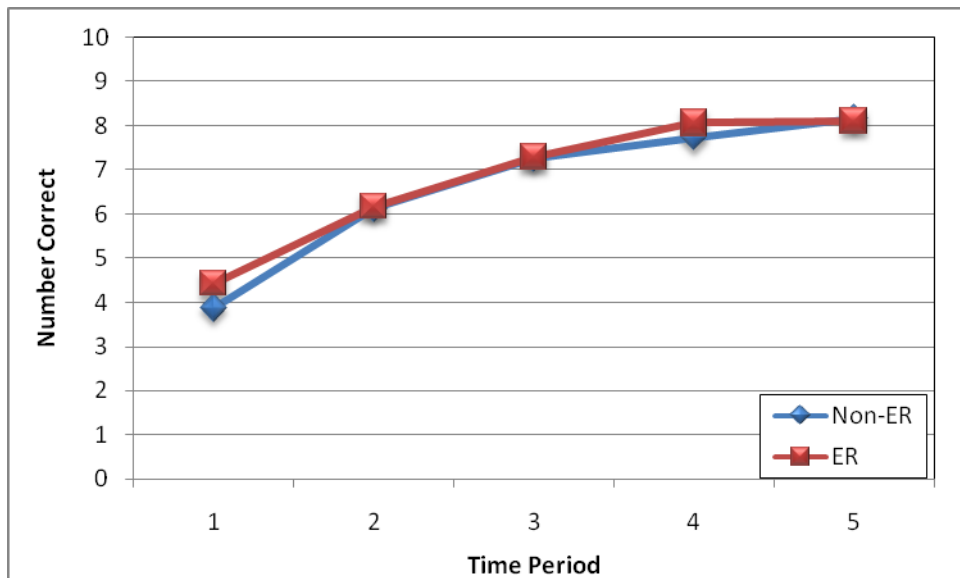
- The variability in scores for ER group was significantly smaller than for the non-ER group at time period 3 ($p = .02$), time period 4 ($p < .01$) and time period 5 ($p < .01$). The two groups were similar on variability of student scores at time periods 1 and 2; however, by time period 3 the variability in group scores decreased more for the ER group. See Appendix A for box plots of subtest scores by group and time period.

Table 10. Print Concepts: Average Number Correct, Standard Deviation and p Values by Group and Time Period

Group	Time Period					Gain
	1	2	3	4	5	
Average Number Correct						
Non-ER	3.9	6.1	7.3	7.7	8.2	4.3
ER	4.4	6.2	7.3	8.1	8.1	3.7
Difference	0.6	0.1	0.0	0.3	0.0	-0.6
p	0.01	0.42	0.43	0.07	0.42	
Standard Deviation						
Non-ER	2.4	2.6	2.5	2.5	2.3	-0.1
ER	2.4	2.6	2.3	2.0	2.0	-0.4
Difference	-0.1	0.0	-0.3	-0.5	-0.3	-0.3
p	0.39	0.46	0.02	0.00	0.01	

*ER N= 190, non-ER N=192

Figure 5. Print Concepts: Average Number Correct by Group and Time period



Scores were grouped into three categories based on the number of items the student was able to identify correctly: 0-2 correct, 3-6 correct, and 7-10 correct. This allows one to look at how students scored at time period 1 and see where those students ended up at time period 5 (see Table 11 and Figure 6).

Of the students who scored below 2 items correct at time period 1 testing,

- in the ER group 60% moved up to scoring between 7 and 10 items correct at time period 5 testing, compared to 65% in the non-ER group;
- in the ER group 30% moved up to scoring between 3 and 6 items correct at time period 5 testing, compared to 28% in the non-ER group; and
- in the ER group 11% remained in the 0-2 items correct group at time period 5 testing, compared to 7% in the non-ER group.

Of the students who scored between 3 and 6 items correct at time period 1 testing,

- in the ER group 88% moved up to scoring between 7 and 10 items correct at time period 5 testing, compared to 87% in the non-ER group;
- in the ER group 12% stayed the same scoring between 3 and 6 items correct at time period 5 testing, compared to 11% in the non-ER group; and
- in the ER group 0% moved down to the 0-2 items correct group at time period 5 testing, compared to 2% in the non-ER group.

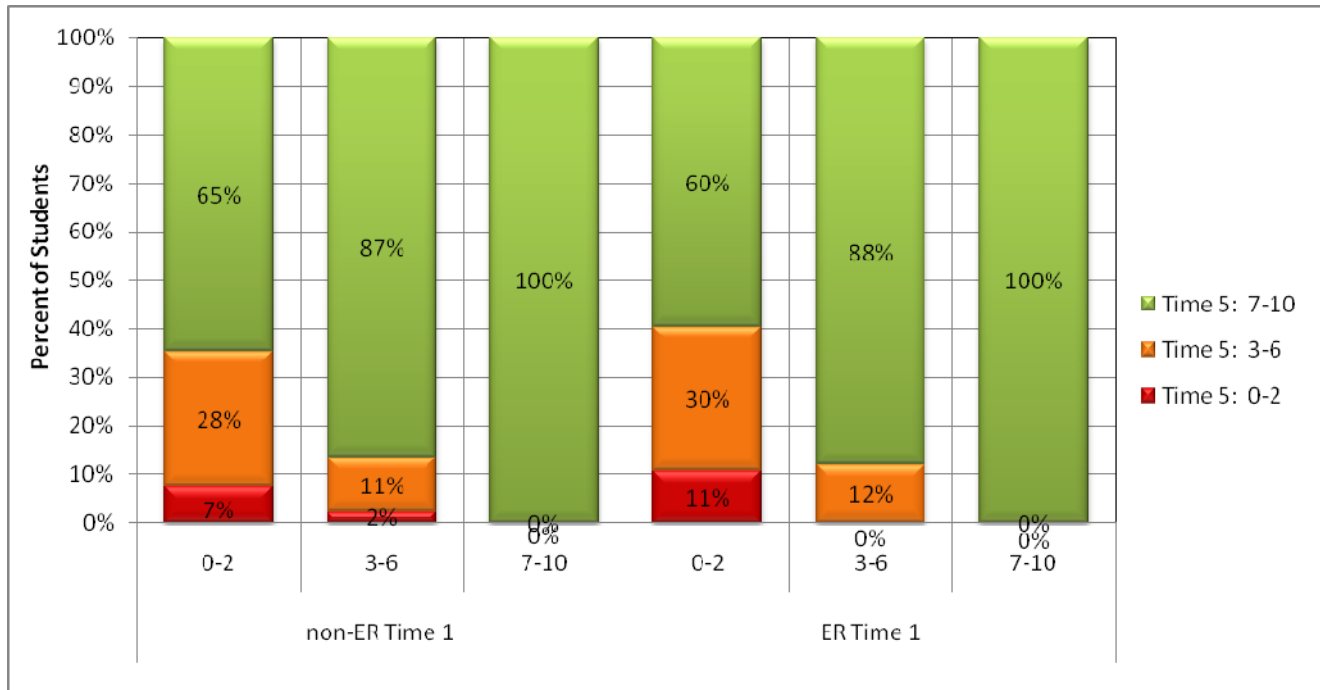
Of the students who scored between 8 and 10 items correct at time period 1 testing,

- in the ER group and the non-ER group 100% stayed the same scoring between 7 and 10 items correct at time period 5 testing.

Table 11. Print Concepts Score Categories by Group and Time Period

Group	Time 1 Score Category	Time 5 Score Category						
		Number of Students				Percent of Students		
		0-2	3-6	7-10	N	0-2	3-6	7-10
non-ER	0-2	5	19	44	68	7%	28%	65%
	3-6	2	10	77	89	2%	11%	87%
	7-10	0	0	35	35	0%	0%	100%
ER	0-2	5	14	28	47	11%	30%	60%
	3-6	0	12	89	101	0%	12%	88%
	7-10	0	0	42	42	0%	0%	100%

Figure 6. Print Concepts Score Categories by Group and Time Period



Rhyming

Average scores and standard deviations by group are presented in Table 12. Figure 7 shows average scores by group over time. At time period 1, ER students scored on average five items correct out of 10. By time period 5 testing, the ER students scored on average 8.8 items correct. ER students gained on average 3.9 points during the school year and non-ER students gained 4.2 points.

Differences in Average Scores

- When controlling for student scores at time period 1, the ER group scored significantly higher at time period 5 on their Rhyming scores than did the non-ER group (Adjusted means: 8.6 vs. 7.9, $F(1, 381) = 8.1, p < 0.01$). The effect size for this difference is 0.29 which corresponds to a Cohen's U_3 index of 62%. The average student in ER classrooms scored higher than 62% of the non-ER students on their Rhyming scores (scored 12 percentile points higher than an average non-ER group student).
- The ER group scored significantly higher in Rhyming at all five time periods:
 - time period 1 - $t(369) = 5.1, p < 0.01$
 - time period 2 - $t(382) = 4.1, p < 0.01$
 - time period 3 - $t(382) = 3.0, p < 0.01$
 - time period 4 - $t(365) = 4.5, p < 0.01$
 - time period 5 - $t(352) = 4.4, p = 0.01$

Differences in Variability

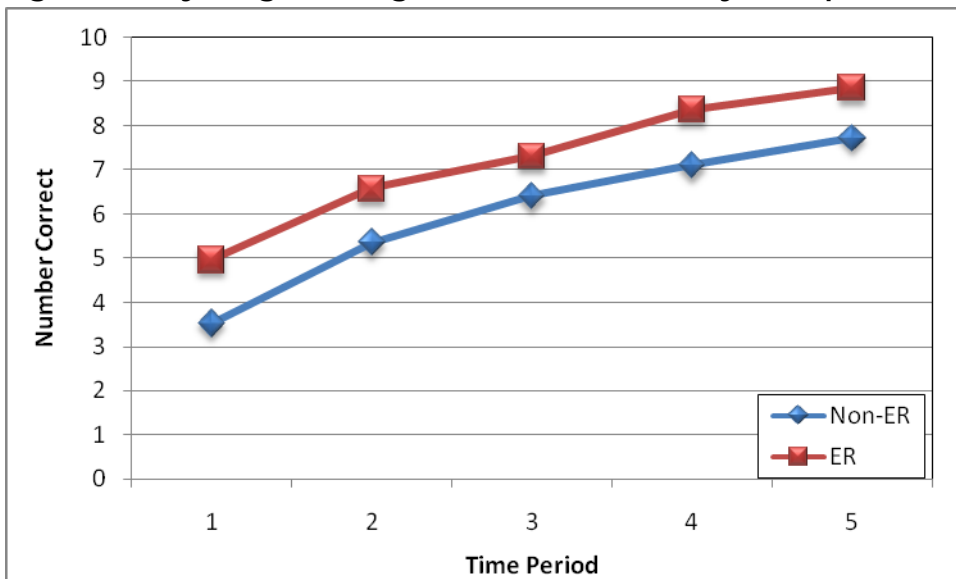
- The variability in scores for ER group was significantly larger than for the non-ER group at time period 1 ($p < .01$) and was statistically smaller at time period 4 ($p < .01$) and time period 5 ($p < .01$). This means there was more spread in student scores for the ER group than for the non-ER group at time period 1 and less spread in student scores for the ER group than for the non-ER group at time periods 4 and 5. See Appendix A for box plots of subtest scores by group and time period.

Table 12. Rhyming: Average Number Correct, Standard Deviation and p Values by Group and Time Period

Group	Time Period					Gain
	1	2	3	4	5	
Average Number Correct						
Non-ER	3.5	5.4	6.4	7.1	7.7	4.2
ER	5.0	6.6	7.3	8.4	8.8	3.9
Difference	1.4	1.2	0.9	1.3	1.1	-0.3
p	0.00	0.00	0.00	0.00	0.00	
Standard Deviation						
Non-ER	2.5	3.0	3.0	3.0	2.8	0.3
ER	3.1	2.8	2.9	2.5	2.1	-0.9
Difference	0.6	-0.2	-0.1	-0.6	-0.7	-1.3
p	0.00	0.14	0.35	0.00	0.00	

*ER N= 194, non-ER N=187

Figure 7. Rhyming: Average Number Correct by Group and Time period



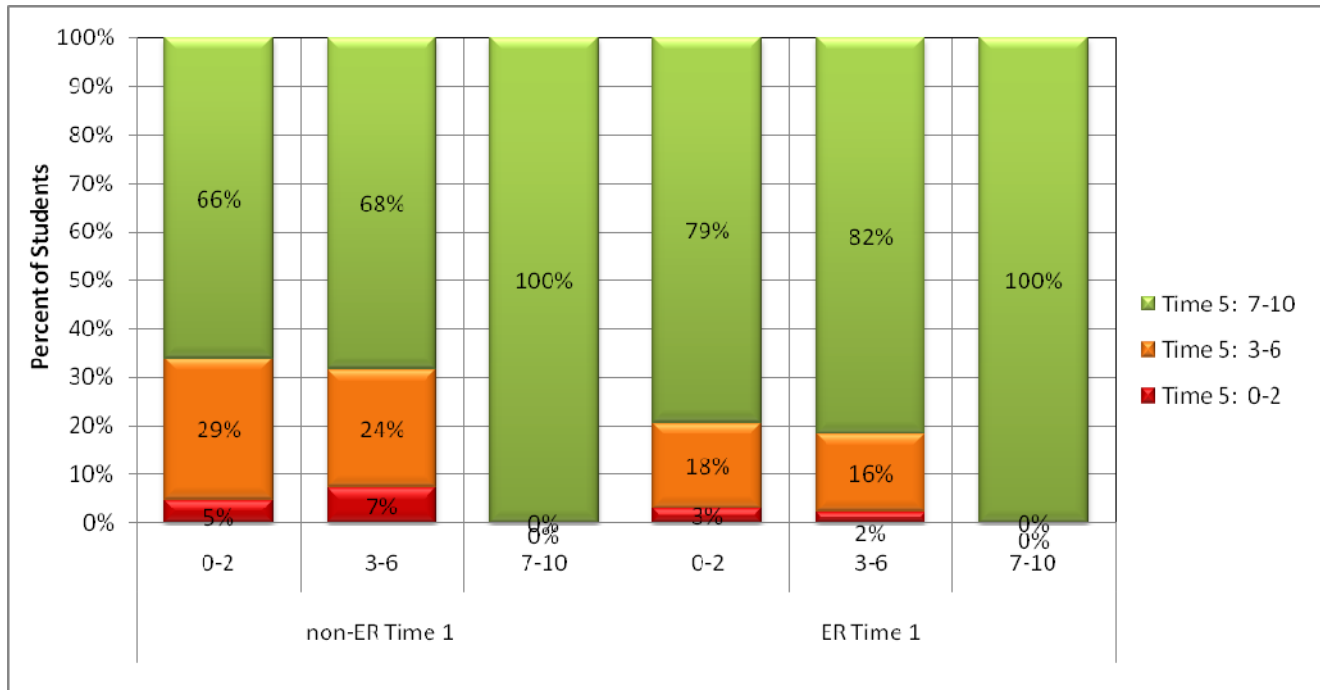
Scores were grouped into three categories based on the number of items the student was able to identify correctly: 0-2 correct, 3-6 correct, and 7-10 correct. This allows one to look at how students scored at time period 1 and see where those students ended up at time period 5 (see Table 13 and Figure 8).

- Of the students who scored below three items correct at time period 1 testing,
- in the ER group 79% moved up to scoring between 7 and 10 items correct at time period 5 testing, compared to 66% in the non-ER group;
 - in the ER group 18% moved up to scoring between 3 and 6 items correct at time period 5 testing, compared to 29% in the non-ER group; and
 - in the ER group 3% remained in the 0-2 items correct group at time period 5 testing, compared to 5% in the non-ER group.
- Of the students who scored between 3 and 6 items correct at time period 1 testing,
- in the ER group 82% moved up to scoring between 7 and 10 items correct at time period 5 testing, compared to 68% in the non-ER group;
 - in the ER group 16% stayed the same scoring between 3 and 6 items correct at time period 5 testing, compared to 24% in the non-ER group; and
 - in the ER group 2% moved down to the 0-2 items correct group at time period 5 testing, compared to 7% in the non-ER group.
- Of the students who scored between 7 and 10 items correct at time period 1 testing,
- in the ER group and the non-ER group 100% stayed the same scoring between 7 and 10 items correct at time period 5 testing.

Table 13. Rhyming Score Categories by Group and Time Period

Group	Time 1 Score Category	Time 5 Score Category						
		Number of Students				Percent of Students		
		0-2	3-6	7-10	N	0-2	3-6	7-10
non-ER	0-2	3	19	43	65	5%	29%	66%
	3-6	7	24	67	98	7%	24%	68%
	7-10	0	0	28	28	0%	0%	100%
ER	0-2	1	6	27	34	3%	18%	79%
	3-6	2	16	80	98	2%	16%	82%
	7-10	0	0	61	61	0%	0%	100%

Figure 8. Rhyming Score Categories by Group and Time Period



Uppercase Letters

Average scores and standard deviations by group are presented in Table 14. Figure 9 shows average scores over time by group. At time period 1, ER students correctly named on average 11.7 letters out of 26. By time period 5 testing, the ER students correctly named on average 23.7 letters. The ER group gained on average 12 points and the non-ER group gained 11.7 points between time period 1 and time period 5 testing.

Differences in Average Scores

- When controlling for student scores at time period 1, the ER group scored significantly higher at time period 5 on their Uppercase Letter scores than did the non-ER group (Adjusted means: 23.6 vs. 22.5, $F(1, 346) = 4.5, p = 0.02$). The effect size for this difference is 0.23 which corresponds to a Cohen's U_3 index of 59%. The average student in ER classrooms scored higher than 59% of the non-ER students on their Uppercase Letters scores (scored nine percentile points higher than an average non-ER group student).
- The ER group scored significantly higher in Uppercase Letters at time period 5 ($t(318) = -2.4, p = 0.01$) and trended toward scoring higher at time period 4 ($t(336) = -1.4, p = 0.08$).

Differences in Variability

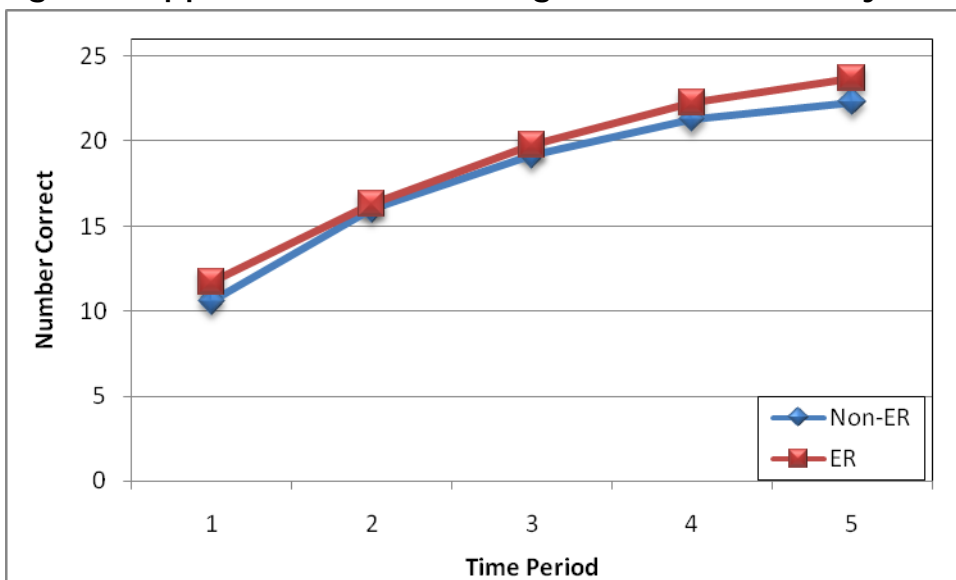
- The variability in scores for ER group and the non-ER group were very similar at time period 1 ($p = .46$). By time period 2, the ER group standard deviation was trending toward being smaller than the non-ER group standard deviation. The standard deviation of the ER group was statistically smaller at time period 3 ($p = .05$), time period 4 ($p < .01$) and time period 5 ($p < .01$). This means there was similar spread in student scores for the ER group and the non-ER group at Time 1 and the spread decreased over time indicating that students were more homogenous in their Uppercase Letter scores over time in the ER group. See Appendix A for box plots of subtest scores by group and time period.

Table 14. Uppercase Letters: Average Number Correct, Standard Deviation and p Values by Group and Time Period

Group	Time Period					Gain
	1	2	3	4	5	
Average Number Correct						
Non-ER	10.6	16.0	19.2	21.3	22.3	11.7
ER	11.7	16.3	19.8	22.3	23.7	12.0
Difference	1.1	0.3	0.6	1.0	1.4	0.3
p	0.14	0.37	0.23	0.08	0.01	
Standard Deviation						
Non-ER	9.6	9.3	8.3	7.2	6.4	-3.2
ER	9.6	8.8	7.3	5.8	4.6	-5.0
Difference	0.0	-0.5	-1.0	-1.4	-1.9	-1.8
p	0.46	0.10	0.05	0.00	0.00	

*ER N= 172, non-ER N=177

Figure 9. Uppercase Letters: Average Number Correct by Group and Time period



Scores were grouped into three categories based on the number of items the student was able to identify correctly: 0-7 correct, 8-16 correct, and 17-26 correct. This allows

one to look at how students scored at time period 1 and see where those students ended up at time period 5 (see Table 15 and Figure 10).

Of the students who scored below eight items correct at time period 1 testing,

- in the ER group 85% moved up to scoring between 17 and 26 items correct at time period 5 testing, compared to 69% in the non-ER group;
- in the ER group 11% moved up to scoring between 8 and 16 items correct at time period 5 testing, compared to 18% in the non-ER group; and
- in the ER group 5% remained in the 0-7 items correct group at time period 5 testing, compared to 12% in the non-ER group.

Of the students who scored between 8 and 16 items correct at time period 1 testing,

- in the ER group and the non-ER group 100% moved up to scoring between 17 and 26 items correct at time period 5 testing.

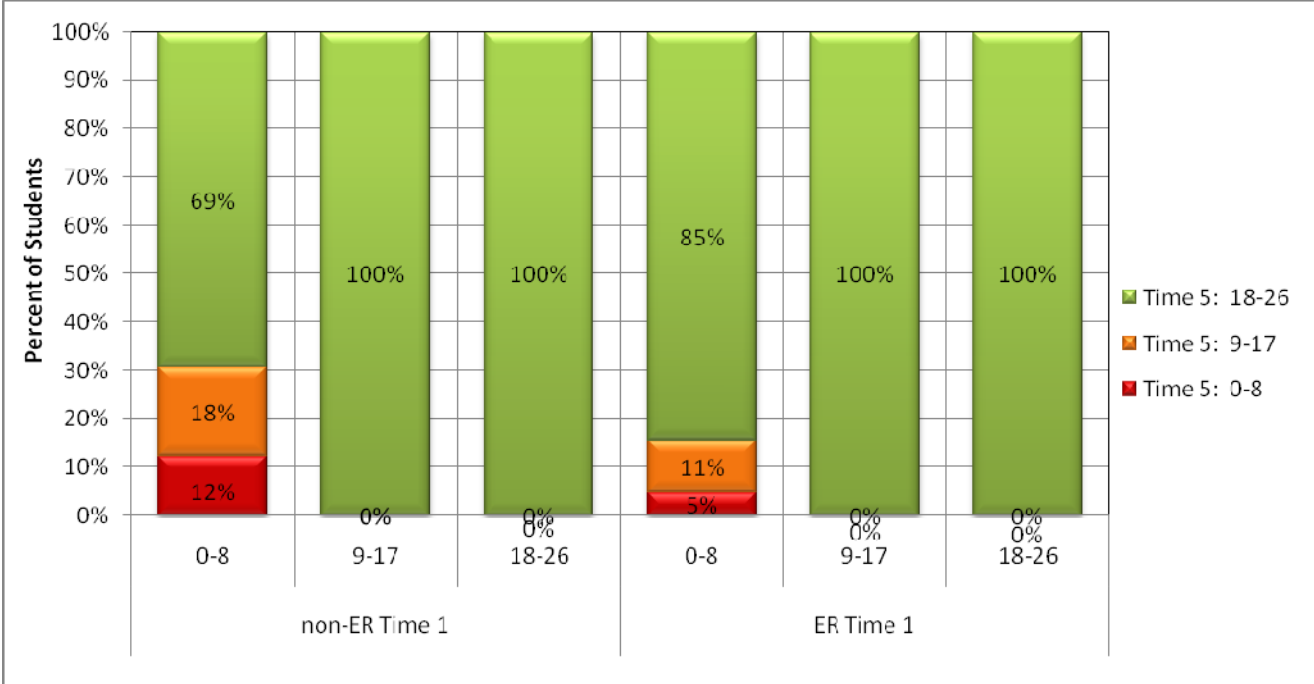
Of the students who scored between 17 and 26 items correct at time period 1 testing,

- in the ER group and the non-ER group 100% stayed the same scoring between 17 and 26 items correct at time period 5 testing.

Table 15. Uppercase Letters Score Categories by Group and Time Period

Group	Time 1 Score Category	Time 5 Score Category						
		Number of Students				Percent of Students		
		0-8	9-17	8-26	N	0-8	9-17	8-26
non-ER	0-8	12	18	68	98	12%	18%	69%
	9-17	0	0	26	26	0%	0%	100%
	8-26	0	0	53	53	0%	0%	100%
ER	0-8	4	9	71	84	5%	11%	85%
	9-17	0	0	31	31	0%	0%	100%
	8-26	0	0	57	57	0%	0%	100%

Figure 10. Uppercase Letters Score Categories by Group and Time Period



Discussion

ER lessons are organized around rhymes and the alphabet, include multiple activities designed to increase student vocabulary, and, to a lesser degree include activities to promote the development of print concepts. Across the five ELQA subtests, the largest effect size was for Rhyming, followed by Uppercase Letters and Picture Naming (Expressive Vocabulary) (see Table 16). The average student in ER classrooms scored higher than 58-62% of non-ER students on these three ELQA assessments. For both the Rhyming and Uppercase Letters assessments, the spread of student scores was more homogenous toward the end of the school year which indicates that there were fewer students who scored low on those assessments in ER Classrooms. This trend of decreased variability in student scores was also evident for Picture Naming; however, it did not reach statistical significance. ER students scored significantly higher at all five time periods on Rhyming. The statistically significant difference at time period 1 indicates that by the time the assessment was administered, ER students had received enough rhyming instruction to move them ahead of students whose teachers were not trained in ER.

Table 16. Adjusted Means, Effect Sizes and *p* values by ELQA Subtest

ELQA Subtest	Adjusted Time 5 Means		Effect Size (Hedges <i>g</i>)	Cohen's U3	<i>p</i>	Means <i>p</i>					Standard Deviation <i>p</i>				
	ER	Non-ER				T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
Rhyming	8.6	7.9	0.29	62%	< .01	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.35	0.00	0.00
Uppercase Letters	23.6	22.5	0.23	59%	0.02	0.14	0.37	0.23	0.08	0.01	0.46	0.10	0.05	0.00	0.00
Picture Naming	23.4	23.1	0.20	58%	0.03	0.38	0.06	0.11	0.04	0.04	0.17	0.02	0.31	0.50	0.09
Receptive Vocabulary	8.8	8.6	0.15	56%	0.08	0.00	0.00	0.00	0.00	0.01	0.49	0.12	0.13	0.28	0.26
Print Concepts	8.0	8.3	-0.15	44%	0.08	0.01	0.42	0.43	0.07	0.42	0.39	0.46	0.02	0.00	0.01

Scores on Receptive Vocabulary were in the expected direction and approached statistical significance when controlling for time period 1 scores. Using unadjusted group means, the ER group scored significantly higher at all five time periods on Receptive Vocabulary. Consider adding more detailed instructions in the curriculum for methods teachers can use to increase student vocabulary. In addition, consider reinforcing the importance of both structured vocabulary instruction and taking advantage of opportunities to introduce and reinforce new vocabulary during ER training.

Scores for Print Concepts were not in the expected direction using the adjusted ANCOVA means. However, using the unadjusted means, ER students scored significantly higher at time 2 and trended toward scoring higher at time 4. By time 5, there were no statistically significant differences between the two groups on Print Concepts. The ELQA Print Concepts assessment consists of ten items used to assess the child's knowledge of various aspects of reading. Students are asked to identify the title on the front cover of a book, identify where one begins to read (first word on page), indicate the direction one reads—that we read words left-to-right, indicate where one

starts reading after completing one line of text—return sweep (start at beginning on next line of text), point to words as they are read (One-to-one match; at least 5 words), understand use of period to end sentence, identify a capital letter, any letter, any word and identify the first letter in a word.

Opportunities for print concepts instruction in the ER curriculum include: the Story Book Caboose lesson activity which is included in all 39 ER lessons, and the Book and Print Awareness activity that occurs in The Choosing Depot but only in the introductory lesson (front and back of book). The Sound The Whistle, Sound Play activity also includes some print concepts related activities. During the ER Professional Development, print concepts identified as important include: “stories and other texts are written left to right, spaces between words matter, and there is a one-to-one correspondence between the words on a page and the words the reader says.” Within each ER lesson, Story Book Caboose lists books related to the lesson rhyme, letter or number. Little is documented in the ER Manual regarding teaching techniques for use when implementing Story Book Caboose other than including “listening and extension” under the activity title. The Book and Print Awareness activity is well-documented in the ER manual and provides teachers direction on how to implement this activity with their students. Consider structuring ER print concept activities to include more specific direction for teachers on how to use a multisensory approach and how to provide repetition for a wider array of print concepts.

The ER Curriculum also includes activities that are not measured by the ELQA (e.g. blending, segmenting, beginning sounds, middle sounds, ending sounds, sound substitution, sound deletion, etc). It would be interesting to collect data to determine if students manifest differences in these early literacy skills.

Overall, the differences between these two demographically similar groups indicate that teachers trained in ER are more effective in increasing student early literacy skills that are central to the ER curriculum (Rhyming and Uppercase Letters) and to a lesser degree vocabulary. Given that data were not collected from teachers regarding their teaching methods, curricula, or behavior one can assume that ER is the reason for these differences. The pattern of findings suggests that ER caused the differences between groups as the largest differences between groups in early literacy skills are the same skills around which ER is organized (rhyming and alphabet). However, this does not establish a causal relationship between ER and student early literacy outcomes. It is also possible that teachers who choose to attend ER training are better teachers than are those who have not attended ER training. Conducting a larger study that includes collecting data on both student early literacy scores over time and on their teachers’ education, teaching methods, curricula and behavior would further establish the relationship between the ER curriculum/training and student scores and teacher behavior.

References

What Works Clearinghouse Procedures and Standards handbook, Version 2 – December 2008. Downloaded from:
http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf

Appendix A. Box Plots

Interpreting box plots

Box plots are a useful type of graph for examining the spread of scores. The top of the box represents the 75th percentile, the bottom of the box represents the 25th percentile, and the line in the middle represents the 50th percentile (the Median). The whiskers (the lines that extend out the top and bottom of the box) represent the highest and lowest values that are not outliers or extreme values. Outliers (values that are between 1.5 and 3 times the interquartile range) are represented by circles beyond the whiskers and extreme values (values that are more than 3 times the interquartile range) are represented by asterisks beyond the whiskers. The numbers beside the outlier and extreme values represent the case numbers for those values.

For example, box charts of student Picture Naming scores by group are presented in Figure 11. Variability at time period 1 testing was very similar between the groups, though the ER group had two extreme scores. At time period 1, both groups had students who maxed out the Picture Naming test and students whose vocabulary score was quite low. The spread of the two groups between the 25th and 75th percentile as well as the median score (50th percentile) were very similar. Over time it is evident that the lower scoring students in the ER group moved upwards into the main body of students. For example outlier student (915) in the ER group who scored less than 5 words correct at time 1, moved into outlier status at times 2 and 3 and into the main distribution of scores by time 4. At time 5 testing, the bottom whisker and the box containing the 25th -75th percentile of students are considerably shorter for the ER group which indicates that fewer students were left behind.

Figure 11. Box Plots: Picture Naming Scores by Group and Time Period

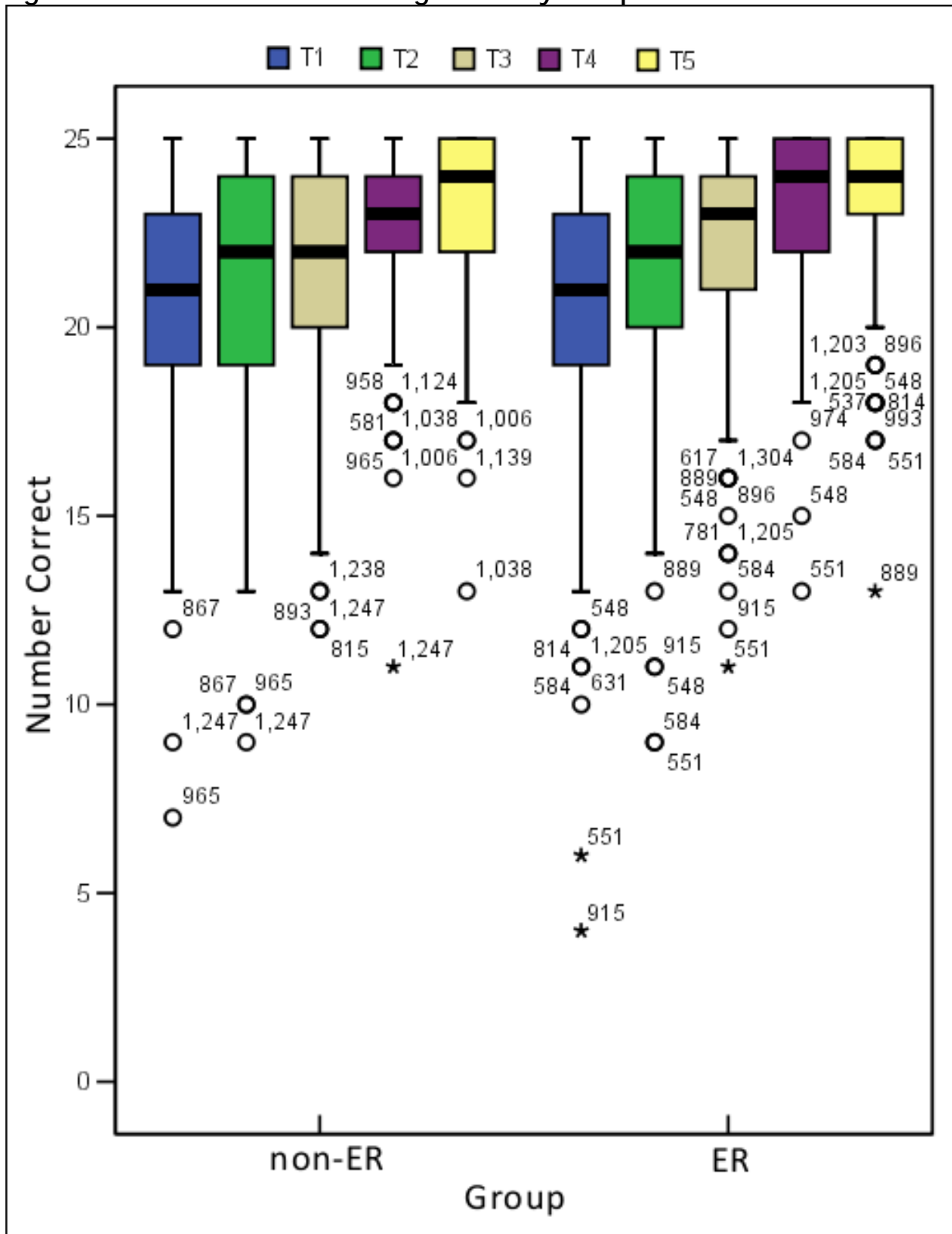


Figure 12. Box Plots: Receptive Vocabulary Scores by Group and Time Period

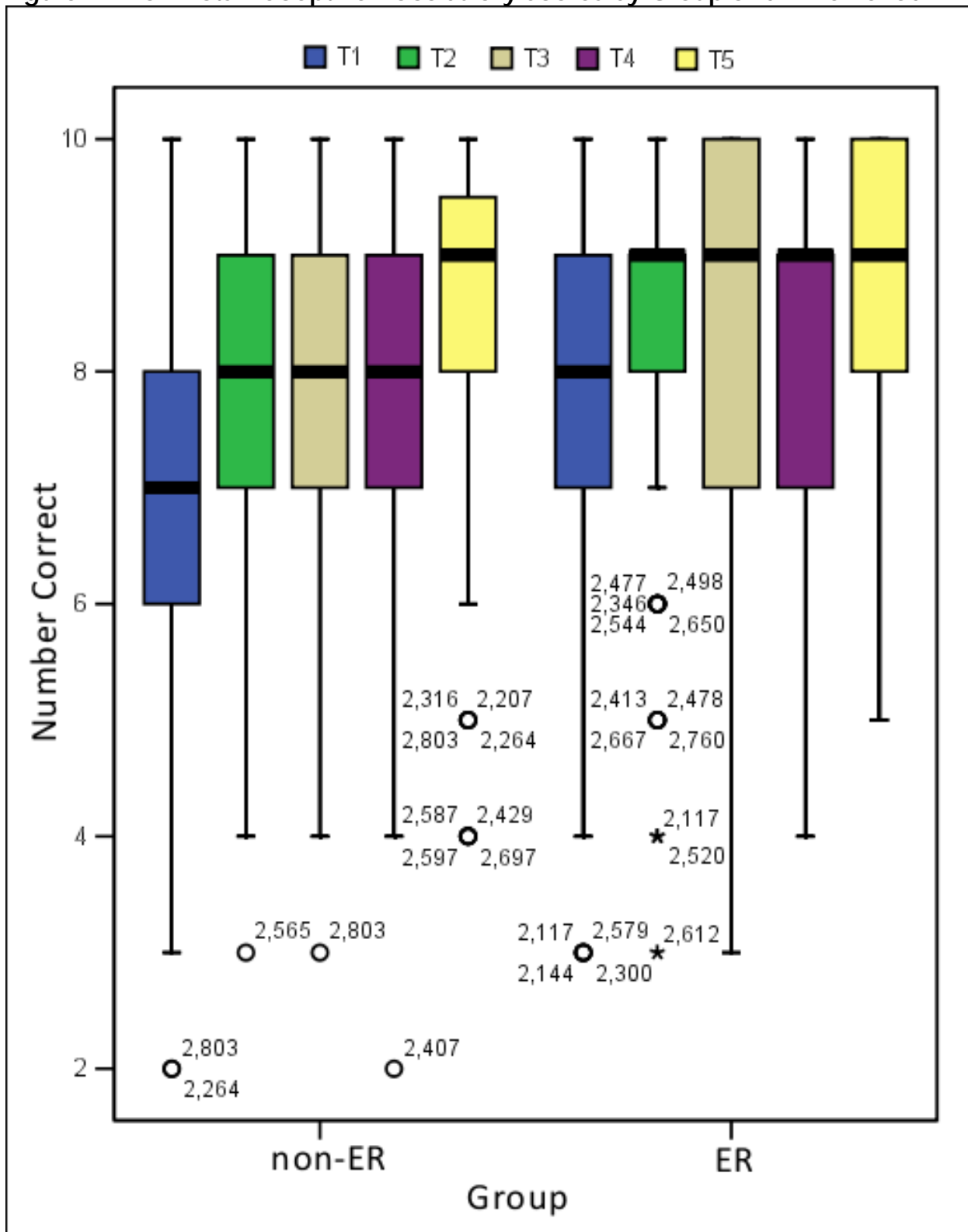


Figure 13. Box Plots: Print Concepts Scores by Group and Time Period

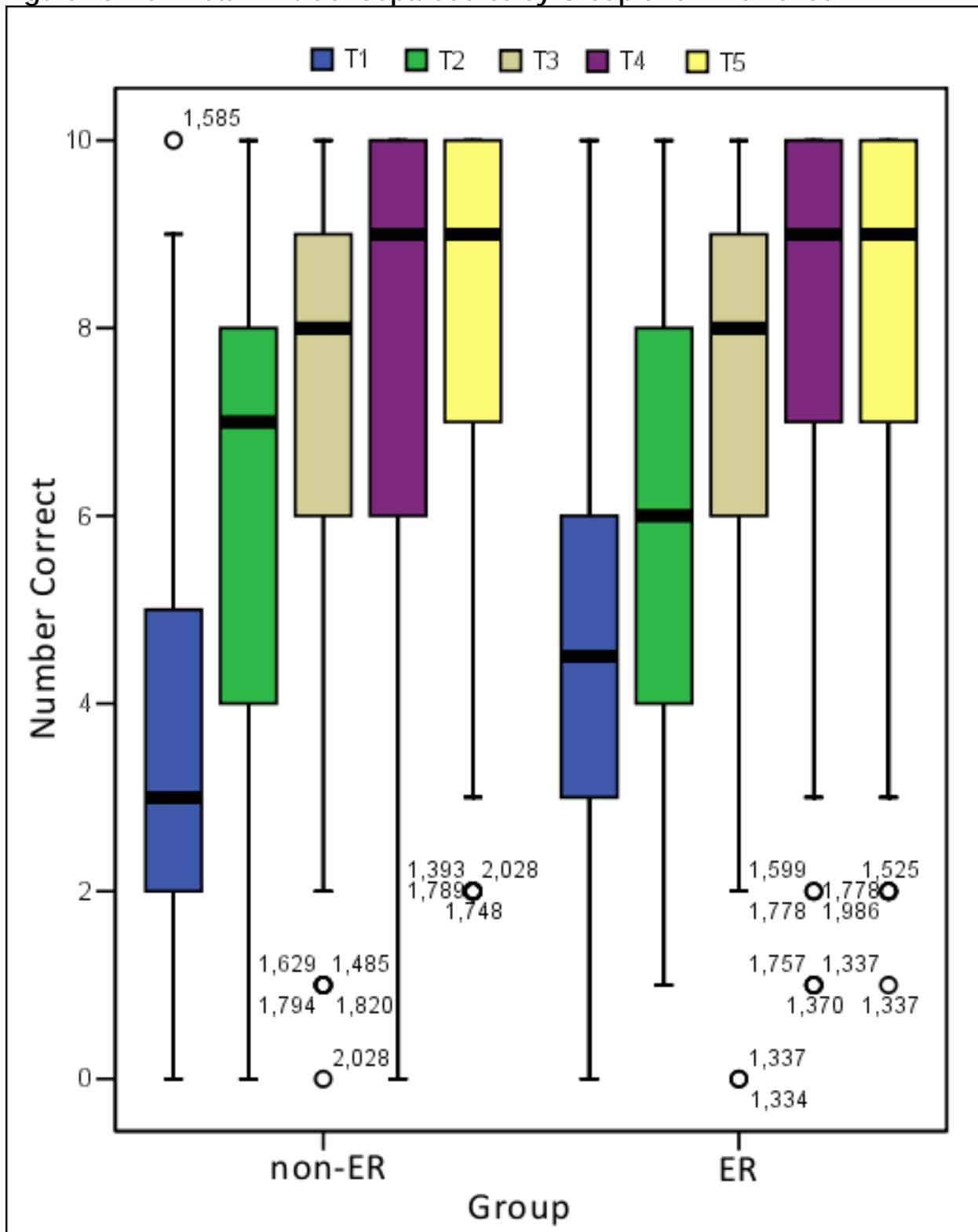


Figure 14. Box Plots: Rhyming Scores by Group and Time Period

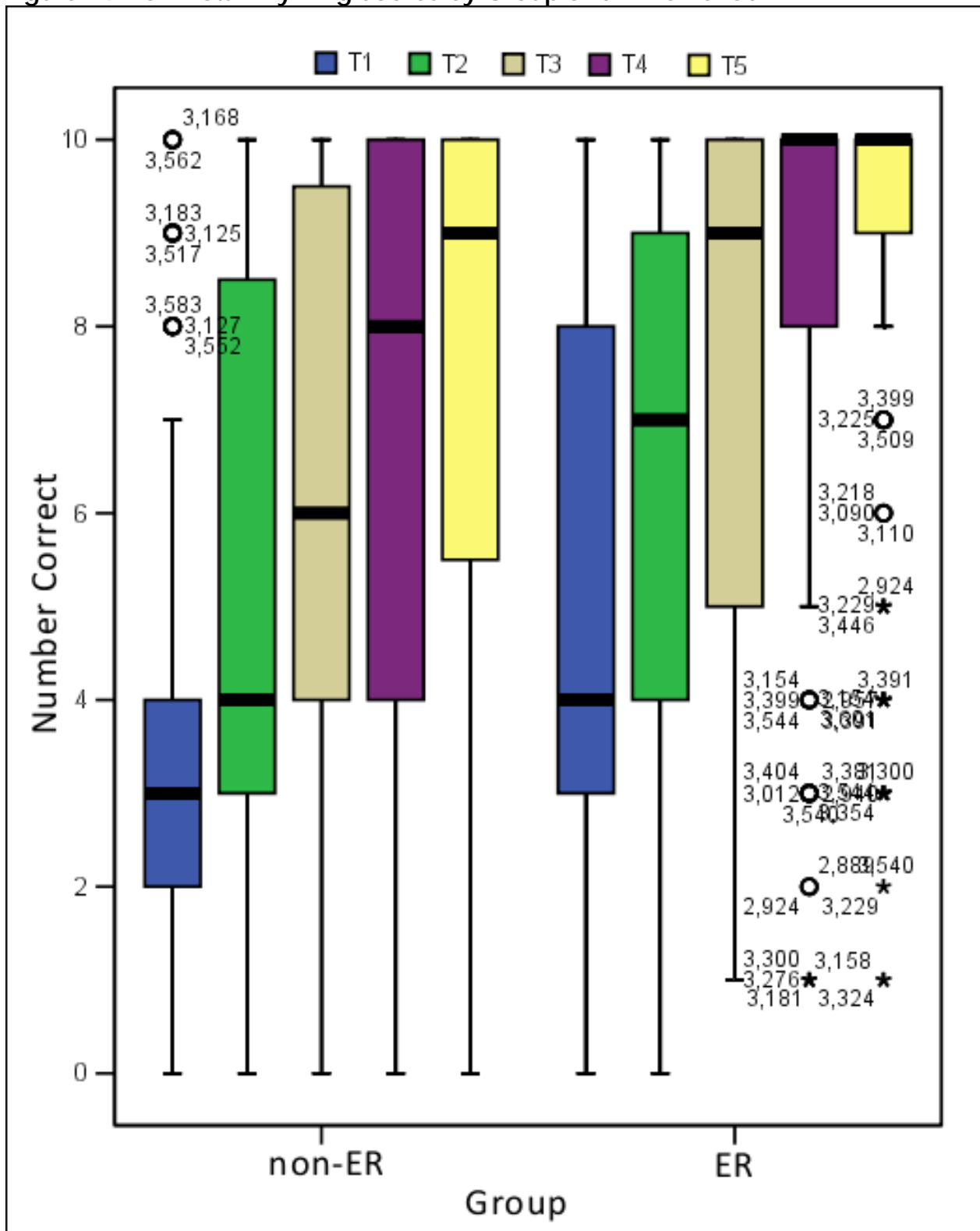


Figure 15. Box Plots: Uppercase Letters Scores by Group and Time Period

