

An Evaluation of the Impact of the Kentucky Reading Project on Teacher and Student Growth 2006-07

Janice F. Almasi, Zsuzsanna Costanzo, Michelle Crout, Susan Thomas Frank, Paige Harrison, Julie Owczarzak, and Jenny Priddy
University of Kentucky

Purpose

The Kentucky Reading Project (KRP) is a fully developed educational intervention, which has been replicated with seven cadres of teachers since its inception in 1998. The purpose of this study was to evaluate the impact of the Kentucky Reading Project, more specifically, this study sought to: (a) evaluate the impact of the Kentucky Reading Project on teacher growth and development, and (b) evaluate the impact of the Kentucky Reading Project on student achievement in reading.

Background

The Kentucky Reading Project (KRP) is an educational intervention developed by the Collaborative Center for Literacy Development (CCLD). Created by Kentucky Senate Bill 186 in 1998, the Collaborative Center for Literacy Development is a consortium of the eight public universities in the Commonwealth of Kentucky and the National Center for Family Literacy. The goal of the KRP is to increase student achievement in literacy by improving knowledge and instructional practices of K-5 teachers through the design and implementation of a comprehensive literacy program in classroom settings. The KRP is a year long, graduate-level reading course, which consists of a two-week summer institute, four follow-up sessions during the year, and at least one coaching visit to each teacher. Faculty members from the eight public universities serve as directors and instructional leaders for the KRP. Reading First coaches from the Kentucky Department of Education work in collaboration with these directors, and training specialists from the National Center for Family Literacy also work with the KRP to provide information, resources, and support regarding family involvement in literacy.

Teachers who participate in the KRP learn how to: (a) apply theory and research to meet the literacy needs of all students; (b) address equity and diversity issues as they pertain to literacy instruction; (c) use phonemic awareness, phonics, and word recognition strategies to develop fluency; (d) identify processes and strategies for teaching reading comprehension; (e) integrate reading and writing; (f) teach reading and writing across the curriculum using narrative and expository texts; (g) use assessment to inform instruction; (h) emphasize family involvement in literacy; (i) provide support for struggling readers; (j) design and manage instruction; and (k) use state curriculum guides. Teams of teachers who participate in the KRP also develop and implement a Literacy Action Plan (LAP) focused on improving student achievement in literacy, teachers' knowledge, and instructional practices in literacy.

Purpose/Research Questions

The goals of this investigation were to determine the impact of the KRP on teacher growth and student achievement in reading. More specifically, research questions include:

1. What is the impact of the KRP on teachers' content knowledge related to teaching reading?
 - (a) What is the impact of the KRP on teachers' content knowledge of comprehension?
 - (b) What is the impact of the KRP on teachers' content knowledge of word analysis?
 - (c) What is the impact of the KRP on teachers' understanding of reading pedagogy?

2. What is the impact of the KRP on students' reading achievement?
 - (a) What is the impact of the KRP on kindergarteners' concepts about print?
 - (b) What is the impact of the KRP on kindergarteners' letter recognition?
 - (c) What is the impact of the KRP on students' (grades 1-5) highest oral reading level?
 - (d) What is the impact of the KRP on students' (grades 1-5) comprehension?
 - (e) What is the impact of the KRP on students' (grades 1-5) strategies for decoding unknown words?

Methods

Research Design

A pretest-posttest one-group design was used to evaluate the impact of the KRP on teacher growth and student reading achievement.

Participants

Two hundred thirty-seven teachers participated in KRP during the 2006-07 academic year. The participants taught a variety of grade levels ranging from preschool through seventh grade. However, the largest percentage of participants taught kindergarten ($n = 40$ or 16.8%), first grade ($n = 34$ or 14.3%), second grade ($n = 30$ or 12.6%), and third grade ($n = 23$ or 9.7%). Most of the participants were classroom teachers ($n = 202$ or 85.2%). Other participants taught Special Education ($n = 15$ or 6.3%) and Title I/Remedial Reading ($n = 10$ or 4.2%). Of these participants, most were in the beginning stages of their teaching careers: 69.2% of the participants had been teaching for five years or less, 49.8% had been teaching for three years or less, and 33.8% had been teaching for two years or less. However, the range of experience varied from one year to 29 years with a mean of 5.6 years of teaching experience.

Teachers were asked to randomly select ten students from their classroom from whom they would gather reading achievement data in the fall of 2006 and the spring of 2007. A total of 2,297 students (517 kindergarteners, 1,780 students in grades 1-6) participated in the evaluation.

Pretest and posttest data were available for 359 students in kindergarten and 416 students in grades one through six.

Measures and Procedures

Several measures were used to determine teacher and student growth and development: (a) the Content Knowledge for Teaching Reading assessment, (b) the Letter Recognition and Concepts About Print subtests of the Observation Survey of Early Literacy Achievement, and (c) the Qualitative Reading Inventory-4.

Content Knowledge for Teaching Reading (CKTR). Teacher growth and development was measured using the Content Knowledge for Teaching Reading (CKTR) assessment (Ball, Phelps, Rowan, & Schilling, 2003). The CKTR was administered prior to the start of the two-week summer institute in the summer of 2006 and at the culmination of the KRP experience in the spring of 2007. Pretest and posttest data were available for 209 of the 237 teacher participants.

The CKTR was designed to measure teacher content knowledge of reading and the effects of this knowledge on instruction. It consists of 48 multiple-choice items representing three domains: (a) Content knowledge of comprehension (13 items), (b) Content knowledge of word analysis (18 items), and (c) Comprehension/knowledge of teaching and content (17 items). Each item is embedded within an authentic classroom scenario. Items pertaining to content knowledge (comprehension and word analysis) require teachers to use their knowledge of reading in the context of teaching situations. Items pertaining to teaching require teachers to use their knowledge of reading pedagogy to select the best teaching actions or moves in each scenario.

Each subscale of the CKTR has adequate reliability (Ball, Phelps, Rowan, & Schilling, 2004). The Comprehension Content Knowledge subscale had the lowest reliability ($\alpha = 0.5984$) while the Word Analysis Content Knowledge subscale had the highest reliability ($\alpha = 0.7291$). The Comprehension/Knowledge of Teaching and Content had a reliability of $\alpha = 0.6678$. Overall, the full scale of the CKTR had high reliability as well ($\alpha = 0.8031$). Thus, the CKTR is a fairly reliable measure of teacher content knowledge for teaching reading.

Observation Survey of Early Literacy Achievement. The Concepts about Print and Letter Identification subtests of Clay's Observation Survey of Early Literacy Achievement (1993) test were individually administered by teachers to determine kindergarten students' ability to recognize letters, directionality, and print conventions. Pretests were administered during the fall and an alternate form of the Concepts about Print (1993) subtest and the Letter Identification subtest were administered as posttests in the spring. Test-retest reliability coefficients for the Concepts About Print subtest ranged from 0.73 – 0.89 (Clay, 1993). When correlated with the Word Reading subtest, the validity of the Concepts About Print subtest was reported as 0.79 (Clay, 1993). Clay (1993) reported the split-half reliability of the Letter Identification subtest as 0.97. Validity was measured by correlating scores on the Letter Identification subtest with scores on the Word Reading subtest and was reported as 0.85.

Qualitative Reading Inventory-4 (QRI-4). Student reading achievement in grades one through six was measured using the Qualitative Reading Inventory-4 (QRI-4), which is an informal reading inventory developed by Leslie and Caldwell (2006). The purpose of the QRI-4, like any other informal reading inventory, is to provide graded word lists and passages to assess a student's oral reading, silent reading, and listening comprehension. Diagnostically the QRI-4 can be used to: (a) estimate reading levels, (b) group students effectively for instruction, (c) choose appropriate textbooks for guided instruction, (d) suggest directions for instructional intervention, (e) compile a profile of a student's reading ability, or (f) suggest student growth over time. Each user chooses different components of the QRI-4 to administer based on his/her desired use. For this study the evaluation team used the QRI-4 to understand how the KRP intervention may have influenced student growth over time. The teachers were to use the data gathered from the QRI-4 to inform and guide their instruction (i.e., estimate reading levels, group students, choose appropriate textbooks for guided instruction, and to inform them about how to help students become better readers). Thus, the QRI-4 is a measure that not only provides performance-based data related to students' growth in reading, but is also of practical use to teachers to help them design instruction to meet the needs of individual students in their classrooms, which is a goal of the KRP intervention.

Each teacher who expected to teach grades 1-7 in the 7th Cadre (2006-07) received a QRI-4 manual (Leslie & Caldwell, 2006) and was trained to administer the QRI during the summer of 2006. Teachers were required to use the QRI-4 to assess ten students who were randomly selected from their classroom. Assessments included measures of oral reading level, word recognition, and comprehension ability (including explicit and implicit comprehension of text). Students read a series of graded passages aloud to determine their highest oral reading instructional level. Analyses of oral reading miscues and comprehension were performed at each student's highest oral reading instructional level to determine the strategies they used during authentic reading to identify words and make sense of text.

Teachers administered the QRI-4 to the same students twice during the school year, once in the fall of 2006 and once in the spring of 2007. Teachers were to use this information to determine progress in students' reading achievement and identify areas of improvement as well as students needing more instruction. More specifically, the QRI-4 was intended to provide educators with a means of identifying a student's reading level, word recognition ability, comprehension ability, and strategy use.

Technical development of the QRI focused on inter-scorer reliability, alternate-form reliability, internal consistency reliability, and criterion-related validity. Leslie and Caldwell (2001) reported high inter-scorer reliability for total miscues ($\alpha = 0.93$) and comprehension questions ($\alpha = 0.98$). The reliability of the QRI to provide accurate instructional reading levels was measured using alternate-form reliability. Livingston's K^2 ranged from 0.71 to 0.84 across all readability levels, which is quite acceptable. Because the QRI-4 has restricted variability, Standard Error of Measurement (SEM) was used to establish the internal consistency reliability. SEM ranged from .13 to .22 across all passages. Concurrent validity was used to establish the criterion-related validity of the QRI to other measures of reading. Students' instructional reading levels in familiar reading material on the QRI-4 were correlated with students' normal curve equivalent scores for total reading on either the California Achievement Test or the Iowa

Test of Basic Skills. Correlations were significant at all grade levels, ranging from .48 in third grade to .86 in first grade. Word recognition scores from the QRI correlated .90 with a combined Word-Identification and Word-Attack scale score from the Woodcock Reading Mastery Test-Revised (WRMT-R). Weighted comprehension scores from the QRI correlated .75 with Passage Comprehension scores of the WRMT-R. Based on this information, the QRI-4 is a reliable and valid measure of oral reading fluency, word recognition, and comprehension ability for students in this evaluation.

Data Analysis

The teacher growth data were analyzed using descriptive statistics, paired samples *t*-tests, and Analysis of Covariance procedures. Kindergarten students' growth was analyzed using descriptive statistics and paired samples *t*-tests.

Student growth in grades one through six was evaluated using slightly different procedures. The QRI-4 is intended to provide insight into a student's independent, instructional, and frustrational oral and silent reading levels with narrative and expository text. However, the measure was only used to determine each student's highest instructional oral reading level, oral reading accuracy at that level, the types of miscues made at that level, and comprehension at that level. Therefore, each individual student's data reflected information about how the child was reading at the level at which instruction should occur. To streamline the data gathering process, and to keep it relevant for teachers, data related to their independent and frustrational levels was not obtained. QRI-4 data were analyzed and interpreted using descriptive statistics, the Wilcoxon Signed Ranks test, and *t*-test procedures. A paired samples *t*-test was used to evaluate hypotheses regarding students' oral reading accuracy from pretest to posttest.

Findings

Teacher Growth and Development

Table 1 displays the means and standard deviations for teachers' scores on each subscale and the full scale of the Content Knowledge for Teaching Reading (CKTR) measure at the time of pretest and posttest. Although 237 teachers participated in the KRP, only 209 completed both the pretest and posttest of the CKTR. Thus, findings are reported for those individuals who completed both pretest and posttest.

Table 1

KRP Teacher Means and Standard Deviations on the Content Knowledge for Teaching Reading Subscales and Full Scale as a Function of Time

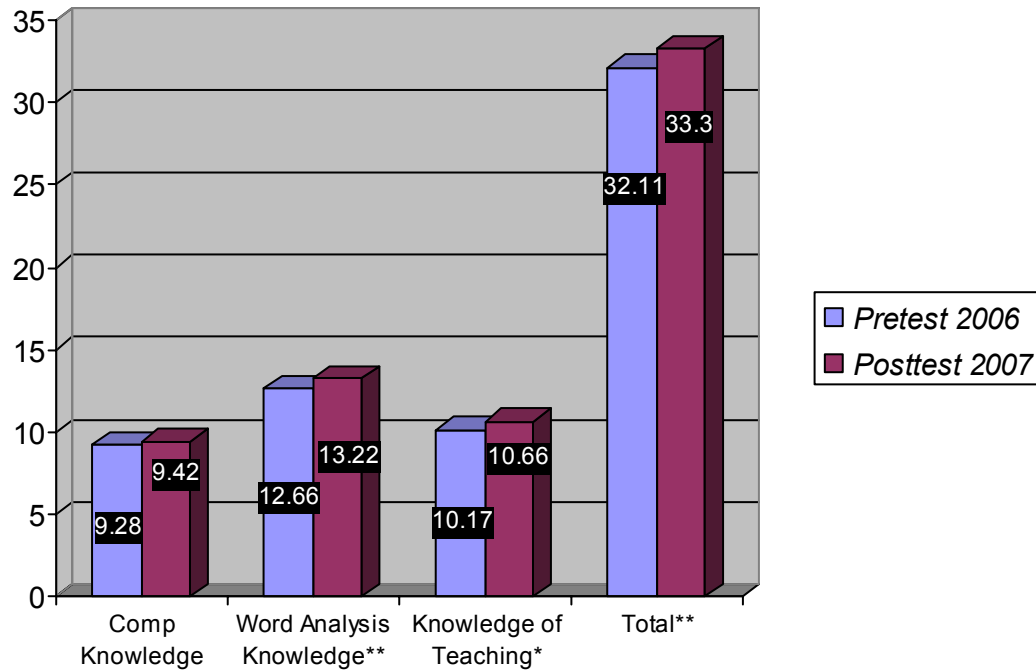
	Content Knowledge for Teaching Reading Subscales							
	Comprehension Content Knowledge (n = 13 questions)		Word Analysis Content Knowledge (n = 18 questions)		Knowledge of Teaching Content (n = 17 questions)		CKTR Total (n = 48 questions)	
	Pretest (n=209)	Posttest (n=209)	Pretest (n=209)	Posttest (n=209)	Pretest (n=209)	Posttest (n=209)	Pretest (n=209)	Posttest (n=209)
Mean	9.28	9.42	12.66	13.22**	10.17	10.66*	32.11	33.30**
Standard Deviation	1.80	1.81	2.16	2.14	2.24	2.30	4.40	4.55

* denotes statistically significant differences ($p \leq .01$)

** denotes statistically significant differences ($p \leq .001$)

Paired samples t -tests were used to compare the pretest and posttest means for each subscale and the full scale of the CKTR. Teachers participating in the KRP demonstrated statistically significant growth from pretest to posttest on all subscales. Although comprehension content knowledge increased slightly, the difference between pretest and posttest was not statistically significant ($t = -0.979$, $df = 208$, $p \leq .33$, two-tailed). Teachers' content knowledge of word analysis, however, did show statistically significant growth ($t = -3.541$, $df = 208$, $p \leq .001$, two-tailed). Similarly, teachers' knowledge of teaching content also showed statistically significant growth ($t = -2.971$, $df = 208$, $p \leq .003$, two-tailed). Teachers' overall scores on the CKTR full scale showed statistically significant growth as well ($t = -4.043$, $df = 208$, $p \leq .001$, two-tailed). Figure 1 visually depicts teacher growth on each subscale and the full scale of the CKTR.

Figure 1. KRP teachers' mean scores on subscales of the Content Knowledge for Teaching Reading measure as a function of time.



* denotes statistically significant differences ($p \leq .01$)

** denotes statistically significant differences ($p \leq .001$)

These findings suggest that teachers who participated in the KRP made significant growth throughout the year related to their content knowledge of word analysis (i.e., knowledge related to decoding and recognizing words) and their knowledge and understanding of teaching (i.e., knowledge of reading pedagogy), which led to statistically significant gains on the full scale of the CKTR.

Student Achievement in Reading

Kindergarten. Kindergarten students' growth and development was measured using the Concepts About Print and Letter Identification subtests of Clay's (1993) Observation Survey of Early Literacy Achievement. Pretest and posttest data were available for 359 of the 517 kindergarteners. Table 2 displays kindergarteners' mean scores on each subtest at the time of the pretest and posttest.

Table 2

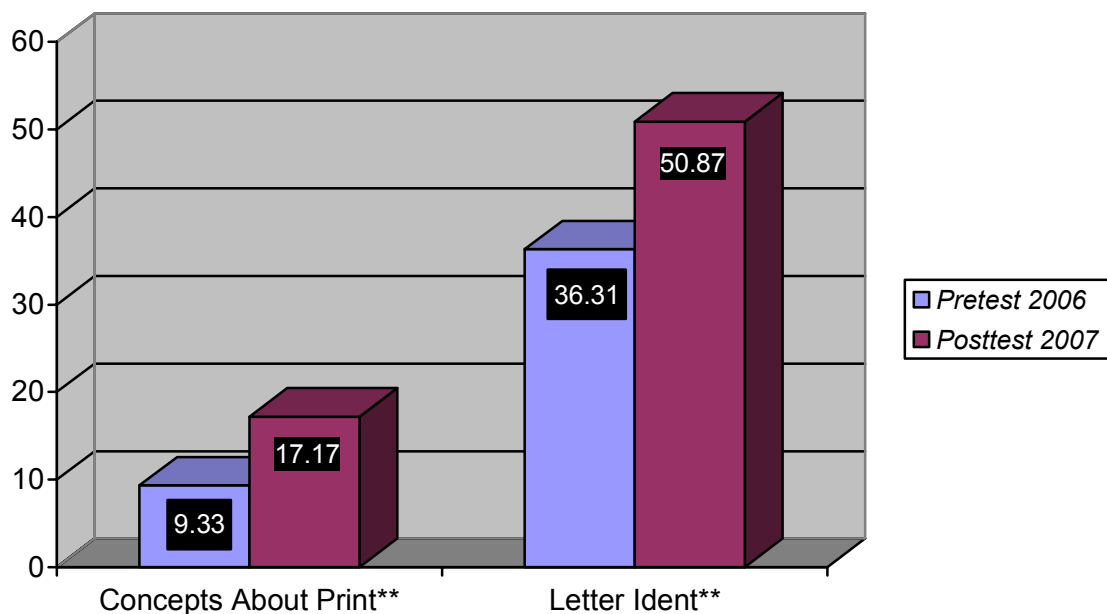
Kindergarteners' Means (and Standard Deviations) for the Concepts About Print and Letter Identification Subtests of the Observation Survey of Early Literacy Achievement as a Function of Time

Concepts About Print		Letter Identification	
Pretest (n = 359)	Posttest (n = 359)	Pretest (n = 359)	Posttest (n = 359)
9.33 (4.85)	17.17** (4.29)	36.31 (17.4)	50.87** (7.56)

** denotes statistically significant differences ($p \leq .001$)

Paired samples *t*-tests were used to determine whether differences in means from pretest to posttest were significant. Statistically significant differences occurred in both the Concepts About Print scores ($t = -32.83$, $df = 358$, $p \leq .001$, two-tailed) and the Letter Identification scores ($t = -18.44$, $df = 358$, $p \leq .001$, two-tailed). Figure 2 provides a visual depiction of these changes over time. Overall these findings suggest that kindergarteners experienced significant growth over time in terms of their ability to recognize letters, determine the direction in which text is read, and the conventions of printed text in English.

Figure 2. Kindergarten students' scores on Concepts About Print and Letter Identification subtests of the Observation Survey of Early Literacy Achievement as a function of time.



** denotes statistically significant differences ($p \leq .001$)

Grades 1-6. Teachers administered reading passages from the QRI-4 to determine the difference between students' highest oral reading level at the time of pretest and their highest oral reading level at the time of posttest. Of the 416 students for whom pretest and posttest data were available, 278 (66.8%) experienced increases in oral reading level, 134 (32.2%) remained at the same level, and 4 (1%) experienced decreases in oral reading level over the year. The mean change in oral reading level for all students was 1.0 (SD = 0.98). The median change in oral reading level was also 1.0 grade levels. The Wilcoxon Signed Ranks Test was used to test whether differences in oral reading levels from pretest to posttest were significant. This procedure indicated that significantly more students experienced positive gains in oral reading level than students who experienced either no gain or negative gains ($Z = -14.56, p \leq .001$).

As students read the texts at their highest oral reading level, teachers recorded the degree to which they read all of the words in the text accurately (oral reading accuracy) and asked students to answer explicit (factual) and implicit (inferential) comprehension questions after they completed their reading. Table 3 displays the means and standard deviations for students' oral reading accuracy (i.e., the percentage of words read correctly at the highest instructional oral reading level), explicit comprehension (i.e., literal understanding of factual details), and implicit comprehension (i.e., ability to make inferences) at pretest and posttest by grade.

Table 3

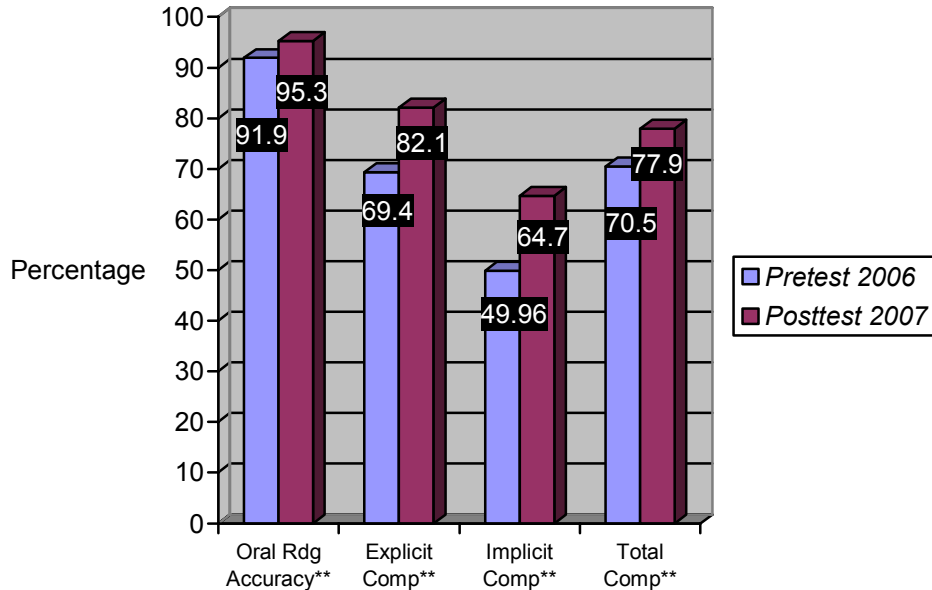
Means and (Standard Deviations) for Oral Reading Accuracy, Explicit Comprehension, and Implicit Comprehension at Students' Highest Instructional Oral Reading Level as a Function of Time

Students	Oral Reading Accuracy		Explicit Comprehension		Implicit Comprehension		Total Comprehension	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
(n = 416)	91.9% (10.29)	95.3%** (5.18)	69.4% (28.83)	82.1%** (18.95)	49.96% (36.40)	64.7%** (32.89)	70.5% (23.03)	77.9%** (15.32)

** denotes statistically significant differences ($p \leq .001$)

Paired samples *t*-tests were used to determine whether differences from pretest to posttest were significant. Students exhibited statistically significant increases from pretest to posttest in terms of oral reading accuracy ($t = -8.571, df = 415, p \leq .001$, two-tailed), explicit comprehension ($t = -9.392, df = 415, p \leq .001$, two-tailed), implicit comprehension ($t = -7.407, df = 414, p \leq .001$, two-tailed), and total comprehension ($t = -6.782, df = 415, p \leq .001$) (see Table 3). Figure 3 depicts these differences visually. Overall these findings suggest students experienced statistically significant gains in their ability to read orally with accuracy and in their explicit and implicit comprehension over time.

Figure 3. Mean scores for oral reading accuracy, explicit comprehension, implicit comprehension, and total comprehension as a function of time.



At the highest instructional oral reading level, teachers were also asked to record the types of miscues students made as they read orally. They recorded six different types of miscues: (a) substitutions (replacing one word for another), (b) mispronunciation (replacing a word with a non-word), (c) insertions (inserting a word where one does not belong), (d) self-corrections (correcting an error on one's own), (e) omissions (failing to read a word in the text), and (f) reversals (transposing two words). Typically, educators hope that over time students decrease the percentage of miscues that are substitutions because this type of miscue tends to be highly correlated with lower levels of comprehension. Alternatively, educators hope that students are able to recognize their oral reading miscues and self-correct them. Thus, educators typically endeavor to decrease the proportion of substitutions while increasing the proportion of self-corrections.

Table 4 displays the means and standard deviations for changes in the percentage of each type of miscue made while reading texts at the highest oral reading level at pretest and posttest.

Table 4
Mean and Standard Deviation for Percentage of Miscues Made at Students' Highest Oral Reading Level as a Function of Time

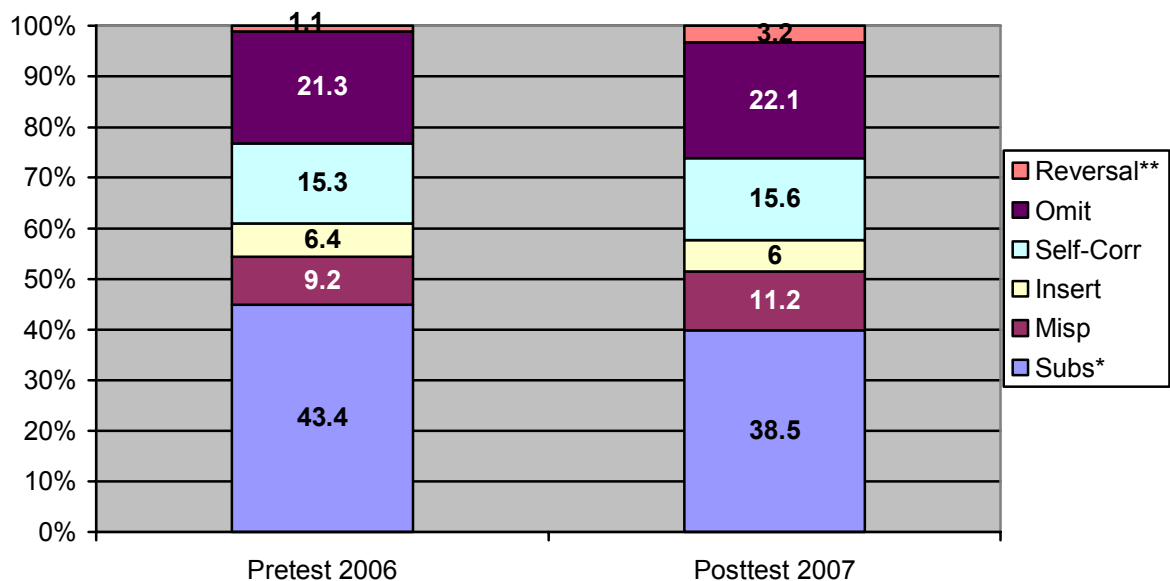
	Types of Miscue											
	Substitution		Mispronunciation		Insertion		Self-correction		Omission		Reversal	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Mean	43.4	38.5*	9.2	11.2	6.4	6.0	15.3	15.6	21.3	22.1	1.1	3.2**
SD	28.9	26.9	16.2	17.0	13.8	12.1	19.3	18.0	25.3	27.5	4.6	11.2

* denotes statistically significant differences ($p \leq .01$)

** denotes statistically significant differences ($p \leq .001$)

Paired samples *t*-tests were used to determine whether differences in percentages of miscues from pretest to posttest were statistically significant. Statistically significant changes occurred in the percentage of substitutions ($t = 2.887$, $df = 415$, $p \leq .004$, two-tailed) and reversals ($t = -3.622$, $df = 415$, $p \leq .001$, two-tailed). Figure 4 represents a visual depiction of these changes over time. These findings suggest that students were making significantly fewer miscues that impede comprehension (i.e., substitutions). However, the number of reversals also increased significantly.

Figure 4. Proportion of each type of miscue at students' highest instructional oral reading level as a function of time.



Cautions

Caution should be used when interpreting the data related to student reading achievement presented above. The 237 participating KRP teachers administered the QRI-4 to 10 students selected randomly from their classes and entered that data on-line. In order to verify the QRI-4 data teachers entered on-line, our evaluation team requested that teachers submit hard copies of the QRI-4 assessments. Data for 1,780 students were submitted by teachers.

In the process of verifying the data to make sure teachers had used the QRI-4 properly and had coded students' oral reading accurately, the evaluation team noted that teachers had made substantive errors in administering the QRI-4, coding students' oral reading miscues, and entering the QRI-4 data on-line. Given this amount of error, the evaluation team decided that the data entered by teachers online was not accurate, and we decided to drop any student data that could not be verified by hand. This decision resulted in a manual re-coding of the data received for all 1,780 students. Ultimately we could only include QRI-4 data in our analyses for which we had both pretest and posttest data that we could verify with the hard copies of the data (n = 416 students). All teacher errors were recorded and the electronic database updated to reflect accurate coding.

Only data from 416 of the 1,780 total students in grades one through six were usable (23.4% of all data). The primary reason why data were not usable was that some portion of the hard copy data was not submitted by teachers for 763 students (42.9% of all data). Without hard copies of the data available for verification we were unable to include much of the data in the present analysis. The remaining data (n = 601 students, 33.7% of all data) were unable to be used due to errors made by teachers during the process of administering the QRI-4 to students. A total of 9 types of teacher administration errors were identified. Two of those teacher administration errors were most common. The most common error was that data for 250 students (14% of all data) were unable to be used because teachers did not recognize that the student was at the frustration level in terms of comprehension and neglected to administer reading passages at a lower level. The second most common error was that data for 182 students (10.2% of all data) were unable to be used because teachers mis-scored students' answers on comprehension questions.

Conclusions

Keeping in mind the cautions discussed above, teachers participating in the KRP during 2006-07 experienced statistically significant increases in word analysis content knowledge and knowledge of teaching. Students also experienced statistically significant growth in terms of oral reading level, oral reading accuracy, explicit comprehension, and implicit comprehension. As well, the processes used by these students changed over time in that they made significantly fewer oral reading errors that impeded comprehension.

Recommendations/Commendations

1. The Kentucky Reading Project is to be commended for continuing to make a positive difference in the lives of teachers and students within the Commonwealth of Kentucky. Within the past year the KRP has implemented several of the recommendations from the 2005-06 evaluation. The most recent evaluation continued to use measures that are able to measure student reading achievement and teacher content knowledge with greater reliability and validity than has ever been possible. As well, the most recent evaluation includes data regarding student growth in kindergarten that is more developmentally appropriate than it has in the past (i.e., the Observation Survey of Early Literacy Achievement). **All personnel associated with the Kentucky Reading Project and CCLD should be commended for working together to implement these changes.**

2. In order to verify that data from student achievement is administered and coded accurately in the future, it is essential that all teachers participating in the KRP submit hard copies of their QRI-4 assessments. Without such information any conclusions regarding student reading achievement must be interpreted with caution. **Thus, it is recommended that all teachers participating in the KRP continue to submit hard copies of all pretest and posttest QRI-4 data they gather.**

3. Teachers continued to have great difficulty administering, coding, and recording student data from the QRI-4 accurately. Learning to administer an informal reading inventory of any type (not just the QRI-4) should be an essential component of every elementary teacher's repertoire. The information gleaned from informal reading inventories such as the QRI-4 enable teachers to: (a) estimate reading levels, (b) group students effectively for instruction, (c) choose appropriate textbooks for guided instruction, (d) suggest directions for instructional intervention, (e) compile a profile of a student's reading ability, and (f) depict student growth over time. Without such information it is unclear how teachers would be able to make instructional decisions to meet students' individual needs.

Thus, it is recommended that the KRP continue to redesign portions of its curriculum to communicate the importance of using classroom-based assessments to inform instruction. The curriculum should also focus on how to administer informal reading inventories such that teachers learn why/how to identify students' independent, instructional, and frustrational reading levels. As well, the curriculum should assist teachers in learning how to use the information gleaned from such assessments to modify instruction to meet students' individual needs.

4. While it is acknowledged that, in reality teachers are required to use many forms of assessment by the state and their district, it is imperative to the KRP evaluation that all teachers use at least one common measure of student achievement. A common measure enables one to evaluate growth in an equivalent manner across the state. The QRI-4 and the Observation Survey of Early Literacy Achievement currently serve that purpose. While state reading assessments could serve a similar purpose they do not provide the

degree of authenticity (i.e., ecological validity and face validity) that these informal measures do. **Thus, it is recommended that the KRP continue to use authentic measures of student achievement that are common across all KRP sites.**

5. **As it has been since 2004, it is again recommended that more systematic evaluation occur for KRP related to teacher and student growth.** The present design (a pretest-posttest one group design) does not permit one to be able to make causal inferences (i.e., that both teacher and student growth are attributable to elements of the KRP curriculum).

While the KRP evaluation in 2006-07 has benefited from the use of more reliable and valid measures of student reading achievement and teacher content knowledge, a more rigorous evaluation would enable causal relationships to be made between KRP and student achievement. An example of a more rigorous evaluation would involve a randomized controlled trial in which data would be gathered using a pretest-posttest control group design (Campbell & Stanley, 1966). Pre- and post-treatment data would be collected for students, parents/guardians, and teacher participants in treatment and control conditions. The effects of the KRP treatment condition would be studied by analyzing pretest and posttest scores of participants in treatment and control conditions to evaluate the impact of the KRP on students' reading achievement, students' motivation for reading, and family involvement in literacy. As well, the impact of the KRP on teacher self-efficacy and teacher adaptability could be evaluated. Thus, these design options would enable student achievement and motivation for reading to be analyzed with respect to changes in teacher behavior.

While the experimental design described above would provide KRP with data that would enable it to generalize its impact, a more rigorous evaluation should also include qualitative investigations of teacher growth and change using observations and interviews. Such rich data sources would provide KRP with a wealth of information about the impact of the program on specific individuals and how their classrooms and teaching were impacted by KRP.

While resources for the evaluation were increased for the 2006-07 evaluation (from \$5,000 to \$35,000), sufficient resources are still not available to conduct the type of rigorous evaluation of KRP that is needed. **Thus, it is also recommended that the annual KRP evaluation budget be increased to at least \$75,000.**

References

- Ball, D. L., Phelps, G., Rowan, B., & Schilling, S. (2003). *Measuring teachers' content knowledge for teaching reading: Elementary reading release items*. Ann Arbor, Michigan: Study of Instructional Improvement.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.

Clay, M. M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.

Leslie, L., & Caldwell, J. (2006). *Qualitative reading inventory – 4*. New York: Longman.