Making Connections

Educator outcomes associated with implementation of Mississippi's K–3 early literacy professional development initiative

Jessica Sidler Folsom
Kevin G. Smith
Florida State University

Kymyona Burk
Nathan Oakley
Mississippi Department of Education

Key findings

In January 2014 the Mississippi Department of Education began providing statewide early literacy professional development to K–3 educators through online modules and face-to-face workshops and providing literacy coaches to target schools (schools identified as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment). This study examined changes in teacher knowledge of early literacy skills and in ratings of quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies in Mississippi between winter 2014 and fall 2015. Among the findings:

- Average educator knowledge increased from the 48th percentile to the 59th percentile on the Teacher Knowledge of Early Literacy Skills survey. Educators who had not yet participated in the professional development program by the end of the study had lower measures of teacher knowledge than those who had completed the program.
- In target schools the average rating of quality of instruction increased from the 31st percentile to the 58th, the average rating of student engagement increased from the 37th percentile to the 53rd, and the average rating of teaching competencies increased from the 30th percentile to the 44th. Teachers who had not yet participated in the professional development program by the end of the study had lower measures of instruction quality, student engagement, and teacher competencies than teachers who had completed the program.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

Institute of Education Sciences
U.S. Department of Education

REL SOUTHEAST
Regional Educational Laboratory
At Florida State University

U.S. Department of Education
Betsy DeVos, *Secretary*

Institute of Education Sciences
Thomas W. Brock, *Commissioner for Education Research*
Delegated the Duties of Director

National Center for Education Evaluation and Regional Assistance
Ricky Takai, *Acting Commissioner*
Elizabeth Eisner, *Acting Associate Commissioner*
Amy Johnson, *Action Editor*
Sandra Garcia, *Project Officer*

REL 2017–270

April 2017

Summary

Substantial research points to the importance of developing strong early literacy skills. However, according to the National Assessment of Educational Progress, between 2007 and 2013, no more than 55 percent of Mississippi grade 4 students were reading at or above the proficiency level that demonstrates solid academic performance for the grade assessed (U.S. Department of Education, 2013). This was of serious concern to educators and policymakers in Mississippi. Thus, in April 2013, Mississippi's Literacy-Based Promotion Act was signed into law with the goal of having every student read at or above grade level by the end of grade 3. In response to the act, in January of 2014 the Mississippi Department of Education began providing early literacy professional development to all K–3 educators using the Language Essentials for Teaching Reading and Spelling program (Moats & Tolman, 2009). Participants received the professional development content across eight modules split into two phases. Each phase included six weeks of online coursework and three days of face-to-face workshops. Typically, educators completed one phase per academic year. Content ranged from learning the foundations of language and reading to teaching comprehension strategies and writing instruction.

At the same time as the professional development, the department provided state literacy coaches to target schools (those most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment). The state literacy coaches spent an average of two to three days per week in each school they served.

In response to a request from the Mississippi Department of Education, the Regional Educational Laboratory (REL) Southeast developed two tools to support the department: the Teacher Knowledge of Early Literacy Skills (TKELS) survey and the Coach's Classroom Observation Tool (CCOT). The Mississippi Department of Education invited all educators of students in kindergarten to grade 3 to complete the TKELS survey at four times between spring 2014 and fall 2015: spring 2014, fall 2014, spring 2015, and fall 2015. State literacy coaches observed classroom instruction in the target schools using the CCOT four times between winter 2014 and spring 2015: winter 2014, spring 2014, fall 2014, and spring 2015. The study team used the results to investigate changes in teacher knowledge and in ratings of quality of early literacy skills instruction, student engagement in early literacy instruction, and teaching competencies as well as the relationship between changes and educators' progress in the Language Essentials for Teaching Reading and Spelling professional development program.

This study presents the results of a systematic investigation of change in educators' knowledge and classroom practices—the targeted areas for change associated with the professional development program that the Mississippi Department of Education used. It does not, however, evaluate the impact of the professional development program. The study design does not allow pre–post changes to be attributed to the program with a high degree of confidence. The four main findings of the study were:

- Between spring 2014 and fall 2015 average teacher knowledge of early literacy skills increased from 49.56 on the Teacher Knowledge of Early Literacy Skills survey to 52.28 (measured in *T*-score points, a standardized score with an average of 50 and standard deviation of 10)—equivalent to answering at least one more item out of 31 correctly. This corresponded to an increase from the 48th percentile to the 59th.

- The increase in teacher knowledge of early literacy skills was associated with progress in the professional development program. Educators who had completed the program scored an average of 2.90 points higher than did educators who had not started it. At the end of the study, educators who had not started the program were in the 54th percentile on the Teacher Knowledge of Early Literacy Skills survey, whereas educators who had completed it were in the 65th percentile.
- Between winter 2014 and spring 2015 in target schools the average ratings of quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies increased. The average rating of quality of instruction increased from the 31st percentile to the 58th percentile, the average rating of student engagement increased from the 37th percentile to the 53rd percentile, and the average rating of teaching competencies increased from the 30th percentile to the 44th percentile.
- The increases in the average ratings of quality of instruction, student engagement, and teaching competencies in target schools were associated with progress in the professional development program. Teachers who had completed the program were rated 0.30 point higher in quality of instruction, 0.22 point higher in student engagement, and 0.41 point higher in teaching competencies than did teachers who had not started it. At the end of the study, teachers who had not started the program were rated in the 42nd percentile for quality of instruction, the 39th percentile for student engagement, and the 38th percentile for teaching competencies, whereas teachers who had completed the program were rated in the 59th percentile for quality of instruction, the 53rd percentile for student engagement, and the 54th percentile for teaching competencies.

The findings suggest that during the period when the Language Essentials for Teaching Reading and Spelling professional development program was implemented, teacher knowledge of early literacy skills, the quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies improved among educators who participated in the program over and above any increases found among educators generally. The findings suggest that progress in the program was associated with improvements in teacher knowledge, quality of instruction, student engagement, and teacher competencies. The study has several implications for alliance members, the Mississippi Department of Education, and the research community:

- The finding that teacher knowledge of early literacy skills can improve is encouraging, because an average of only 54 percent of items were answered correctly across the entire study. Mississippi may consider working on ways to continue improving teacher knowledge of early literacy skills.
- Ultimately, the goal of the Literacy-Based Promotion Act is for all Mississippi students to read at grade level by the end of grade 3. Although improvements in teacher knowledge and instruction are theoretically likely to result in student improvements, future research could consider including measures of K–3 student achievement.
- To determine whether the professional development or coaching efforts were effective, researchers could employ, whenever feasible, a study design that allows for the study of causal relationships (for example, a waitlist randomized controlled trial). This study gives some positive support to the notion that such a study may be worthwhile in that it found changes in a positive direction.

Contents

Why this study?

The ability to read is arguably the most important skill needed for success as an adult. Substantial research points to the importance of developing strong early literacy skills because they are closely linked to reading achievement in the primary school grades and serve as the basis for successful performance in school and beyond (National Early Literacy Panel, 2008; Foorman et al., 2016; National Reading Panel, 2000).

As summarized in a recent National Center for Education Evaluation report (U.S. Department of Education, 2016a), despite research that suggests that teachers' content knowledge and practice do not appear to be strongly and consistently linked to student achievement, one way that many state and local education agencies try to improve the educational attainment of their students is through professional development efforts for teachers. There has been a growing consensus that teachers' knowledge is essential for effective professional development, yet until recently there has been limited evidence to inform the design and delivery of professional development programs. Summarizing across three randomized controlled trials on content-focused professional development (one on reading and two on math), the report drew two primary conclusions:

- Teachers' knowledge and aspects of practice improved.
- Student achievement was not positively affected by the end of the year in which the professional development was implemented.

One way that many state and local education agencies try to improve the educational attainment of their students is through professional development efforts for teachers

However, the report notes that research should look to improve understanding of which aspects of teacher knowledge and practice professional development should focus on, and to improve the impact of professional development on teacher knowledge and practice. This study, while not as rigorous as the three summarized in the report, does address measurement of teacher knowledge and practice, as well as one way that a state has been trying to improve professional development efforts to improve student achievement.

According to the National Assessment of Educational Progress, Mississippi students have historically been below the national average on reading assessments (U.S. Department of Education, 2013). Until 2007, more than half of Mississippi grade 4 students were performing below basic, meaning that they were unable to locate relevant information, make simple inferences, and use their understanding of the text to identify details that support a given interpretation or conclusion or to interpret the meaning of a word as it is used in the text. Between 2007 and 2013 the maximum percentage of grade 4 students reading at or above basic was 55 percent. This was of serious concern to educators and policymakers in Mississippi.

The Mississippi Department of Education believes in a logic model that posits that providing teacher professional development and literacy coaching for target schools will improve teacher knowledge/skill in the near term and improve early childhood literacy in the midterm, which will also lead to better education and life outcomes in the long term. Thus, in April 2013, Mississippi Senate Bill 2347, the Literacy-Based Promotion Act, was signed into law with the goal of having every student read at or above grade level by grade 3. The act requires the Mississippi Department of Education to provide technical assistance and professional development for grade K–3 educators (teachers, coaches, and administrators) as local school districts administer the provisions of the legislation.[1]

According to the Mississippi Department of Education, before the Literacy-Based Promotion Act, the department focused on implementing newly adopted content standards for English language arts and math but did not focus targeted instruction for teachers on specific components of literacy instruction statewide. Although some districts may have offered training related to the components of effective literacy instruction, professional development in early literacy had not been a statewide focus in recent years. In January 2014 the department began providing professional development to all grade K–3 educators through a series of online modules and face-to-face workshops over the course of multiple years using the Language Essentials for Teaching Reading and Spelling program (Moats & Tolman, 2009; box 1).

Box 1. Description of Language Essentials for Teachers of Reading and Spelling professional development program

The Mississippi Department of Education completed an independent, comprehensive selection process to identify a vendor to provide a K–3 literacy professional development training system. The department selected Cambium Learning, which provided the Language Essentials for Teaching Reading and Spelling (Moats & Tolman, 2009) system.

The Language Essentials for Teaching Reading and Spelling program connects research to practice using accessible language and interactive exercises. The content was delivered over a course of eight modules of print and online materials and two face-to-face workshops. Each module builds on previous models; the eight modules (modules 1–7 and module 9; the state determined that the content of module 8 was not applicable to its goals) are provided in two phases.

In the first phase participants are given access to online courses for modules 1–3 for six weeks. Participants come together for three days of in-person training to reinforce the content from modules 1–3 with the addition of module 7. The intended outcomes for modules 1–3 and 7 are to:
- Explore and understand the essential "ingredients" of language.
- Identify the major brain processing systems involved in reading.
- Recognize and reproduce all speech sounds of English, understanding their features.
- Understand the challenges English learner students may face.
- Understand the importance of phonological awareness in reading and spelling instruction.
- Understand English orthography and letter patterns within words.
- Understand phoneme-grapheme correspondences.
- Understand morphology.
- Understand effective, systematic phonics instruction.
- Understand the sequence and substance of concept development in code-based instruction.

In the second phase participants are given access to online courses for modules 4–6 for six weeks. They come together for another round of three days of in-person training to reinforce the content from modules 4–6 with the addition of module 9. The intended outcomes for modules 4–6 and 9 are to:
- Understand varied approaches to vocabulary instruction.
- Understand techniques for fostering word use, knowledge of word relationships, and awareness of word structure.
- Understand the rationale for fluency components in lesson design.
- Learn and practice techniques for speed drills, repeated readings, and simultaneous/alternate reading.
- Learn approaches and strategies for teaching comprehension at the phrase, sentence, paragraph, and passage level.
- Learn questioning techniques and strategies useful before, during, and after readings.
- Understand the component skills that underlie composition.
- Understand the framework for analyzing writing samples.

(continued)

Cambium Learning delivered the training on the program throughout the year, including summers. Training began in January 2014 and was ongoing throughout Mississippi as of June 2016. The first phase of training was generally provided in one academic year and the second in the next academic year. Because enrollment was staggered, educators could enroll in phase 1 at any time but had to complete phase 1 before phase 2. The interval between phase 1 and phase 2 varied. For example, one educator may have taken phase 1 in spring 2014 and then taken phase 2 in summer 2014, whereas another educator may have taken phase 1 in summer 2014 and phase 2 in summer 2015. The original agreement between the Mississippi Department of Education and Cambium Learning was to provide professional development to 10,000 individuals, with targets for each fiscal year (Mississippi Department of Education, 2013; see figure below). Cambium Learning provided and supervised all professional development; actual completion rates were not available to the study team.

Targets for teacher progress in the Language Essentials for Teaching Reading and Spelling professional development program among Mississippi grade K–3 educators over three academic years, 2013/14–2015/16

Percent of educators
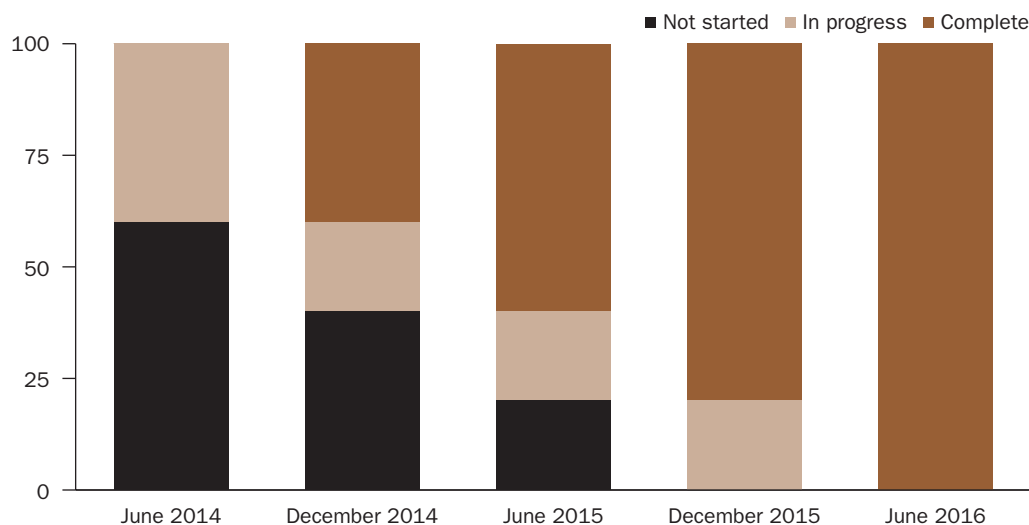
Note: The Language Essentials for Teaching Reading and Spelling professional development program was developed by Moats and Tolman (2009).

Source: Authors' interpretation of Mississippi Department of Education (2013).

Training was offered regionally in groups of no more than 50 participants to all K–3 teachers and administrators, K–8 special education teachers, preservice teachers, and higher education faculty in colleges of education. This report focuses only on K–3 educators (including special education teachers, literacy coaches, and administrators). The training was free and was mandatory only for educators in target schools (schools with the highest percentage of students in the lowest two achievement levels on the statewide literacy assessment). Although the training was not mandatory for other educators, the Mississippi Department of Education and districts highly publicized and encouraged participation. The department also allowed educators to receive continuing education unit credits toward teaching certification as an incentive for participating.

At the same time as the professional development, the department provided state literacy coaches for target schools (schools most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment). The state literacy coaches received the same professional development (though earlier) than the target school teachers. The literacy coaches met monthly for additional professional development and problem solving. Topics of monthly meetings included the research-based role of literacy coaching to making data-based decisions for literacy instruction. The state literacy coaches spent an average of two to three days each week in each school they served.

In light of the Literacy-Based Promotion Act's requirement for professional development, the Regional Educational Laboratory (REL) Southeast Improving Literacy Research Alliance and Improving Schools in Mississippi Research Alliance—both of which have Mississippi Department of Education administrators as members—requested an analysis of teacher knowledge and early literacy skills instruction. Specifically, the alliances were interested in changes in teacher knowledge and classroom instruction and the relationship between changes and educators' progress in the professional development program.

To measure teacher knowledge and classroom instruction, the Mississippi Department of Education requested technical analytic assistance in the development of a teacher knowledge survey and a classroom observation tool as well as analysis of the resulting data

Although there have been several studies of teacher knowledge of early literacy skills, previous tools have typically focused on one or two specific skills (for example, phonological awareness; Binks-Cantrell, Joshi, & Washburn, 2012). Moreover, prior teacher knowledge assessments have typically not been fully tested for validity through substantive item response theory analysis (Binks-Cantrell et al., 2012), despite calls from the field for valid and reliable instruments with adequate estimates of reliability and validity (Brady et al., 2009; Cunningham, Zibulsky, & Callahan, 2009). Therefore, as part of this study, the REL Southeast developed the Teacher Knowledge of Early Literacy Skills (TKELS) survey to include teaching, knowledge, and application questions related to a broad range of early literacy skills (comprehension, writing and grammar, fluency, vocabulary, spelling, phonological and phonemic awareness, and phonics). As part of the development, the current study completed a full psychometric analysis (described in appendix A) of the TKELS survey and developed equated forms (see appendix B)—much like equated forms used for progress monitoring with students.

What the study examined

To measure teacher knowledge and classroom instruction, the Mississippi Department of Education, through the research alliances, asked REL Southeast to provide technical analytic assistance in the development of a teacher knowledge survey and a classroom observation tool. REL Southeast developed the tools, and the department used them for its own data collection purposes. The data were provided to the study team for analysis. Thus, one of the primary contributions of this project is not just the study findings but also the development of the TKELS survey and the Coach's Classroom Observation Tool (CCOT; box 2).

Although this study does not evaluate the impact of the Language Essentials for Teaching Reading and Spelling professional development program using a causal research design, it highlights a systematic investigation of change in educators' knowledge of early literacy skills and classroom practices—the targeted areas for change associated with the program. This study addressed four questions.
- How did average knowledge of early literacy skills among Mississippi K–3 educators change between spring 2014 and fall 2015?

- Was change in average knowledge of early literacy skills associated with educators' progress in the professional development program?
- How did average ratings of the quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies change between winter 2014 and spring 2015 among Mississippi K–3 teachers in target schools?
- Was change in the average ratings of quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies associated with progress in the professional development program among teachers in target schools?

This study does not provide causal conclusions about the efficacy of the professional development program, coaching, or the Literacy-Based Promotion Act. In the absence of a causal study, such as a randomized controlled trial, the effectiveness of the training cannot be determined. However, it is still useful to examine trends in educator outcomes during the time that the program was offered. For example, the findings can help indicate whether the program is potentially promising and should be the focus of additional research and testing. Additionally, this report may be beneficial to staff development practitioners or researchers who might consider using the measurement tools developed for the Mississippi Department of Education and described in this report. Further, this study serves as an example of how states or districts can conduct research on their own staff development initiatives.

This study does not provide causal conclusions about the efficacy of the professional development program, but it does examine trends in educator outcomes during the time that the program was offered

Box 2 summarizes the study data and methods; appendix A provides further details.

Box 2. Data and methods

Data
The Mississippi Department of Education collected all data used in this study. There were two primary sources of data: the Teacher Knowledge of Early Literacy Skills (TKELS; see appendix B) survey and the Coach's Classroom Observation Tool (CCOT; see appendix C). The data gathered through the TKELS survey were intended to be representative of all K–3 educators (general and special education teachers, literacy coaches, and administrators) across the state of Mississippi. The data gathered through the CCOT were intended to be representative of teachers in Mississippi target schools (schools identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment).

Teacher Knowledge of Early Literacy Skills. Regional Educational Laboratory (REL) Southeast assisted the Mississippi Department of Education in developing the TKELS survey, an online survey of teacher knowledge of early literacy skills (see appendix A for details on the development of the survey and appendix B for the survey itself). The survey includes questions related to knowledge, application, and teaching of comprehension, fluency, writing and grammar, vocabulary, spelling, phonological and phonemic awareness, and phonics. The number of items answered correctly out of 31 (raw score) is converted to a T- score, which is a standardized measurement scale with a range of 20–80, mean of 50, and standard deviation of 10. The Mississippi Department of Education solicited responses to the survey four times

(continued)

Box 2. Data and methods *(continued)*

between spring 2014 and fall 2015. See appendix D for descriptive statistics for the surveys. Among the 7,638 unique TKELS respondents, 40 percent took the survey at multiple survey administrations.

Classroom observations. REL Southeast assisted the Mississippi Department of Education in developing the CCOT, a tool for the state literacy coaches to use when observing classroom instruction (see appendix A for details on the development of the tool and appendix C for the tool itself, including operational definitions). Three types of data from the CCOT were used: ratings of quality of early literacy skills instruction (on a scale of 1 = weak, 2 = low average, 3 = high average, 4 = excellent), student engagement during early literacy skills instruction (on a scale of 1 = low, 2 = medium, 3 = high), and teaching competencies (each competency was rated on a scale of 1 = never exhibited, 2 = exhibited less than half the time, 3 = exhibited half the time, 4 = exhibited more than half the time, 5 = exhibited all the time). The ratings of quality of instruction and student engagement are averages across all instructional events, and the rating of teaching competencies is the average of all teaching competencies ratings (see appendix A). Four times between winter 2014 and spring 2015, state literacy coaches randomly selected one teacher, per grade, per school to whom they were assigned, to observe using the CCOT during the 90-minute literacy instruction block. See appendix D for descriptive statistics for the observations. Among the 316 teachers observed with the CCOT, 80 percent were observed more than once.

Methods

Trends in teacher knowledge of early literacy skills and trends in instructional practices were analyzed in two phases. The first phase determined whether teacher knowledge or instructional practices changed over time and, if so, the nature of the change. The second phase examined whether changes were associated with educators' progress in the Language Essentials for Teaching Reading and Spelling professional development program (Moats & Tolman, 2009). Each time data were collected through the TKELS survey or the CCOT, educators reported their progress in the program. For study purposes, professional development progress was broadly defined into three categories:
- *Not started:* The educator had not accessed any of the online modules or face-to-face workshops.
- *In progress:* The educator had started at least one of the online modules (educators could not attend a face-to-face workshop until they had completed the corresponding online modules).
- *Complete:* The educator had completed both sets of online modules and both sets of face-to-face workshops.

For each outcome (teacher knowledge, quality of instruction, student engagement, and teaching competencies), the study team produced an overall growth estimate for educators who had not started the program, a growth difference for educators who were in progress with the program, and a growth difference for educators who had completed the program. The overall growth and the growth difference for each level of progress were added together to determine the amount of growth associated with level of progress in the program.

Where applicable, corresponding percentiles are provided to assist in interpreting the magnitude of change from the beginning to end of the study and the magnitude of the difference between educators who had not started the professional development program and those who had completed it. A percentile is the percentage of scores at or below a given value.

What the study found

This section details the results of the analyses corresponding to the four research questions outlined above.

Mississippi K–3 educators showed increased knowledge of early literacy skills between spring 2014 and fall 2015

Across the four administration windows, the mean TKELS score was 51.24 points, which is equivalent to answering 17–18 items correctly out of 31 and corresponds to the 55th percentile.

Not accounting for progress in the professional development, in spring 2014 the average TKELS score across all educators in the state was 49.55, which is equivalent to answering 16–17 items correctly out of 31 and corresponds to the 48th percentile (figure 1; see also table A7 in appendix A). There was a slight decrease (0.23 point) between spring 2014 and fall 2014, but then the average score began to increase. By fall 2015 it was 52.28, which is equivalent to answering 18–19 items answered correctly and corresponds to the 59th percentile. This was an increase of 2.73 points over the course of the study, which is equivalent to answering approximately one more question correctly. This small but statistically significant increase indicates positive change in teacher knowledge of early literacy skills among Mississippi K–3 educators.

Change in teacher knowledge of early literacy skills was associated with educators' progress in the professional development program

Change in teacher knowledge of early literacy skills was associated with educators' progress in the Language Essentials for Teaching Reading and Spelling professional development program

The percentage of respondents to the TKELS survey who had not started the Language Essentials for Teaching Reading and Spelling professional development program decreased over time, and the percentage who were in progress or who had completed the program

Figure 1. The average score on the Teacher Knowledge of Early Literacy Skills survey among Mississippi grade K–3 educators increased between spring 2014 and fall 2015

Teacher Knowledge of Early Literacy Skills score
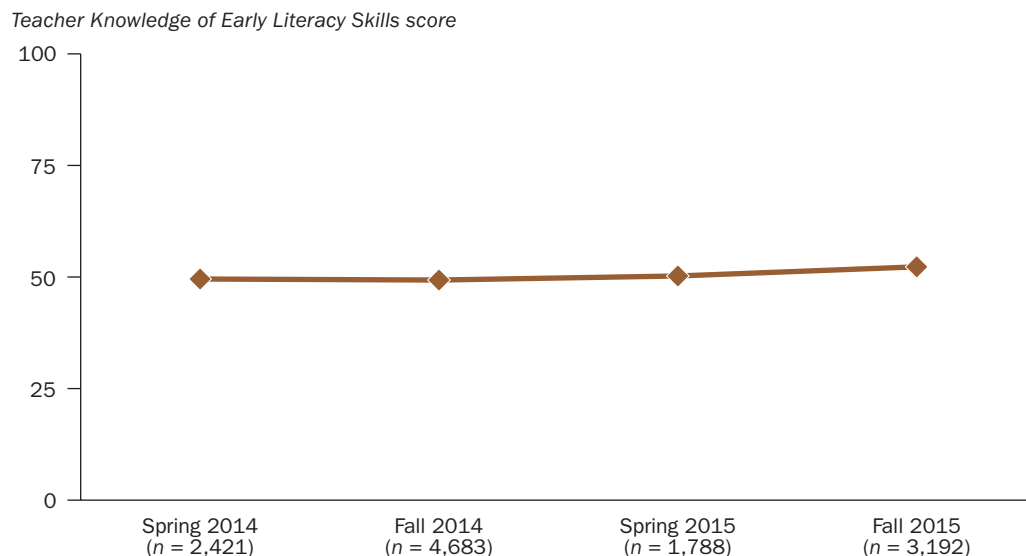
Source: Authors' analysis of data from Mississippi Department of Education (2015b).

increased over time (figure 2). In spring 2014, 90 percent of respondents had not started the program, and less than 1 percent had completed it. By fall 2015, 51 percent of respondents had not started the program, and 29 percent had completed it.

Completing more of the professional development program was associated with increases in teacher knowledge. The overall TKELS score growth and the overall differences in scores between educators who were in progress and educators who had completed the program were statistically significant. After typical growth in teacher knowledge among individuals who had not started the program at each administration window was controlled for, educators who were in progress scored an average of 1.42 points higher on the TKELS survey than did educators who had not started. This is roughly equivalent to answering one more item out of 31 correctly. Similarly, educators who completed the program scored an average of 2.90 points higher than did educators who had not started. This is roughly equivalent to answering two more items correctly. By fall 2015, on average, educators who had not started the program were in the 54th percentile, educators who were in progress were in the 60th percentile, and educators who had completed the program were in the 65th percentile.

Completing more of the professional development program was associated with increases in teacher knowledge

Among Mississippi target schools, K–3 teachers showed gains in the ratings of quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies between winter 2014 and spring 2015

This section describes the results of the analyses to determine whether there were changes in the ratings of quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies.

Figure 2. The percentage of Mississippi grade K-3 educators who were in progress or had completed the Language Essentials for Teaching Reading and Spelling professional development program increased among survey respondents between spring 2014 and fall 2015
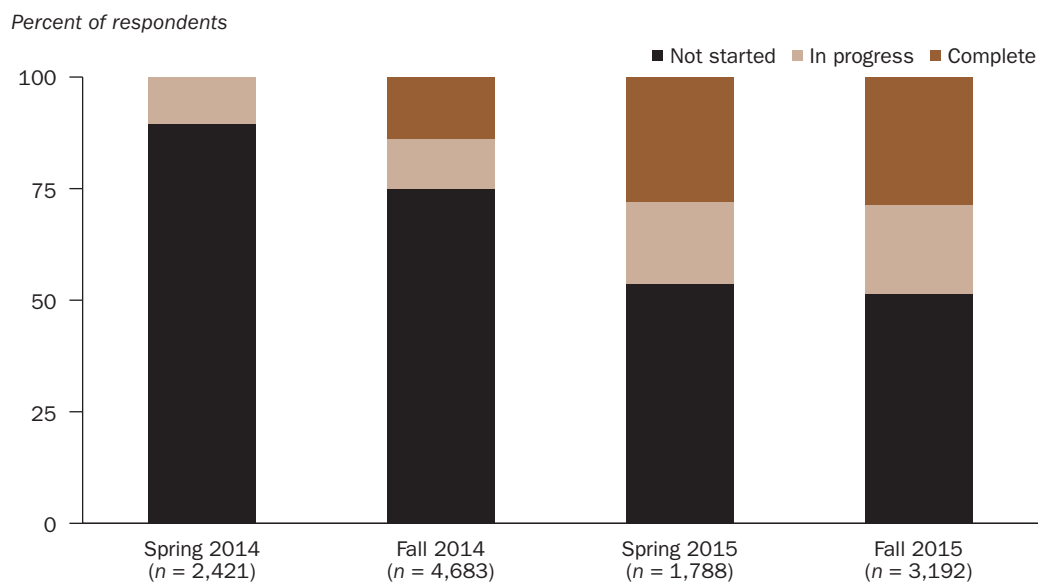
Percent of respondents

Note: The Language Essentials for Teaching Reading and Spelling professional development program was developed by Moats and Tolman (2009).

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

Ratings of quality of early literacy skills instruction increased. The percentage of teachers rated high average to excellent in quality of literacy instruction increased between winter 2014 and spring 2015, while the percentage of teachers rated weak or low average decreased (figure 3).

In winter 2014 the average teacher rating for quality of early literacy skills instruction was 2.18, which corresponds to the 31st percentile (figure 4). The average rating increased steadily at approximately 0.15 point per observation period. By spring 2015 it had increased 0.45 point to 2.63, which corresponds to the 58th percentile.
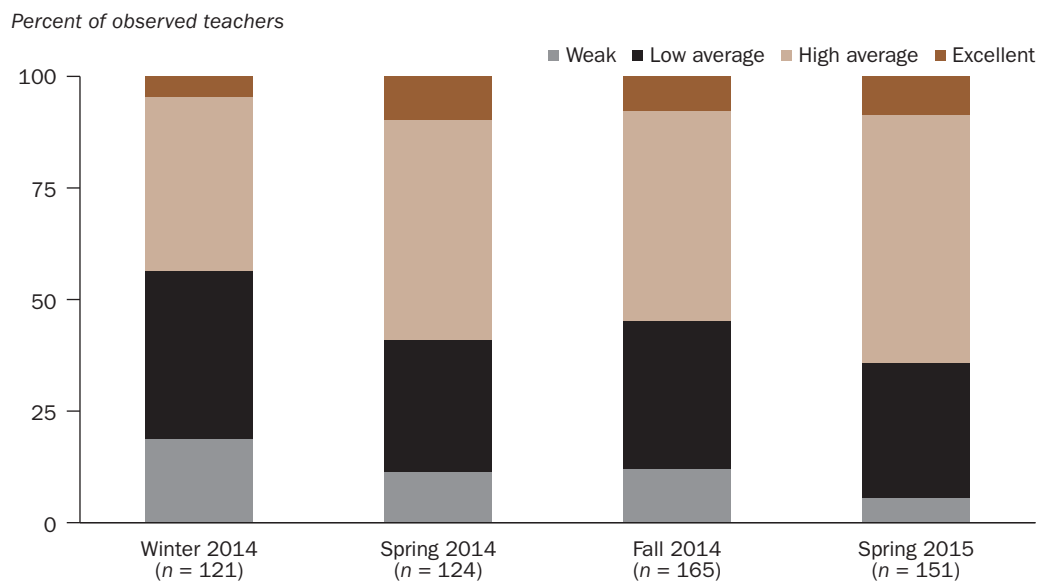
Ratings of student engagement during early literacy skills instruction increased. The percentage of teachers rated high in student engagement increased between winter 2014 and spring 2015, while the percentage rated low decreased (figure 5).

In winter 2014 the average student engagement rating was 2.15, which corresponds to the 37th percentile (figure 6). The average rating increased steadily at approximately 0.09 point per observation window. By spring 2015 it had increased 0.27 point to 2.42, which corresponds to the 53rd percentile.

Ratings of teaching competencies increased. In winter 2014 the average teaching competency rating was 3.20, which corresponds to the 30th percentile (figure 7). The average rating increased steadily at approximately 0.23 point per observation period. By spring 2015 it had increased 0.69 point to 3.89, which corresponds to the 44th percentile.

Ratings of quality of early literacy skills instruction, of student engagement during early literacy skills instruction, and of teaching competencies increased

---

Figure 3. The percentage of Mississippi grade K–3 teachers in target schools who were rated high average to excellent in quality of early literacy skills instruction increased between winter 2014 and spring 2015

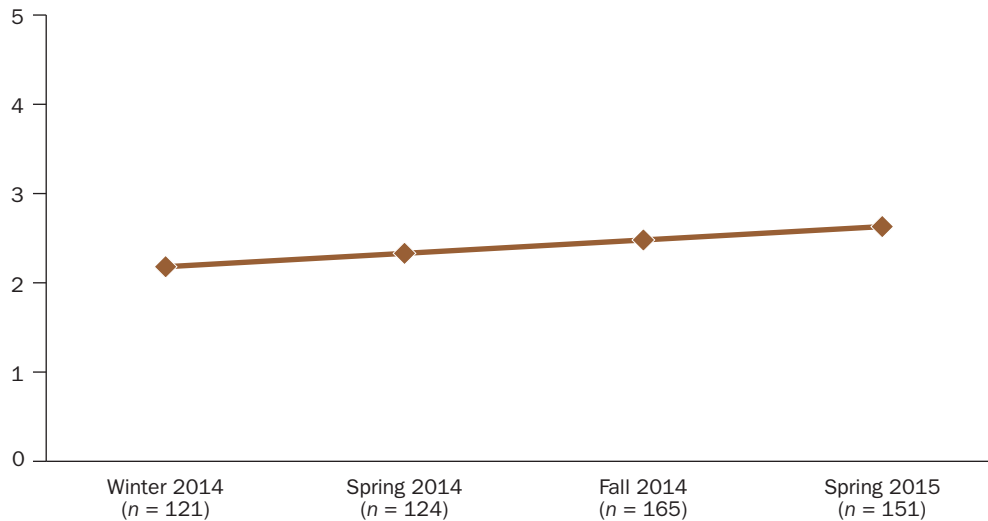*Percent of observed teachers*

Note: Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. At the conclusion of each observation the literacy coach assigned an overall quality of early literacy skills instruction rating of weak, low average, high average, or excellent.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Figure 4. The average rating of quality of early literacy skills instruction among Mississippi grade K–3 teachers in target schools increased between winter 2014 and spring 2015

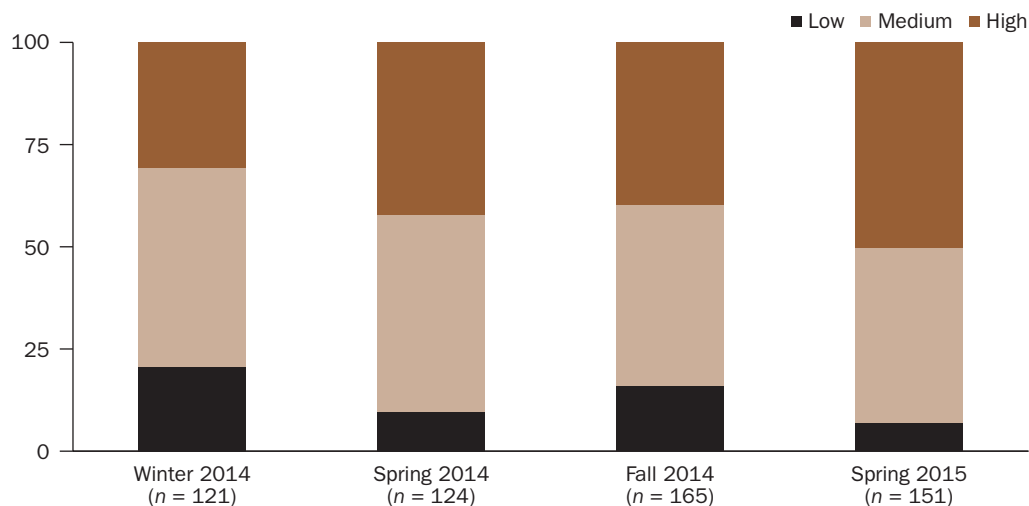*Average rating of quality of early literacy skills instruction*

Note: Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. During the observation, the literacy coach rated the quality of each instructional event as weak, low average, high average, or excellent. The observation tool automatically calculated the average quality rating for each instructional category across all instructional events. At the conclusion of each observation the literacy coach assigned an overall quality of early literacy skills instruction rating of weak, low average, high average, or excellent. For the growth models the study team calculated a single mean quality score as the mean of the overall quality rating and the average quality rating across all instructional categories.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Figure 5. The percentage of Mississippi grade K–3 teachers in target schools who were rated high in student engagement during early literacy skills instruction increased between winter 2014 and spring 2015
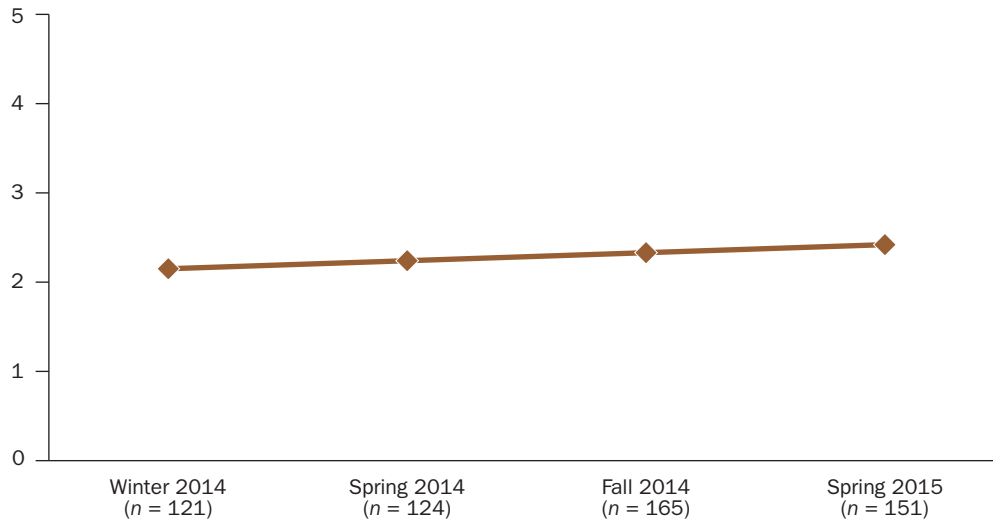
Percent of observed teachers

Note: Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. At the conclusion of each observation the literacy coach assigned an overall student engagement during early literacy skills instruction rating of low, medium, or high.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Figure 6. The average rating of student engagement during early literacy skills instruction among Mississippi grade K–3 teachers in target schools increased between winter 2014 and spring 2015

Average rating of student engagement during early literacy skills instruction

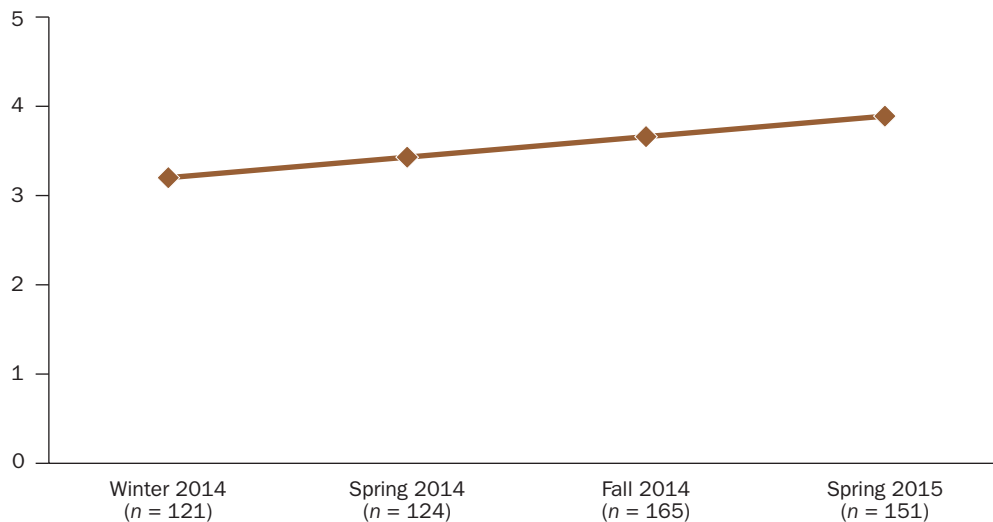| Winter 2014 (n = 121) | Spring 2014 (n = 124) | Fall 2014 (n = 165) | Spring 2015 (n = 151) |

Note: Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. During the observation the literacy coach rated student engagement for each instructional event as low, medium, or high. The observation tool automatically calculated the average student engagement rating for each instructional category across all instructional events. At the conclusion of each observation the literacy coach assigned an overall student engagement rating of low, medium, or high. For the growth models the study team calculated a single mean student engagement score as the mean of the overall student engagement rating and the average student engagement rating across all instructional categories.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Figure 7. The average rating of teaching competencies among Mississippi grade K–3 teachers in target schools increased between winter 2014 and spring 2015

Average rating of teaching competencies

| Winter 2014 (n = 121) | Spring 2014 (n = 124) | Fall 2014 (n = 165) | Spring 2015 (n = 151) |

Note: Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. The teaching competencies rating is an average of 30 items related to planning, management, instruction, monitoring of student learning, and personal characteristics. For each item the literacy coach rated the extent to which the teacher exhibited the competency, ranging from never to all the time.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Among Mississippi target schools, changes in average ratings of quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies were associated with teachers' progress in the Language Essentials for Teaching Reading and Spelling professional development program
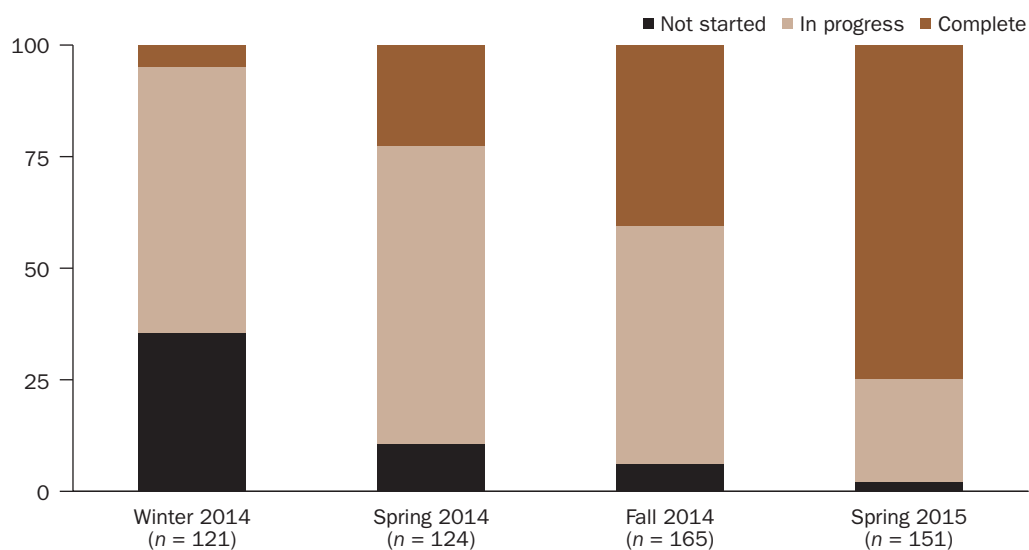
The percentage of observed teachers in target schools who had not started the professional development decreased over time, and the percentage who were in progress or who had completed the Language Essentials for Teaching Reading and Spelling professional development program increased over time (figure 8). In winter 2014, 36 percent of observed teachers had not started the program, and 6 percent had completed it. By spring 2015, 2 percent of observed teachers had not started the program, and 75 percent had completed it. These completion rates are different from those for the teacher knowledge analyses because the observations were conducted only in target schools. The professional development program was mandatory to educators in target schools and was offered to them before educators in other schools.

Change in the average rating of quality of early literacy skills instruction was associated with target school teachers' progress in the professional development program. Completing more of the professional development program was associated with increases in ratings of quality of instruction (figure 9). Across all observation windows, 62 percent of teachers who had completed the program were rated high average or excellent, compared with 24 percent of teachers who had not started the program.

Completing more of the professional development program was associated with increases in ratings of quality of instruction

Figure 8. The percentage of teachers who were in progress or had completed the Language Essentials for Teaching Reading and Spelling professional development program increased among observed Mississippi grade K–3 teachers in target schools between winter 2014 and spring 2015

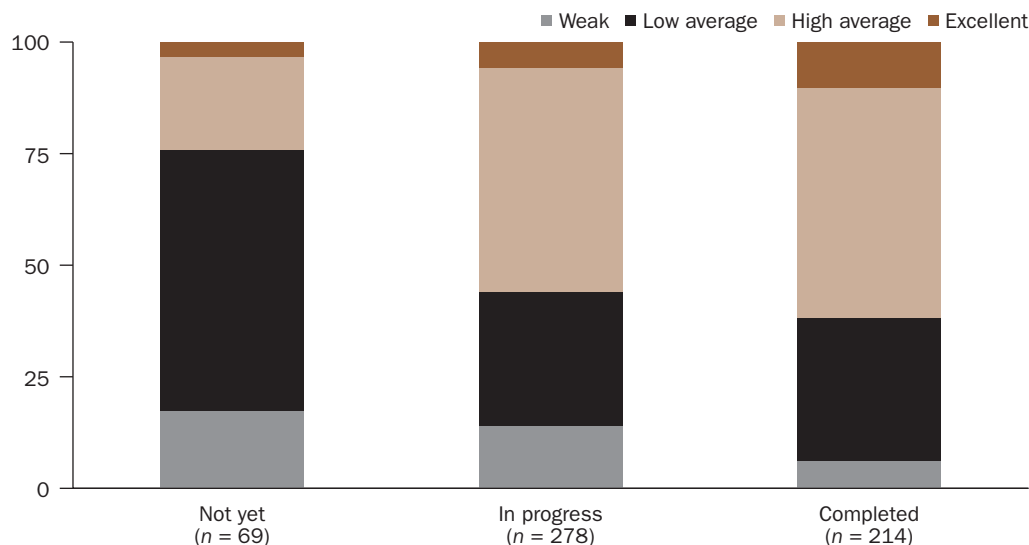*Percent of observed teachers*

Note: The Language Essentials for Teaching Reading and Spelling professional development program was developed by Moats and Tolman (2009).

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Figure 9. The percentage of Mississippi grade K–3 teachers in target schools who were rated high average to excellent in quality of early literacy skills instruction was higher among teachers who were in progress or who had completed the Language Essentials for Teaching Reading and Spelling professional development program than among teachers who had not started it, winter 2014–spring 2015

Percent of observed teachers

Note: The Language Essentials for Teaching Reading and Spelling professional development program was developed by Moats and Tolman (2009). Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. At the conclusion of each observation the literacy coach assigned an overall quality of early literacy skills instruction rating of weak, low average, high average, or excellent.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

After typical growth among teachers who had not started the professional development program at each observation was controlled for, teachers who had completed the program were rated an average of 0.30 point higher in quality of early literacy skills instruction than were teachers who had not started it. The overall growth in ratings for all teachers was statistically significant. The difference in ratings between teachers who had completed the program and those who had not started it was statistically significant. The difference in ratings between teachers who had not started the program and those who were in progress was not statistically significant. By spring 2015, on average, teachers who had not started the program were in the 42nd percentile, teachers who were in progress were in the 50th percentile, and teachers who had completed the program were in the 59th percentile.
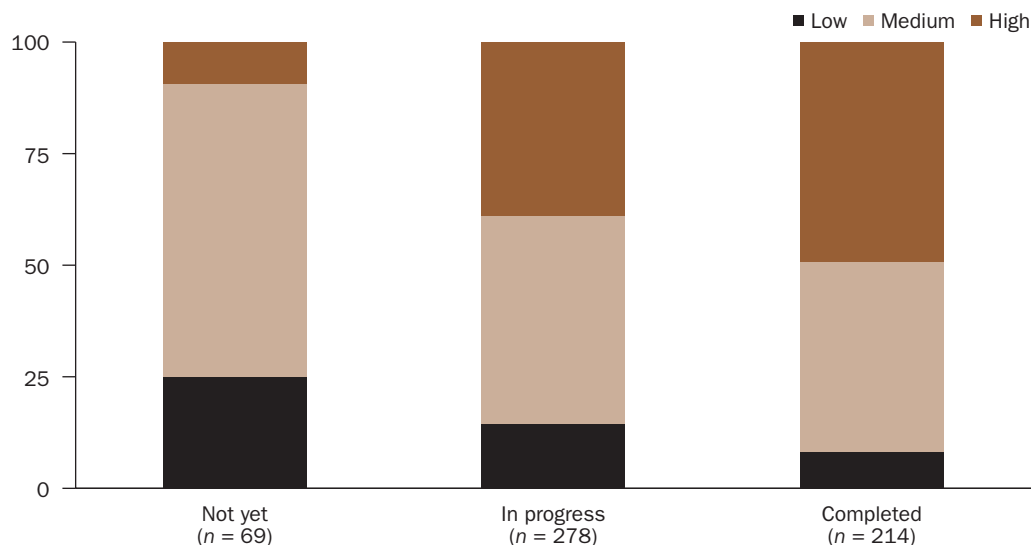
Change in the average rating of student engagement during early literacy skills instruction was associated with target school teachers' progress in the professional development program. Completing more of the professional development program was associated with increases in ratings of student engagement (figure 10). Across all observation windows, 49 percent of teachers who had completed the program were rated high, compared with 9 percent of teachers who had not started the program.

After typical growth among individuals who had not started the professional development program at each administration window was controlled for, teachers who had started the

Completing more of the professional development program was associated with increases in ratings of student engagement

Figure 10. The percentage of Mississippi grade K–3 teachers in target schools who were rated high in student engagement during early literacy skills instruction was higher among teachers who were in progress or who had completed the Language Essentials for Teaching Reading and Spelling professional development program than among teachers who had not started it, winter 2014–spring 2015

Percent of observed teachers

Note: The Language Essentials for Teaching Reading and Spelling professional development program was developed by Moats and Tolman (2009). Target schools are those identified by the Mississippi Department of Education as being most in need based on the percentage of students in the lowest two achievement levels on the statewide literacy assessment. At the conclusion of each observation the literacy coach assigned an overall student engagement during early literacy skills instruction rating of low, medium or high.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

professional development program were rated 0.18 point higher in student engagement during early literacy skills instruction than were teachers who had not started it. This was a statistically significant difference. Likewise, teachers who had completed the program were rated 0.22 point higher than were teachers who had not started it. This was a statistically significant difference. The overall growth in student engagement ratings was statistically significant for all teachers. By spring 2015, on average, teachers who had not started the program were in the 39th percentile, teachers who were in progress were in the 51st percentile, and teachers who had completed the program were in the 53rd percentile.

Completing more of the professional development program was associated with increases in ratings of teaching competencies

Change in the average rating of teaching competencies was associated with target school teachers' progress in the professional development program. Completing more of the professional development program was associated with increases in ratings of teaching competencies. After typical growth among individuals who had not started the program at each observation window were controlled for, teachers who were in progress were rated an average of 0.33 point higher in teaching competencies than were teachers who had not started it. This was a statistically significant difference. Similarly, teachers who completed the program were rated an average of 0.41 point higher than were teachers who had not started it. This was a statistically significant difference. By spring 2015, on average, teachers who had not started the program were in the 38th percentile, teachers who were in progress were in the 52nd percentile, and teachers who had completed the program were in the 54th percentile.

Implications of the study findings

The findings suggest that during the period when the Language Essentials for Teaching Reading and Spelling professional development program was implemented, teacher knowledge of early literacy skills, the quality of early literacy skills instruction, student engagement during early literacy skills instruction, and teaching competencies improved among educators who participated in the program over and above any increases found among educators generally. The findings suggest that progress in the program was associated with improvements in teacher knowledge, quality of instruction, student engagement, and teacher competencies.

However, it is important to note that while the findings were statistically significant, some of the findings may be viewed as substantively small differences. Thus, policymakers may want to consider whether some of the statistically significant findings reflect a meaningful amount of change or are large enough compared with what they may have expected to see happen. In light of the small differences, policymakers may want to consider whether the effort and resources invested in the professional development and state literacy coaches was worthwhile or sufficient and to consider what is necessary to achieve their policy goal of having every student read at or above grade level by the end of grade 3.

The study has several implications for alliance members, the Mississippi Department of Education, and the research community.

The finding that teacher knowledge of early literacy skills can improve is encouraging, because only 54 percent of items were answered correctly on average across the entire study. Although a seemingly low percentage, this is not contrary to findings from similar measures. The TKELS measures developed for this study were based on measures from previous studies such as Binks-Cantrell et al. (2012) and Bos et al. (2001) and included many of the same items as those studies. Binks-Cantrell et al. (2012) reported an average of 63 percent of items answered correctly; their findings were only slightly better than earlier findings by Bos et al. (2001), who found that preservice teachers answered 53 percent of items correctly and in-service teachers answered 60 percent of items correctly.

Ultimately, the goal of the Literacy-Based Promotion Act is for all Mississippi students to read at grade level by the end of grade 3. Although improvements in teacher knowledge and instruction are theoretically likely to result in student improvements, future research could consider including measures of K–3 student achievement. Although state departments of education often have large-scale initiatives to influence student achievement through a variety of mechanisms, such as professional development, they are rarely accompanied by a systematic investigation of achievement of the targeted change. This study does not evaluate the impact of the professional development program, but it highlights a systematic investigation of change in educators' knowledge and teachers' classroom practices—the targeted areas for change associated with the professional development program. The Language Essentials for Teaching Reading and Spelling professional development program and this study have included a widespread population of K–3 teachers, special educators, and administrators. Future research may consider narrowing the target population—for example, by focusing on only K–3 general education teachers or all K–3 teachers in a specific district. That is, by narrowing the target population, it would be feasible to develop a

This study does not evaluate the impact of the professional development program, but it highlights a systematic investigation of change in educators' knowledge and teachers' classroom practices—the targeted areas for change associated with the professional development program

sample with comprehensive contact information and ensure that all teachers in the sample are invited (and reminded) to participate.

To improve the reliability and validity of the classroom observations, future research may consider employing an outside observer not familiar with the teacher, the state literacy coach, or the teacher's professional development status. Outside observers may provide a more objective perspective. Additionally, future research may consider observations of classrooms that are not served by state literacy coaches. Observing schools with and without state literacy coaches would serve two purposes. First, it would provide a more accurate representation of all Mississippi elementary schools. Second, it would provide an opportunity to determine whether changes were associated with both coaching and professional development.

Lastly, and most important, to determine whether the professional development or coaching efforts were effective, researchers could employ, whenever feasible, a study design that allows for the examination of causal relationships (for example, a waitlist randomized controlled trial). In that this study found changes in a positive direction, it gives some support to the notion that such an examination might be worthwhile.

Limitations of the study

This study has several main limitations.

The research design of the study does not allow for evaluation of the impact or effectiveness of the Language Essentials for Teaching Reading and Spelling professional development program

First, the research design of the study does not allow for evaluation of the impact or effectiveness of the Language Essentials for Teaching Reading and Spelling professional development program. Therefore, it is not possible to determine whether the professional development caused the increases in teacher knowledge or improvement in the teaching practices that were observed. Moreover, while the study answered the stated research questions, it was not designed to answer what may be the most important question to policymakers and educators —especially beyond Mississippi—which is whether the teachers' improved knowledge and practices translated or will translate into improved student performance.

Second, because there was no state database of educators' email addresses, the Mississippi Department of Education was unable to ensure that all educators were invited to take the survey. The department emailed superintendents and curriculum coordinators, who then emailed principals, who then emailed teachers. Ideally, every K–3 educator across the state would have responded at each survey administration. It is unknown how many educators received an invitation to take the survey. Moreover, even if an educator received an invitation, there was no way to require the educator to take the survey. It is possible that those who responded were generally higher achieving, more confident educators.[2] Those who were less confident may have not taken the survey or responded only when the Mississippi Department of Education, districts, or schools made a more deliberate effort to recruit respondents. Because individual scores were not shared with educators, schools, or districts, there were no incentives to take the survey or complete the survey with accuracy. And because the surveys were completed online, there was no way to prevent educators from looking up answers or collaborating with their colleagues when answering questions.

Third, although there were 7,638 unique respondents, representing 68 percent of Mississippi K–3 educators, only 40 percent of respondents took the survey at multiple survey administrations. Only 5 percent of respondents took the survey at all four survey administrations.

Missing data can have serious impacts on the conclusions inferred from the analyses. One option to address the missing data would be to retain only cases with all complete data. However, it may introduce bias if those who responded at all administration windows were different from those who responded fewer than four times. Moreover, the precision of the estimates for professional development exposure could be reduced because the standard error (confidence band around the estimates) increases as sample size decreases. Likewise, a smaller sample size reduces the statistical power to detect statistically significant effects (Karahalios et al., 2013). Recent work comparing methods for handling missing data have shown that even with an overall missing data rate at 60 percent (similar to this study), the primary impact of missing data is on the precision of the estimates—the standard errors are larger—but the estimates themselves are not grossly affected when using analytic techniques similar to those used in this study (Dong & Peng, 2013). Nevertheless, the estimates for the relationship between teacher knowledge and progress in the professional development program may in fact be higher or lower than reported in this study.

Fourth, the study findings apply only to educators who participated in the study and may not be generalizable to all K–3 educators in Mississippi or in the nation. The TKELS survey was developed using questions that had previously been used in other research tools. Thus, the items had face validity. The final items used, combined into the two forms, had sufficient technical properties for use in this study; however, the piloting and norming of the TKELS survey took place within Mississippi. Therefore, there is no normative sample to tie the scores to besides the first set of Mississippi elementary educators who responded to the survey.

Moreover, it is unknown whether the TKELS respondents were representative of the full population of Mississippi elementary educators. That is, according to the original agreement with the vendor (the organization hired to administer the training, Cambium Learning), 60 percent of educators should have started the professional development program and 40 percent should have completed it by December 2014 (between fall 2014 and spring 2015 in the study's data collection periods). However, among survey respondents, a far smaller percentage had started or completed the professional development program. Likewise, according to the original agreement, by December 2015 (after fall 2015 in the study's data collection period), all educators should have started the program, and 80 percent of educators should have completed it; a smaller percentage had started or completed the professional development program by this deadline. This suggests that the proportion of survey respondents who were in progress or who had completed the program was less than the targeted percentage of educators and may not be an accurate representation of the full population of educators who participated in the professional development program.

Fifth, state literacy coaches conducted the observations, and they may have been biased toward seeing positive changes in teachers' quality, student engagement, and teaching competencies. Moreover, because the schools with state literacy coaches were generally those in most need of improvement, it is likely that teachers in those schools were most in need of the professional development, and thus the gains would be more apparent.

Sixth, the instructional findings have limited generalizability. Although the observations are representative of the schools the state literacy coaches served, not all schools had a state literacy coach. Observations did not occur in classrooms that did not have a state literacy coach. Therefore, the overall growth observed may be related to the coaching rather than to the professional development program.

Appendix A. Data and methods

This appendix describes the data sources and details of the study methods and results.

Data

This section describes the data used in the study and how the data were accessed. Tables with descriptive statistics for all measures are provided in appendix D.

Assessment of teacher knowledge. Regional Educational Laboratory (REL) Southeast developed the Teacher Knowledge of Early Literacy Skills (TKELS) survey at the request of the Mississippi Department of Education, which was interested in having a psychometrically valid assessment of teacher knowledge of early literacy skills. Importantly, the department wanted a tool with multiple forms to reduce the risk of testing effects, as educators would complete the assessment more than once.

REL Southeast compiled a bank of 75 items from previous research studies (Binks-Cantrell, Joshi, & Washburn, 2012; Bos, Mather, Dickson, Podhajski, & Chard, 2001; Carlisle, Correnti, Phelps, & Zeng, 2009; Carlisle, Kelcey, Rowan, & Phelps, 2011; Cunningham, Perry, Stanovich, & Stanovich, 2004; Mather, Bos, & Babur, 2001; Moats & Foorman, 2003; Reutzel et al., 2011; Salinger et al., 2010; Spear-Swerling & Cheesman, 2012). All items were in the public domain or were obtained from the study authors and used with their permission. All items were coded by content (comprehension, fluency, writing and grammar, vocabulary, spelling, phonological and phonemic awareness, and phonics) and type (application, knowledge, and teaching/pedagogy) and cross-walked with each of the Language Essentials for Teaching Reading and Spelling (Moats & Tolman, 2009) modules. The 75 items were randomly assigned, within content and type, into two forms of 45 items each. Each form had 30 unique items and 15 common items.

The first administration of the TKELS survey (spring 2014) served as the pilot of the 75 items. The survey software randomly assigned respondents to a form at each administration window. At the close of the administration window, all 75 items were simultaneously analyzed. First, unidimensionality was established by comparing a single-factor model to a two-factor model (application/knowledge and teaching/pedagogy), and to a bi-factor model using Mplus (Muthén & Muthén, 2015). The single-factor model provided the best fit to the data.

Next, all items were entered into a two-parameter item response theory model using BILOG-MG 3 (Zimowski, Muraki, Mislevy, & Bock, 2003). Items with poor discrimination parameters (less than 0.29) or extreme difficulty parameters (greater than ± 3.25) were removed. From the original 75 items piloted, 54 were retained, with an overall reliability of .76. The item response theory ability (theta) estimates from the retained items were converted to a *T*-score (mean = 50, standard deviation = 10, minimum = 20, and maximum = 80) and retained as the scores for survey administration 1.

The remaining 54 items were assigned to two forms (31 items per form, 23 unique items and 8 common items; see appendix B) with attention to distribute content and type between the two forms while maintaining consistency according to test characteristic curves. The mean discrimination parameters were 0.69 (standard deviation = 0.31) for form A and 0.65

(standard deviation = 0.28) for form B, and mean difficulty parameters were −0.03 (standard deviation = 1.46) for form A and −0.01 (standard deviation = 1.48) for form B.

Finally, item response theory true-score equating was conducted using PIE for PC GUI (Kolen & Brennan, 2004) to create a raw-score to T-score conversion table (see table B3 in appendix B; mean = 50, standard deviation = 10) between the two new forms and the single analysis of all 54 items. The two equated forms of 31 items each were randomly assigned by the survey software to respondents at each administration of the TKELS survey. Across all administrations, the mean TKELS score was 51.24 (standard deviation = 13.85); this was equivalent to answering 17–18 items correctly out of 31.

In summary, the TKELS survey includes two psychometrically equated short forms (forms A and B) with scores linked to the full TKELS survey, which was normed in spring 2014. At each survey administration the survey software randomly assigned respondents to form A or B; there were no controls to ensure that respondents received the same form in consecutive administration windows. The TKELS raw score is converted via item response theory true-score equating to a T-score (range = 20–80, mean = 50, and standard deviation = 10). A T-score of 50 is equivalent to answering approximately 50 percent of items correctly, a T-score of 20 is equivalent to answering 6 percent or less of the items correctly, and a T-score of 80 is equivalent to answering 88 percent or more of the items correctly.

Each iteration of the TKELS survey was administered online using Qualtrics. The Mississippi Department of Education decided when to administer each round of the TKELS survey; the first administration occurred around the same time that educators were notified that they could register for the professional development. The department solicited responses to the TKELS survey via a chain of emails through district superintendents, curriculum coordinators, and school administrators, to teachers. Additionally, respondents from each prior administration were automatically emailed through Qualtrics. At each administration, educators were randomly assigned to a form by the survey software to control for form effects confounding with administration windows. Because respondents had equal probability of receiving either form at every administration, it is possible that respondents received the same form in consecutive administration windows. Although this would increase the potential for testing effects, because it was random, there is no reason to believe testing effects would be systematically influencing growth estimates.

The number of respondents varied at each administration (n = 1,789 to 4,712; table A1), representing 16–42 percent of all K–3 educators in Mississippi. There were 12,155 responses across all administrations, representing 7,638 unique individuals. This represents approximately 68 percent of the estimated 11,203 K–3 educators in Mississippi (Mississippi Department of Education, 2014). Responses were received from all 148 Mississippi districts, though not every district was represented in each administration. The percentage of districts represented ranged from 88 percent to 97 percent across observation windows (see table A1).

At the completion of each administration of the TKELS survey, the Mississippi Department of Education provided REL Southeast with the data that included a unique educator identifier, item-level responses to the TKELS survey, educator-reported district, progress in the professional development program, and demographic information, including education and teaching experience. Germaine to this report are only the TKELS scores, district, and

Table A1. Survey administration dates and number of respondents

Survey administration window number	Survey administration window descriptor	Approximate dates	Number of districts represented	Percent of districts represented	Number of respondents	Percent of K–3 educators[a]
1	Spring 2014	March–April 2014	130	88	2,457	21.9
2	Fall 2014	October–November 2014	144	97	4,712	42.1
3	Spring 2015	April–May 2015	140	95	1,789	16.0
4	Fall 2015	September–October 2015	144	97	3,197	28.5

Note: The number of respondents in this table are the total number of respondents in each administration window. In some cases, insufficient data were available for the analysis (for example, because respondents did not finish the survey). The number of responses used in the analyses is reported in figures 1 and 2 in the main text.

a. Based on the estimate of 11,203 educators in Mississippi from Mississippi Department of Education (2014).

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

professional development progress. Characteristics of the sample of respondents at each administration are presented in table A2.

To determine whether the responding sample was similar to that of Mississippi's elementary educators, the characteristics of the full sample of respondents was compared with the characteristics of Mississippi elementary educators responding to the 2011–12 Schools and Staffing Survey (U.S. Department of Education, 2016b).[3] Chi-square tests revealed that the responding sample was comparable to Mississippi elementary educators in terms of highest degree obtained, race, and gender but not ethnicity (table A3). The sample of respondents had a smaller proportion of Hispanic educators than did the general population of Mississippi elementary educators.

Coach's Classroom Observation Tool. The Mississippi Department of Education hired state literacy coaches to service K–3 teachers, elementary schools, and districts most in need of improvement. Because some districts or schools may have employed literacy coaches who were not official state literacy coaches, the report specifically refers to the state literacy coaches. Hiring and training of state literacy coaches began in fall 2013 and is ongoing; therefore, the number of coaches varied at each observation window, and the number of teachers observed, and data received varied at each observation window. Generally, each coach served two schools and completed eight observations at each window (one per grade K–3 per school). However, some coaches served schools that did not include all grades K–3 (for example, a grade 3–5 elementary school). Although all the state literacy coaches were trained in the Language Essentials for Teaching Reading and Spelling professional development program, the coaching was not specific to the program's curriculum. In 2013/14 approximately 29 coaches served 50 schools, and in 2014/15 approximately 51 coaches served 87 schools. The Mississippi Department of Education hired new coaches throughout the school year, so it is not possible to determine exactly how many coaches were part of the team at each observation window. Likewise, not every coach was fully trained and eligible to submit observation data at each observation window. The number of coaches submitting observations is listed in table A4.

Table A2. Characteristics of the sample of Teacher Knowledge of Early Literacy Skills survey respondents

Characteristics	Spring 2014		Fall 2014		Spring 2015		Fall 2015	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total number of respondents	2,457		4,712		1,789		3,197	
Highest degree								
Undergraduate	1,365	55.6	2,566	54.5	861	48.1	1,642	51.4
Master's	960	39.1	1865	39.6	788	44.0	1282	40.1
Specialist	68	2.8	153	3.2	63	3.5	126	3.9
Ed.D./Ph.D.	13	0.5	35	0.7	15	0.8	33	1.0
Race								
White	1,900	77.3	3,441	73.0	1,315	73.5	2,274	71.1
Black	401	16.3	914	19.4	254	14.2	517	16.2
Other[a]	25	1.0	66	1.4	30	1.7	55	1.7
Ethnicity								
Hispanic	18	0.7	26	0.6	19	1.1	21	0.7
Not Hispanic	1,840	74.9	3,479	73.8	1,311	73.3	2,291	71.7
Gender								
Male	23	0.9	124	2.6	39	2.2	69	2.2
Female	2,345	95.4	4,368	92.7	1,623	90.7	2,865	89.6
Current position								
Kindergarten teacher	607	24.7	973	20.6	373	20.8	683	21.4
Grade 1 teacher	589	24.0	957	20.3	338	18.9	639	20.0
Grade 2 teacher	451	18.4	927	19.7	359	20.1	611	19.1
Grade 3 teacher	463	18.8	909	19.3	305	17.0	596	18.6
Special educator	119	4.8	456	9.7	197	11.0	299	9.4
Literacy coach	19	0.8	50	1.1	29	1.6	39	1.2
Administrator	20	0.8	46	1.0	52	2.9	102	3.2
Other	189	7.7	394	8.4	136	7.6	226	7.1

Note: The number of respondents in this table is the total number of respondents. In some cases, insufficient data were available for the analysis (for example, because respondents did not finish the survey). The number of responses used in the analyses is reported in figures 1 and 2 in the main text. Percentages are representative of the full responding sample for that administration window. Percentages may not sum to 100 because not all respondents provided information or because of rounding.

a. Includes Asian, American Indian/Alaska Native, Native Hawaiian or other Pacific Islander, and other.

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

The REL Southeast developed the Coach's Classroom Observation Tool (CCOT) at the request of the Mississippi Department of Education, which was interested in having a uniform, research-based tool to conduct formal observations of classroom literacy instruction. The CCOT combines aspects of the Timed Observations/Student Engagement and Teacher Competency Checklist (Foorman, Goldenberg, Carlson, Saunders, & Pollard-Durodola, 2004; Foorman & Schatschneider, 2003; Foorman et al., 2006) and the Instructional Content Emphasis-Revised (Edmonds & Briggs, 2003).

REL Southeast provided the Mississippi Department of Education with the CCOT as a fillable PDF (see appendix C) that state literacy coaches used to enter the observation data. Additionally, the REL Southeast produced a training manual and trained the state literacy coaches through a series of workshops to conduct the observations. To establish interrater reliability, the state literacy coaches were required to meet an 80 percent agreement rate on a training video in order to use the CCOT. Additionally, the state literacy coaches video

Table A3. Characteristics of the sample compared with characteristics of all Mississippi elementary educators according to the 2011–12 Schools and Staffing Survey

Value	Total number of respondents	Percent of respondents	SASS estimated percentage	Chi square test
Total number of respondents	12,155			
Highest degree				
Undergraduate	6,434	54	62	
Master's	4,895	41	30	0.15
Specialist	410	3	3	
Ed.D./Ph.D.	96	1	<1	
Race				
White	8,930	80	78	
Black	2,086	19	22	0.64
Other[a]	176	2	<1	
Ethnicity				
Hispanic	84	1	8	0.01
Not Hispanic	8,921	99	92	
Gender				
Male	255	2	5	0.20
Female	11,201	98	95	

SASS is Staffing and Schools Survey (U.S. Department of Education, 2016b).

Note: Percentages may not sum to 100 because of rounding.

a. Includes Asian, American Indian/Alaska Native, Native Hawaiian or other Pacific Islander, and other.

Source: Authors' analysis of data from Mississippi Department of Education (2015b) and computations by the National Center for Education Statistics PowerStats on U.S. Department of Education (2016b).

Table A4. Classroom observation dates and number of observations received

Observation window number	Observation window descriptor	Approximate dates	Number of coaches conducting observations	Number of schools with observation data	Percent of target schools with observation data	Number of teachers observed
1	Winter 2014	January–February 2014	22	40	80	121
2	Spring 2014	April–May 2014	25	41	82	124
3	Fall 2014	October–November 2014	38	57	66	165
4	Spring 2015	March–April 2015	33	50	57	151

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

recorded one observation session per school per observation window, and the video was shared with a randomly selected state literacy coach who completed an observation. The live and video observations were compared to compute interrater reliability. The mean agreement between the live and video observations was 78 percent.

The Mississippi Department of Education adopted the CCOT as the official classroom observation tool for state literacy coaches and required all state literacy coaches to use the CCOT to collect data on one randomly selected teacher per grade (K–3 only), per school served by the coach.[4] Data were collected four times between winter 2014 and spring 2015 (see table A4). Coaches were instructed to observe the same teacher within

the academic year to the fullest extent possible: 14 percent of teachers were observed four times, 7 percent were observed three times, 60 percent were observed twice, and 20 percent of teachers were observed only once. As with the TKELS survey, the Mississippi Department of Education decided when coaches would conduct the observations. Infrastructure was in place for observations before the surveys; therefore, the first round of observations occurred before the first administration of the TKELS survey.

During the observations, the state literacy coach created a running record of instruction. For each instructional event (defined as a distinct or unique activity where the content, grouping, and materials are coordinated around a certain domain or component of reading instruction), the state literacy coach coded the main instructional category, grouping of instruction, student engagement, and quality of instruction. At the completion of the observation, the state literacy coach gave an overall student engagement and overall quality score. Additionally, the state literacy coach completed the checklist of teaching competencies. The Mississippi Department of Education provided REL Southeast with copies of the PDF documents, which were compiled for analysis. At each observation window, the number of state literacy coaches on staff and trained to conduct observations varied. Therefore, the number of teachers observed also varied (n = 121 to 165). Six hundred observations took place across all observation windows, representing 316 unique teachers.

Observations were conducted in 63 schools (50 schools were served in 2013/14 and 87 were served in 2014/15), though not every school was represented at each observation. The number of schools represented at each observation window ranged from 40 to 57 (see table A4). Forty-two schools were represented in 2013/14, representing 84 percent of the target schools. Sixty-one schools were represented in 2014/15, representing 70 percent of target schools. Characteristics of the sample of observed teachers at each observation window are presented in table A5. Across observation windows, the average observation length was 72–86 minutes. Because state literacy coaches conducted the observations, observation data were limited to teachers only in schools with a state literacy coach.

For study purposes, the CCOT produced three sets of variables: quality of instruction, student engagement, and teaching competencies. Each of these sets of variables is described next.

Table A5. Characteristics of the sample of observed teachers, by observation window

Characteristic	Winter 2014 (n = 121)			Spring 2014 (n = 124)			Fall 2014 (n = 165)			Spring 2015 (n = 151)		
	Mean	SD	Number of teachers	Mean	SD	Number of teachers	Mean	SD	Number of teachers	Mean	SD	Number of teachers
Number of students in class	21	4		20	4		21	4		21	3	
Total minutes of observation	85.8	22.9		85.9	30.9		72.1	18.9		71.9	19.8	
Kindergarten			29			30			41			39
Grade 1			28			32			41			38
Grade 2			32			31			42			40
Grade 3			32			31			41			34

SD is standard deviation.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Quality of instruction. For each instructional event, and at the completion of an observation, the state literacy coaches rated the quality of the literacy instruction on a four-point scale: 1 = weak, 2 = low average, 3 = high average, 4 = excellent.

The observation tool automatically calculated the mean quality rating for each instructional category across all instructional events. This resulted in 22 quality variables—20 for each instructional category, one mean rating across all instructional categories (mean = 2.47, standard deviation = 0.61), and one overall quality rating assigned by the observer (mean = 2.49, standard deviation = 0.85). The correlation between the mean quality rating and the overall quality rating ranged from .72 to .82 across the four observations. Thus, for the growth models, the study team calculated a single mean quality score as the mean of the overall quality rating and the average quality rating across all instructional categories. Across all observations the mean quality score was 2.46 (standard deviation = 0.65).

Student engagement. For each instructional event, and at the completion of an observation, the state literacy coaches rated student engagement on a three-point scale: 1 = low engagement, 2 = medium engagement, 3 = high engagement.

The observation tool automatically calculated the mean student engagement rating for each instructional category across all instructional events. This resulted in 22 engagement variables—20 for each instructional category, one mean rating across all instructional categories (mean = 2.34, standard deviation = 0.49), and one overall quality rating assigned by the observer (mean = 2.30, standard deviation = 0.679). The correlation between the mean and the overall quality rating ranged from .78 to .79 across the four observation windows. Thus, for the growth models, the study team calculated a single mean engagement score as the mean of the overall engagement rating and the average engagement rating across all instructional categories. Across all observations the mean quality score was 2.32 (standard deviation = 0.54).

Teaching competencies. The state literacy coaches used a checklist to rate teachers' competencies on 30 items related to planning, management, instruction, monitoring of student learning, and personal characteristics. Each item was on a five-point scale: 1 = never, 2 = less than half the time, 3 = half the time, 4 = greater than half the time, 5 = all the time.

Following the procedures conducted by previous researchers using the same tool (Foorman et al., 2004; Foorman & Schatschneider, 2003; Foorman et al., 2006), the study team conducted a reliability analysis of the 30-item checklist. Cronbach's alpha ranged from .95 to .96 across the four observation windows. This suggested sufficient reliability to calculate a single mean score from the 30 items. Across all observations, the mean score on the checklist was 3.61 (standard deviation = 0.83).

Methods

This section describes the analytic methods used in the study.

Multilevel random effects growth curve models with time-varying covariates for teacher knowledge growth. To describe teacher knowledge growth, a series of three-level (time, teacher, and district[5]) random effects growth models were built and tested (Raudenbush &

Bryk, 2002). Time-varying covariates were used to account for differential teacher knowledge growth depending on progress in the professional development program (McCoach & Kaniskan, 2010).

In the model-building process the first focus was on establishing the level-1 model that accurately reflected the shape of the growth trajectory. If the level-1 model is misspecified, incorrect parameter estimates and errors of inferences can occur in the upper levels of the model (McCoach & Kaniskan, 2010). Time was treated as a time-variant predictor and was centered at the first TKELS survey administration—spring 2014. Growth represented change by survey response window—spring 2014, fall 2014, spring 2015, and fall 2015. Because there were four observation windows of data included in the analysis, both linear and quadratic growth could be tested. Model results suggested that the addition of a quadratic growth term best fit the data (table A6).

Progress in the professional development program was coded using two dummy codes—in progress and complete—with not started as the referent group. "In progress" denotes that an educator had started at least one of the online modules (educators could not attend a face-to-face workshop until they had completed the corresponding online content). "Complete" denotes an educator who completed both sets of online modules and both sets of face-to-face workshops. Progress was reported at each administration window and entered as a time-varying covariate. Model results suggested that adding the time-varying covariates resulted in a better model fit (see table A6). In summary, five models were built and tested:

- Null model–random intercept: this model was used to identify the variance components.
- Model 1: null model plus fixed linear growth.
- Model 2: model 1 plus fixed quadratic growth.
- Model 3: model 2 with random linear and quadratic growth.
- Model 4: model 3 plus time-varying covariates for progress in the professional development program.

Table A6. Fit indices for models of teacher knowledge

Model	Degrees of freedom	Akaike information criteria	Bayesian information criteria	Log likelihood ratio test	Deviance	Chi square test[a]	Degrees of freedom
Null model	4	95,564	95,593	−47,778			
Model 1	5	95,396	95,433	−47,693	95,386	169.967***	1
Model 2	6	95,384	95,429	−47,686	95,372	13.168***	1
Model 3	16	95,266	95,385	−47,617	95,234	138.13***	10
Model 4	18	95,202	95,335	−47,583	95,166	68.498***	2

*** Significant at $p < .001$.

a. The chi-square test evaluates the difference between model fit of the previous model to the current model. For example, the chi-square test for model 2 evaluates the difference from model 1.

Note: The null model is the random intercept model used to identify the variance components. Model 1 adds fixed linear growth. Model 2 adds fixed quadratic growth. Model 3 changes to random linear and quadratic growth. Model 4 adds the time-varying covariates for progress in the professional development.

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

All models were analyzed using R's Package lme4 (Bates, Maechler, Bolker, & Walker, 2014). In the model-building process, model fit was examined according to Akaike information criteria, Bayesian information criteria, log-likelihood, deviance, and chi-square tests.

Both the growth and intercept estimates were tested as fixed and random effects. The final model used to test for the association between teacher knowledge and professional development, based on model fit indices, was model 4, represented by:

Level-1 (time) $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (A1)
$$Y_{tij} = \pi_{0ij} + \pi_{1ij}(wave_{tij}) + \pi_{2ij}(wave_{tij}^2) + \pi_{3ij}(InProgress_{tij}) + \pi_{4ij}(Completed_{tij}) + e_{tij}$$

Level-2 (educator)
$$\pi_{0ij} = \beta_{00j} + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + r_{1ij}$$
$$\pi_{2ij} = \beta_{20j} + r_{1ij}$$

Level-3 (district)
$$\beta_{00j} = \gamma_{000} + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + u_{10j}$$
$$\beta_{20j} = \gamma_{200} + u_{10j}$$

where Y_{tij} is the teacher knowledge t-score at time t for educator i in district j, $wave_{tij}$ is the measurement occasion at time t for school i in district j, $wave_{tij}^2$ is the measurement occasion squared at time t for school i in district j, π_{3ij} is the difference when an educator is in the process of completing the professional development program, and π_{4ij} is the difference when an educator has completed the professional development program. Both π_{3ij} and π_{4ij} were treated as fixed effects.

The null model suggested that 56 percent of total variance was between educators, 4 percent was between districts, and the remaining 40 percent was within educator over time.

Model 3 (table A7) addressed research question 1 on how knowledge of early literacy skills among Mississippi's K–3 educators changed between spring 2014 and fall 2015. On average, teacher knowledge exhibited nonlinear change ($\beta_{10j} = -0.80$ and $\beta_{20j} = 0.57$). According to fitted means, there was a slight decrease (0.23 point) between administration window 1 and administration window 2 and increasingly positive changes between administration window 2 and administration window 3 and between administration window 3 and administration window 4 (see figure 1 in the main text). In administration window 1 (spring 2014) the average teacher knowledge score was 49.55 (β_{00j}); by administration window 4 (fall 2015) it was 52.28. Thus, the average change, based on fitted means, between administration window 1 and administration window 2 was 2.73 points, which is equivalent to 0.27 standard deviation.

Model 4 (see table A7) addressed research question 2 on whether the change in teacher knowledge of early literacy skills was associated with educators' progress in the professional development program. The model suggested that the average knowledge score among educators who had not started the professional development program was 49.56

at administration 1 (β_{00j}) and followed a similar nonlinear change pattern as in model 3, (β_{10j} = –1.34 and β_{20j} = 0.62). When an educator was in progress (π_{3ij}), there was an associated increase of 1.42 points, and when an educator had completed the professional development program (π_{4ij}), there was an associated increase of 2.90 points. The final model had a pseudo R^2 of .14, suggesting that 14 percent of the variance in teacher knowledge score was accounted for by time and progress in the professional development program. Pseudo R^2 indicates the amount of variation in teacher knowledge statistically accounted by the models. A larger pseudo R^2 means that more variation is accounted for and associated with model predictors—in this case, progress in the professional development program. A smaller pseudo R^2 means that more variation remains that has not been statistically accounted for. In this model, a pseudo R^2 of .14 indicates 86 percent of the variation in teacher knowledge is still not accounted for and is likely related to other factors.

Multilevel random effects growth curve models with time-varying covariates for changes in classroom instruction. To describe changes in classroom instruction—measured by quality of instruction, student engagement, and teaching competencies—a model-building approach similar to the one used for modeling teacher knowledge growth was used. For each classroom instruction variable (quality of instruction, student engagement, and teaching competencies), a series of three-level (time, teacher, and school) random effects models were built and tested (Raudenbush & Bryk, 2002).

Again, a three-level model is preferred over simpler models because it accounts for both intra- and inter-individual changes while still accounting for the nested nature of the data. That is, rather than looking at overall changes in the group of individuals observed, changes at the individual teacher level were modeled. Findings are representative of an

Table A7. Teacher knowledge model results

Fixed effects					Random effects			
Coefficient	Estimate	Standard error	Degrees of freedom	*t*-value	Groups	Coefficient	Variance	Standard deviation
Model 3: final growth model								
(Intercept)	49.55	0.38	113.85	131.29***	Educator	(Intercept)	74.32	8.62
Wave	−0.80	0.43	124.11	−1.85		Wave	5.85	2.42
Wave^2	0.57	0.13	109.60	4.43***		Wave^2	0.61	0.78
					District	(Intercept)	9.45	3.07
						Wave	5.24	2.29
						Wave^2	0.43	0.65
					Residual		67.77	8.23
Model 4: final growth model with time-varying covariates for progress in the professional development program								
(Intercept)	49.56	0.38	113.00	132.16***	Educator	(Intercept)	73.77	8.59
Wave	−1.34	0.43	117.00	−3.12**		Wave	5.85	2.42
Wave^2	0.62	0.13	96.00	4.93***		Wave^2	0.61	0.78
In progress	1.42	0.33	7,860.00	4.27***	District	(Intercept)	9.21	3.04
Complete	2.90	0.35	5,565.00	8.23***		Wave	4.65	2.16
						Wave^2	0.37	0.61
					Residual		67.84	8.24

** Significant at $p < 01$; *** significant at $p < .001$.

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

average teacher in an average school (of the schools observed, which were primarily low performing) at the specific observation window. Moreover, using three-level data allows for missing level-1 data (time); as long as the level-1 equation is correctly specified, parameter estimates are not substantively different (Zaidman-Zait & Zumbo, 2013). This means that not every person had to be observed at every observation window, and information was not lost because of missing data points. This increases power and the ability to detect smaller effects.

Time-varying covariates were used to account for differential changes in teaching practices depending on progress in the professional development program (McCoach & Kaniskan, 2010). Time was treated as a time-variant predictor and was centered at observation window 1 (winter 2014). Growth represented change by observation window. Because there were four observation windows in the analysis, both linear and quadratic growth was tested. For all variables the addition of the quadratic term did not result in improved model fit, so it was not included in the final model (table A8).

Table A8. Fit indices for models of instructional ratings

Model	Degrees of freedom	Akaike information criteria	Bayesian information criteria	Log likelihood ratio test	Deviance	Chi-square test[a]	Degrees of freedom
Quality of instruction							
Null model	4	1,078	1,096	−535	1,070		
Model 1	5	1,045	1,067	−517	1,035	35.562***	1
Model 2	6	1,047	1,073	−517	1,035	0.111	1
Model 3	9	1,048	1,087	−515	1,030	4.918[b]	4
Model 4	11	984	1,031	−481	962	68.050***	2
Student engagement							
Null model	4	879	896	−435	871		
Model 1	5	861	883	−426	851	19.687***	1
Model 2	6	863	889	−426	851	0.048	1
Model 3	9	869	908	−425	851	0.281[b]	4
Model 4	11	825	872	−401	803	48.063***	2
Teaching competencies							
Null model	4	1,258	1,275	−625	1,250		
Model 1	5	1,220	1,242	−605	1,210	39.950***	1
Model 2	6	1,222	1,248	−605	1,210	0	1
Model 3	9	1,200	1,239	−591	1,182	27.865***[b]	4
Model 4	11	1,135	1,182	−557	1,113	68.753***	2

*** Significant at $p < .001$.

Note: The null model is the random intercept model used to identify the variance components. Model 1 adds fixed linear growth. Model 2 adds fixed quadratic growth. Model 3 removed the quadratic term and changes the growth term to random. Model 4 adds the time-varying covariates for progress in the professional development.

a. The chi-square test evaluates the difference between model fit of the previous model to the current model except for model 3 as discussed below. For example, the chi-square test for model 2 evaluates the difference from model 1.

b. The chi-square test for model 3 tests the difference between model 1 and 3 as the quadratic term from model 2 was removed due to nonsignificance.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Progress in the professional development program was coded using two dummy codes—in progress and completed—with not started as the referent group. "In progress" denotes that a teacher had started at least one of the online modules or at least one of the face-to-face workshops. "Completed" denotes a teacher who completed both sets of online modules and both sets of face-to-face workshops. Progress was reported for each teacher at each observation window and was thus entered as a time-varying covariate. Model results suggested that adding in the time-varying covariates resulted in a better model fit (see table A8). In summary, five models were built and tested:

- Null model—random intercept: this model was used to identify the variance components.
- Model 1: null model plus fixed linear growth.
- Model 2: model 1 plus fixed quadratic growth.
- Model 3: model 1 with random linear growth.
- Model 4: model 3 plus time-varying covariates for progress in the professional development.

All models were analyzed using R's Package lme4 (Bates et al., 2014). In the model-building process, model fit was examined according to Akaike information criteria, Bayesian information criteria, log likelihood, deviance, and chi-square tests. Both the growth and intercept estimates were tested as fixed and random effects. The final model used to test for the association between quality of instruction, student engagement, teaching competencies and professional development, based on model fit indices (see table A8), was model 4, represented by:

Level-1 (time) $\hfill$ (A2)
$$Y_{tij} = \pi_{0ij} + \pi_{1ij}(wave_{tij}) + \pi_{2ij}(InProgress_{tij}) + \pi_{3ij}(Completed_{tij}) + e_{tij}$$

Level-2 (teacher)
$$\pi_{1ij} = \beta_{10j} + r_{1ij}$$

Level-3 (school)
$$\beta_{10j} = \gamma_{100} + u_{10j}$$

where Y_{tij} is the classroom practice variable of interest (quality of instruction, student engagement, or teaching competencies) at time t for teacher i in school j, $wave_{tij}$ is the measurement occasion at time t for teacher i in school j, π_{2ij} is the difference when a teacher is in the progress of completing the professional development, and π_{3ij} is the difference when a teacher has completed the professional development. Both π_{2ij} and π_{3ij} were treated as fixed effects.

The primary parameters of interest in the final model were the time-varying covariates for progress in the professional development program. These parameters are the "complete" and "in progress" effects in model 4 (tables A9–A11) and are parameters π_{2ij} and π_{3ij} in equation A2. These parameters identified how quality of instruction, student engagement, or teaching competencies changed when a teacher was in progress or had completed the professional development program.

Quality of instruction. For quality of instruction the null model suggested that 51 percent of the total variance was between teachers, 5 percent was between schools, and the remaining 44 percent was within teacher over time.

Model 3 (see table A9) addressed research question 3 on how ratings of quality of early literacy skills instruction changed. Model results suggested that the ratings of quality of instruction steadily increased (β_{10j} = 0.15). In observation window 1 (winter 2014) the average quality of instruction rating was 2.18 (β_{00j}). According to fitted means, by observation window 4 (spring 2015) it was 2.63. This is reflective of the quality of early literacy skills instruction moving from low average to high average.

Model 4 (see table A9) addressed research question 4 on whether the change in the average rating of quality of early literacy skills instruction was associated with teachers' progress in the professional development program. The model suggested that the average rating among teachers who had not started the program was 2.06 at observation 1 (β_{00j}) and followed a similar linear growth as in model 3, (β_{10j} = 0.10). When a teacher was in progress (π_{2ij}), there was an associated non–statistically significant increase of 0.15 point, and when a teacher had completed the program (π_{3ij}), there was an associated statistically significant increase of 0.30 point. The final model had a pseudo R^2 of .07, suggesting that 7 percent of the variance in ratings of quality of instruction was accounted for by time and progress in the professional development program.

Student engagement. For student engagement the null model suggested that 45 percent of the total variance was between teachers, 5 percent was between schools, and the remaining 50 percent was within teacher over time.

Model 3 (see table A10) addressed research question 3 on how ratings of student engagement during early literacy skills instruction changed. Model results suggested that ratings of student engagement steadily increased (β_{10j} = 0.09). At observation window 1 (winter 2014) the average student engagement rating was 2.15 (β_{00j}). According to fitted means, by observation window 4 (spring 2015) it was 2.42. Although this increase was statistically significant, student engagement was still categorized as "medium."

Table A9. Quality of instruction model results

Fixed effects					Random effects			
Coefficient	Estimate	Standard error	Degrees of freedom	*t*-value	Groups	Coefficient	Variance	Standard deviation
Model 3: final growth model								
(Intercept)	2.18	0.08	18.42	28.58***	Teacher	(Intercept)	0.24	0.49
Wave	0.15	0.03	14.03	5.15***		Wave	0.00	0.01
					School	(Intercept)	0.07	0.26
						Wave	0.01	0.08
					Residual		0.17	0.41
Model 4: final growth model with time-varying covariates for progress in the professional development program								
(Intercept)	2.06	0.10	37.60	21.70***	Teacher	(Intercept)	0.22	0.47
Wave	0.10	0.03	24.40	3.06**		Wave	0.00	0.01
In progress	0.15	0.08	353.80	1.84	School	(Intercept)	0.06	0.25
Complete	0.30	0.10	449.40	3.05**		Wave	0.01	0.09
					Residual		0.17	0.42

** Significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Table A10. Student engagement model results

	Fixed effects				**Random effects**			
Coefficient	**Estimate**	**Standard error**	**Degrees of freedom**	***t*-value**	**Groups**	**Coefficient**	**Variance**	**Standard deviation**
Model 3: final growth model								
(Intercept)	2.15	0.05	16.37	40.44***	Teacher	(Intercept)	0.13	0.36
Wave	0.09	0.02	12.91	4.45***		Wave	0.00	0.00
					School	(Intercept)	0.02	0.15
						Wave	0.00	0.02
					Residual		0.14	0.37
Model 4: final growth model with time-varying covariates for progress in the professional development program								
(Intercept)	2.02	0.07	40.50	27.95***	Teacher	(Intercept)	0.12	0.35
Wave	0.06	0.03	35.80	2.29*		Wave	0.00	0.00
In progress	0.18	0.07	322.70	2.49*	School	(Intercept)	0.02	0.14
Complete	0.22	0.08	441.30	2.60**		Wave	0.00	0.03
					Residual		0.15	0.39

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

Model 4 (see table A10) addressed research question 4 on whether change in average rating of student engagement was associated with teachers' progress in the professional development program. The model suggested that the average rating among teachers who had not started the program was 2.02 at observation window 1 (β_{00j}) and followed a similar linear growth as in model 3, ($\beta_{10j} = 0.06$). When a teacher was in progress (π_{2ij}), there was an associated statistically significant increase of 0.18 point, and when a teacher had completed the program (π_{3ij}), there was an associated statistically significant increase of 0.22 point. The final model had a pseudo R^2 of .01, suggesting that 1 percent of the variance in ratings of student engagement was accounted for by time and progress in the professional development program.

Teaching competencies. For teaching competencies the null model suggested that 56 percent of the total variance was between teachers, 14 percent was between schools, and the remaining 30 percent was within teacher over time.

Model 3 (see table A11) addressed research question 3 on how ratings of teaching competencies during early literacy skills instruction changed. Model results suggested that ratings of teaching competencies steadily increased ($\beta_{10j} = 0.23$). At observation window 1 (winter 2014) the average rating of teacher competencies was 3.20 (β_{00j}). According to fitted means, by observation window 4 (spring 2015), it was 3.89. This is reflective of ratings of teaching competencies moving from less than half the time to half the time.

Model 4 (see table A11) addressed research question 4 on whether change in average ratings of teaching competencies was associated with teachers' progress in the professional development program. The model suggested the average rating among teachers who had not started the program was 2.99 at observation window 1 (β_{00j}) and followed a similar linear growth as in model 3, ($\beta_{10j} = 0.15$). When a teacher was in progress (π_{2ij}), there was an associated statistically significant increase of 0.33 point, and when a teacher had

Table A11. Teaching competencies model results

Fixed effects					Random effects			
Coefficient	Estimate	Standard error	Degrees of freedom	*t*-value	Groups	Coefficient	Variance	Standard deviation
Model 3: final growth model								
(Intercept)	3.20	0.13	10.13	24.09***	Teacher	(Intercept)	0.50	0.71
Wave	0.23	0.05	12.87	4.23***		Wave	0.04	0.19
					School	(Intercept)	0.30	0.55
						Wave	0.04	0.21
					Residual		0.14	0.38
Model 4: final growth model with time-varying covariates for progress in the professional development program								
(Intercept)	2.99	0.13	16.60	22.14***	Teacher	(Intercept)	0.45	0.67
Wave	0.15	0.05	16.60	3.02**		Wave	0.03	0.17
In progress	0.33	0.09	324.90	3.55***	School	(Intercept)	0.22	0.47
Complete	0.41	0.11	388.60	3.70***		Wave	0.03	0.16
					Residual		0.16	0.40

** Significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data from Mississippi Department of Education (2015a).

completed the program (π_{3ij}), there was an associated statistically significant increase of 0.41 point. The final model had a pseudo R^2 of .16, suggesting that 16 percent of the variance in ratings of teaching competencies was accounted for by time and professional development progress.

Appendix B. Teacher Knowledge of Early Literacy Skills survey

This appendix includes both of the Teacher Knowledge of Early Literacy Skills survey forms (tables B1 and B2) and the raw-to-scale score conversions (table B3) based on the psychometric analysis described in appendix A. All items were sourced from other research tools either in the public domain or used with permission (Binks-Cantrell et al., 2012; Bos et al., 2001; Carlisle et al., 2009; Carlisle et al., 2011; Cunningham et al., 2004; Mather et al., 2001; Moats & Foorman, 2003; Reutzel et al., 2011; Salinger et al., 2010; Spear-Swerling & Cheesman, 2012; see appendix A for details on item selection and modification).

Because the forms were equated—meaning that educators would receive the same score regardless of form—respondents were randomly assigned (by the survey software) to a form at each survey administration. This ensured that a particular form was not associated with a particular administration window. Approximately half of respondents received form A, and the other half received form B at each administration. There were no controls to ensure that the same form was not administered twice in a row to the same respondent. That is, each respondent had an equal chance of receiving either form at each administration, so it is possible respondents may have received the same form in consecutive survey administration windows.

Table B1. Teacher Knowledge of Early Literacy Skills survey form A

Item	Question	Choice A	Choice B	Choice C	Choice D	Correct answer
1	What does morphemic analysis help students do?	identify letter-sound correspondence	blend speech sounds	examine words for meaningful parts	separate syllables into onsets and rimes	C
2	What is a requirement of a syllable?	it contains at least one consonant letter	it contains no more than one vowel letter	it be a pronounceable unit	it contains no more than one phoneme	C
3	What can sentence combining help students learn to do?	question the text	correct grammatical errors	form complex sentence structures	analyze word structure	C
4	What is vocabulary instruction in the primary grades most concerned with teaching students?	highly frequent words	base words and meaningful parts (e.g., prefixes, suffixes)	decodable words	word meanings	D
5	Which of the following is NOT an irregular, high frequency word?	when	does	were	said	A
6	If "tife" is a word, the letter "i" would probably sound like the "i" in which word?	if	beautiful	find	ceiling	C
7	How should writing lessons be explicitly taught?	by explaining and modeling a task, skill, or strategy, and providing feedback while students write	by engaging students in correcting sample sentences on a daily basis	by explaining a task, skill, or strategy, and giving students an opportunity to practice	by engaging students in shared or interactive writing	A
8	Which of the following sets of words would be best for a teacher to use when providing students with examples of words conforming to the "silent e" phonics generalization?	time, make, cube, done	lake, breathe, raise, fate	brake, use, hope, shine	tree, lie, blue, toe	C
9	As a teacher reads aloud to his students from a social studies text he comments aloud, "This word pioneer is in bold print so that means it is an important word," and "The chapter headings in the book can help me understand the main ideas in the book, so I will be sure to read them." The teacher is helping students improve their comprehension of informational text primarily how?	teaching them how to use graphic organizers	modeling attention to useful features of informational text	improving students' recall of the details of the text	teaching them how to infer word meanings from context	B
10	Two or three times each week Mrs. Hruby teaches "phonics through spelling" with her students. She pronounces words sound-by-sound as her students listen, write the appropriate letters, and then blend the letters to identify the words. Why is this activity likely to be effective?	reinforces students' recognition of common spelling patterns	requires students to use letter-sound relationships to blend unfamiliar words	reviews and strengthens students' ability to recognize and blend word chunks	prepares students to combine letter-sound relationships with meaning-based clues	B

(continued)

Table B1. Teacher Knowledge of Early Literacy Skills survey form A *(continued)*

Item	Question	Choice A	Choice B	Choice C	Choice D	Correct answer
11	Why are there two n's in "running"?	because the base word ends in a single consonant preceded by a single vowel	because the final consonant is always doubled when adding -ing	because the letter u has many different pronunciations	because the consonant n is not well articulated	A
12	Mr. Lewis' class has been learning spelling rules for adding "ing" to base words. He is looking for groups of words that illustrate the various rules to give his students a complex challenge. Which of the following groups of words would be best for this purpose?	hopping, running, sending, getting	hoping, buying, caring, baking	seeing, letting, liking, carrying	all of the word sets are useful for this purpose	C
13	Mrs. Card wants to help her students become good spellers. Which activity should Ms. Card do?	pronounce a word and have students write each sound	display letter cards and have students pronounce the sounds	say each sound of a word and have students say the word	ask students whether pairs of spoken words rhyme	A
14	Why is metacognition important in reading comprehension?	it helps students to monitor their own comprehension	it makes the teacher aware of when the students are experiencing difficulty during reading	it prompts students to create mental images	it causes automatic processing of the text so that students can make meaning of the text	A
15	Teachers often read texts aloud as students follow along before the students try to read the text themselves. Which of the following is the best reason why teachers might do this?	to teach comprehension strategies directly	to model their expert decoding skills to students	to present a challenge to the students to read the text quickly	to demonstrate appropriate phrasing and expression for the text	D
16	What is a reading method that focuses on teaching the application of phonemes to letters called?	phonics	phonemics	orthography	phonetics	A
17	What would the open syllable of the nonsense word "botem" most likely rhyme with?	coat	hot	rah	low	D
18	After reading a story, what should the discussion focus on in order to maximize comprehension?	sequencing the events of the story	the most important parts of the story	the details of the story	the characters in the story	B
19	Which of the following is an example of reading comprehension instruction that helps to promote active construction of meaning?	independent silent reading	doing a think aloud	sounding out difficult words	looking up words in a dictionary	B
20	What is the most important reason that oral segmentation and oral blending activities should be a part of reading instruction in the primary grades?	strengthen students' fluency development through oral practice	help students hear and identify short and long vowel sounds	allow students to hear the mistakes of other students	give students practice with skills they will use in silent reading	D
21	Which word(s) is/are phonetically irregular?	done	give	peach	a and b	D

(continued)

Table B1. Teacher Knowledge of Early Literacy Skills survey form A *(continued)*

Item	Question	Choice A	Choice B	Choice C	Choice D	Correct answer
22	Following her lesson on recognizing diphthongs in words, Mrs. Byrnes wants to provide her students with additional practice. Which type of text should she select to provide the best practice?	predictable text with repetitious phrases	authentic text from children's literature	text with a high percentage of selected decodable words	none of the above	C
23	Mr. Kubota teaches his grade 3 students to decode unfamiliar words by breaking words into parts such as word root, prefix, and/or suffix (e.g., un-imagine-able). Which skill is he teaching?	structural analysis	analyze the meaning of the word parts	syllabication	chunking the word	A
24	What is the difference between sight words and vocabulary words?	sight words are learned through decoding and vocabulary words are not	sight words are learned on sight and vocabulary words are learned by decoding	sight words are related to recognition and vocabulary words are related to meaning	none of the above	C
25	A teacher assigns pairs of students to reread a text aloud to each other three times. What skill will this activity strengthen most effectively?	choral reading	text comprehension	fluency development	automatic word recognition	C
26	How many morphemes are in the word "unhappiness"?	2	3	4	5	B
27	Which phonemic awareness activity would be the most difficult for a student?	blending phonemes into real words	blending onset-rime units into real words	deleting a phoneme and saying the word that remains	segmenting words into phonemes	C
28	Mrs. Newswander begins a writing lesson by creating with the students a web that contains the word "said," surrounded by words like shouted, sulked, and replied. She did this to teach students:	prewriting	drafting	revising	editing	C
29	How many phonemes are in the word "box"?	1	2	3	4	D
30	Decoding skills will benefit a student's understanding of text only if the words he or she decodes are what?	recognized at sight	encountered several times	included in the student's oral vocabulary	also defined by context clues	C
31	Which of the following is a nonsense word that does not follow English spelling patterns?	shease	toyn	squive	clow	B

Source: Authors' modification of items sourced from previous research (Binks-Cantrell et al., 2012; Bos et al., 2001; Carlisle et al., 2009; Carlisle et al., 2011; Cunningham et al., 2004; Mather et al., 2001; Moats & Foorman, 2003; Reutzel et al., 2011; Salinger et al., 2010; Spear-Swerling & Cheesman, 2012).

Table B2. Teacher Knowledge of Early Literacy Skills survey form B

Item	Question	Choice A	Choice B	Choice C	Choice D	Correct answer
1	What is the rule for using a "ck" in spelling?	when the vowel sound is a diphthong	when the vowel sound is short	when the vowel sound is long	all of the above	B
2	Decoding skills will benefit a student's understanding of text only if the words he or she decodes are what?	recognized at sight	encountered several times	included in the student's oral vocabulary	also defined by context clues	C
3	Which word contains a consonant digraph?	flop	bang	sink	box	B
4	Which is a distinguishing characteristic of phonemic awareness instruction?	uses printed letters	uses two cueing systems	does not use printed letters	links meaning to sound	C
5	Which strategy for building students' phonemic awareness is the least likely to support beginning reading skills?	teaching blending and segmenting of phonemes in words	beginning phonemic awareness instruction in preschool	teaching letter sounds in combination with phoneme manipulation	teaching 3 or more types of phoneme manipulation skills at a time	D
6	Mrs. Funke is teaching her students to identify multisyllable words. Which is an appropriate first step for her to do?	model analyzing words for familiar prefixes and suffixes	show students how to blend individual letter-sounds, left-to-right	model how to look for little words in big words	demonstrate sequentially blending onsets and rimes	A
7	What is one reason that teaching students the meanings of a new word's parts (affixes and root words) is useful for vocabulary development?	helps students learn alternate spellings for words	helps students use the new word to understand the sentence	helps students decode multisyllabic new words	helps students comprehend other new words	D
8	How many phonemes are in the word "box"?	1	2	3	4	D
9	Which set of words is decodable?	bed, the, sit	side, some, roam	wash, boil, gave	chap, slew, soft	D
10	What would the word be if you say the word "ice," and then reverse the order of the sounds?	easy	sea	size	sigh	D
11	After reading a story, what should the discussion focus on in order to maximize comprehension?	sequencing the events of the story	the most important parts of the story	the details of the story	the characters in the story	B
12	If "tife" is a word, the letter "i" would probably sound like the "i" in which word?	if	beautiful	find	ceiling	C
13	Which of the following is the most effective instructional strategy for helping students simultaneously strengthen word recognition, fluency, and comprehension?	calling on students one at a time to read aloud from a story	having students read words from a word wall	having students select their own books and read them silently	having students "echo-read" paragraphs that the teacher has read aloud	D
14	Which word is an example of this spelling rule: double the final consonant of a closed syllable that ends in one consonant when adding a suffix beginning with a vowel?	ripple	accommodate	grassy	winning	D

Table B2. Teacher Knowledge of Early Literacy Skills survey form B *(continued)*

Item	Question	Choice A	Choice B	Choice C	Choice D	Correct answer
15	Mrs. Newswander begins a writing lesson by creating with the students a web that contains the word, said, surrounded by words like shouted, sulked, and replied. She did this to teach students:	prewriting	drafting	revising	editing	C
16	Which of the following words has an example of a final stable syllable?	wave	bacon	paddle	napkin	C
17	What can sentence combining help students learn to do?	question the text	correct grammatical errors	form complex sentence structures	analyze word structure	C
18	Which of the following is the best description of reading fluency?	reading fluency is the ability to read grade-appropriate text with good comprehension and a high degree of engagement	reading fluency is the ability to read grade-appropriate text with a high degree of accuracy and comprehension	reading fluency is the ability to read individual words, including both real words and nonsense words, with a high degree of accuracy	reading fluency is the ability to read grade-appropriate text accurately, effortlessly, and with appropriate intonation and expression	D
19	Which of the following is the most effective strategy for teaching new vocabulary words?	direct instruction in varied contexts or subjects, and indirect instruction through use of new words in conversation and topically-related texts	direct instruction in the context or subject during which the words are most often used, and indirect instruction through the use of text with controlled vocabulary	direct instruction only; indirect vocabulary instruction is ineffective	indirect instruction only; direct vocabulary instruction is ineffective	A
20	Why may students confuse the sounds /b/ and /p/ or /f/ and /v/?	students are visually scanning the letters in a way that letters are misperceived	the students can't remember the letter sounds so they are randomly guessing	the speech sounds within each pair are produced in the same place and in the same way, but one is voiced and the other is not	the speech sounds within each pair are both voiced and produced in the back of the mouth	C
21	What is the most important reason that oral segmentation and oral blending activities should be a part of reading instruction in the primary grades?	strengthen students' fluency development through oral practice	help students hear and identify short and long vowel sounds	allow students to hear the mistakes of other students	give students practice with skills they will use in silent reading	D
22	Mrs. Ellefsen is determined to increase her students' blending abilities. She has them sort words according to spelling patterns and they are doing well. What else could Mrs. Ellefsen do to increase her students' blending abilities?	have students read widely from easy texts	segment words orally for students to write the word spellings from dictation	model a word spelling strategy for students (e.g., see the word, spell the word, write the word)	explicitly teach students how to blend sounds to pronounce words	D

(continued)

Table B2. Teacher Knowledge of Early Literacy Skills survey form B *(continued)*

Item	Question	Choice A	Choice B	Choice C	Choice D	Correct answer
23	Which of the following is a noun phrase?	wrote the word	beside the stream	an ill-conceived idea	before entering the house	C
24	Which of these would be the final step a teacher would use in an instructional sequence designed to increase students' ability to make inferences about what they read independently?	teacher modeling	student guided practice	student application	teacher direct explanation	C
25	Mr. Kubota teaches his grade 3 students to decode unfamiliar words by breaking words into parts such as word root, prefix, and/or suffix (e.g., un-imagine-able). Which skill is he teaching?	structural analysis	analyze the meaning of the word parts	syllabication	chunking the word	A
26	Mr. Willard is planning a repeated reading activity to strengthen his students' fluency skills. Which of the following reading materials would be most effective for the activity?	a list of words from the social studies textbook	a list of high-frequency words from a teaching manual	two paragraphs from a grade-level text	two paragraphs from a grade 4 level text	C
27	Which of the following words contains a schwa sound?	cotton	phoneme	stopping	preview	A
28	If a student can read a list of words very rapidly and accurately without having to consciously decode, what is the student demonstrating?	comprehension	metacognition	automaticity	vocabulary skills	C
29	Fluency serves as a bridge between which two processes?	word recognition and comprehension	comprehension and vocabulary	phonological awareness and comprehension	word recognition and vocabulary	A
30	Mrs. Jackson's students need to improve their fluency skills. Which of the following activities should she include in her lesson plans for the 90-minute reading period?	students will repeatedly read a text in pairs for 20 minutes	students will read a 20-word list repeatedly until they can read it in 10 seconds	students will repeatedly read a text silently for 50 minutes	students will read along with books on tape for the entire 90 minutes	A
31	What does automaticity in reading refer to?	process complex information with little effort or attention	understand the meaning of the word upon seeing it in text	use the next step in a series of steps that have been memorized	apply an effective comprehension strategy when needed	A

Source: Authors' modification of items sourced from previous research (Binks-Cantrell et al., 2012; Bos et al., 2001; Carlisle et al., 2009; Carlisle et al., 2011; Cunningham et al., 2004; Mather et al., 2001; Moats & Foorman, 2003; Reutzel et al., 2011; Salinger et al., 2010; Spear-Swerling & Cheesman, 2012).

Table B3. Raw-to-scale score conversions for the Teacher Knowledge of Early Literacy Skills survey

Teacher Knowledge of Early Literacy Skills survey form A			Teacher Knowledge of Early Literacy Skills survey form B		
Raw score	Theta equivalent	T-score	Raw score	Theta equivalent	T score
0	—	20	0	—	20
1	−4.033	20	1	−4.235	20
2	−3.257	20	2	−3.409	20
3	−2.839	22	3	−2.956	20
4	−2.546	25	4	−2.639	24
5	−2.311	27	5	−2.388	26
6	−2.107	29	6	−2.173	28
7	−1.920	31	7	−1.979	30
8	−1.740	33	8	−1.796	32
9	−1.563	34	9	−1.621	34
10	−1.384	36	10	−1.449	36
11	−1.201	38	11	−1.278	37
12	−1.012	40	12	−1.105	39
13	−0.817	42	13	−0.928	41
14	−0.618	44	14	−0.745	43
15	−0.417	46	15	−0.552	44
16	−0.212	48	16	−0.347	47
17	−0.002	50	17	−0.128	49
18	0.214	52	18	0.104	51
19	0.440	54	19	0.353	54
20	0.678	57	20	0.618	56
21	0.932	59	21	0.900	59
22	1.203	62	22	1.200	62
23	1.496	65	23	1.516	65
24	1.815	68	24	1.851	69
25	2.166	72	25	2.212	72
26	2.562	76	26	2.614	76
27	3.021	80	27	3.080	80
28	3.583	80	28	3.655	80
29	4.340	80	29	4.430	80
30	5.582	80	30	5.697	80
31	—	80	31	—	80

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

Appendix C. Coach's Classroom Observation Tool

This appendix includes the Coach's Classroom Observation Tool (CCOT), which Regional Educational Laboratory (REL) Southeast created by combining aspects of the Timed Observations/Student Engagement and Teaching Competency Checklist (Foorman et al., 2004; Foorman & Schatschneider, 2003; Foorman et al. (2006) and the Instructional Content Emphasis-Revised (Edmonds & Briggs, 2003). REL Southeast provided Mississippi state literacy coaches with a manual and training on the CCOT before they used this tool (see appendix A for details). The first section of this appendix includes the CCOT record form, and the second section includes the definitions for CCOT. The full manual that accompanies the CCOT is available on request from the first author.

Coach's Classroom Observation Tool record form

Observation Record Form					
Observation start time:	Observation end time:		Total number of minutes (auto calculated):		
Activity start time	Summary of activity	Instructional category (see manual)	Grouping (WC, SG, P, I, D)	Student engagement (H–3, M–2, L–1)	Quality (E–4, HA–3, LA–2, W–1)

Grouping codes: WC is whole class, SG is small group, P is pairs, I is independent, and D is individualized 1:1 teacher instruction.

Student engagement codes: H (3) is high, M (2) is medium, and L (1) is low.

Quality codes: E (4) is excellent, HA (3) is high average, LA (2) is low average, and W (1) is weak.

GLOBAL CLASSROOM RATINGS AND CHECKLIST OF TEACHING COMPETENCIES

Timed observation

What proportion of time...	>50%	>20%	>10%	<10%	0%	n/a
Was an instructional aide (or other classroom assistant) present in the classroom? (check the appropriate box)						
Was instruction conducted in a language other than English? (check the appropriate box)						

What proportion of time was spent in the following groupings:	These values are automatically calculated based on the data in the observation record form
Whole class	
Small group	
Pairs	
Independent	
Individualized/differentiated	

What proportion of time was spent on the following content:	These values are automatically calculated based on the data in the observation record form
Concepts of print/book awareness/conventions	
Phonemic/phonological awareness	
Alphabetic letter recognition and reproduction	
Alphabetic instruction/grapho-phonemic correspondences	
Word work/study/phonics (with text)	
Structural analysis/morphology	
Spelling	
Spelling in the context of reading	
Writing, composition/more than single words	
Grammar/capitalization/punctuation/mechanics	
Students reading their own writing	
Oral language	
Vocabulary	
Fluency	
Reading text/books beyond the word level	
Preview to prepare for reading	
Reading comprehension (during or after reading)	
Giving directions/passing out materials/referring to Teacher Edition (relevant to the reading instruction)	
Feedback—in response to reading relevant comment/event	
Productive non-reading instruction	
Unproductive non-reading instruction	

Based on what you observed today, how would you rate the quality of instruction in:	4—Excellent	3—High average	2—Low average	1—Weak	n/a
	These values are automatically calculated based on the data in the observation record form				
Concepts of print/book awareness/conventions					
Phonemic/phonological awareness					
Alphabetic letter recognition and reproduction					
Alphabetic instruction/grapho-phonemic correspondences					
Word work/study/phonics (with text)					
Structural analysis/morphology					
Spelling					
Spelling in the context of reading					
Writing, composition/more than single words					
Grammar/capitalization/punctuation/mechanics					
Students reading their own writing					
Oral language					
Vocabulary					
Fluency					
Reading text/books beyond the word level					
Preview to prepare for reading					
Reading comprehension (during or after reading)					

Student engagement

Based on what you observed today, how would you rate student engagement during instruction in:	3—High	2—Medium	1—Low	n/a
	These values are automatically calculated based on the data in the observation record form			
Concepts of print/book awareness/conventions				
Phonemic/phonological awareness				
Alphabetic letter recognition and reproduction				
Alphabetic instruction/grapho-phonemic correspondences				
Word work/study/phonics (with text)				
Structural analysis/morphology				
Spelling				
Spelling in the context of reading				
Writing, composition/more than single words				
Grammar/capitalization/punctuation/mechanics				
Students reading their own writing				
Oral language				
Vocabulary				

Based on what you observed today, how would you rate student engagement during instruction in:	3—High	2—Medium	1—Low	n/a
	These values are automatically calculated based on the data in the observation record form			
Fluency				
Reading text/books beyond the word level				
Preview to prepare for reading				
Reading comprehension (during or after reading)				
Overall classroom rating				

Based on what you observed today, how would rate the:	Circle one			
Overall quality of early literacy instruction?	1—Weak	2—Low average	3—High average	4—Excellent
Overall level of student engagement during early literacy instruction?	1—Low	2—Medium	3—High	

Checklist of teaching competencies						
Based on what you observed today, identify how often the teacher displayed the following teaching competencies:	**Check the appropriate box**					
Planning	**All of the time**	**More than half of the time**	**Half of the time**	**Less than half of the time**	**Never**	**No opportunity to observe**
The lesson sequence is followed appropriately						
The teacher plans effectively for instruction, has read the teacher's guide, and uses appropriate and effective examples, anecdotes, analogies, and well-chosen books and materials)						
The teacher seems to be organized and has all the materials necessary for instruction easily accessible						
Management						
The teacher maintains a classroom environment that minimizes distractions and is appropriate for learning						
The teacher has well established instructional routines						
The teacher maximizes the amount of time available for instruction (brief transitions, appropriate timing and pacing of teaching)						
The teacher has clearly stated classroom rules and procedures and communicates expectations about classroom behavior						
The teacher manages student behavior effectively in order to avoid disruptions and to provide productive learning opportunities						
Instruction						
The teacher presents and delivers the lesson effectively (using eye contact, variation of voice tone, animation)						
The teacher provides many and equal opportunities for students to participate						
The teacher emphasizes appropriate content						
The teacher provides sufficient practice						
The teacher provides more practice than necessary						
The teacher models thinking and learning						
The teacher is aware of lesson objectives						
The teacher draws upon students' home and neighborhood experiences						
The teacher draws upon students' home country, beliefs, experiences, and languages						
Monitoring of student learning						
The teacher records student progress efficiently						
The teacher uses the data in order to make judgments and decisions about student performance						
The teacher monitors student responses effectively						
The teacher provides clear, direct, and frequent feedback to the students						
The feedback provided by the teacher is appropriate to the lesson objective						
The teacher provides feedback in a positive manner						
The teacher assigns tasks that are relevant to instructional goals and objectives						
The teacher corrects students' errors to extend instruction						

Checklist of teaching competencies						
Based on what you observed today, identify how often the teacher displayed the following teaching competencies:		Check the appropriate box				
Personal characteristics	All of the time	More than half of the time	Half of the time	Less than half of the time	Never	No opportunity to observe
The teacher is knowledgeable about how children learn to read						
The teacher is enthusiastic about the instructional approach used						
The teacher is aware of the students' level of ability and skill development and plans the instruction appropriately						
The teacher has high but realistic expectations regarding students' learning and their progress						
The teacher is generally motivated and keeps students actively involved by maintaining an enthusiastic learning atmosphere						

Coach's Classroom Observation Tool definitions

COACH'S CLASSROOM OBSERVATION TOOL DEFINITIONS AND INDICATORS
Adapted from: Edmonds & Briggs (2003), Foorman & Schatschneider (2003), and Foorman et al. (2006)
This codebook is intended to give the functional definitions and examples for each of the codes to be used in the Observation Record Form and Global Classroom Ratings. Use this codebook carefully when coding the notes in your observation. When in doubt on a rating, if you are struggling between a higher and lower rating, go with the lower rating; you are struggling because clearly something has not been met with the higher rating, so do not give the higher rating even if the lower rating seems unfair.

Main instructional category

For each instructional event, determine the domain of the main instructional content. Determine the broad category the instruction reflects within the following:

Code	Description
Concepts of print, book awareness, conventions	The teacher is acquainting the children with print. This includes things like conventions of print use and the format of a book, including title page, author, reading left to right, table of contents, genre of book, etc. Teachers may emphasize how print works: moving top to bottom, left to right, recognizing that printed words are separated by spaces, recognizing upper and lower case letters in isolation or in printed words, and recognizing the significance of capitalization, punctuation marks (periods, exclamation marks, quotation marks), and the location of punctuation marks. Intent is not comprehension, but familiarity with features of text. This can be taught during read-alouds.
Phonemic or phonological awareness	This refers to all instruction that is targeted at directing the children's attention to the sounds in language, and/or manipulating these sounds. These activities DO NOT include explicit reference to printed text. The student has the ability to recognize the sounds in spoken language and how they can be segmented (pulled apart), blended (put back together), and manipulated (added, deleted, and substituted). Instruction is characterized by the absence of print and is based on spoken language. Note, in some cases text may be presented but the teacher is not providing explicit instruction about the link between the sounds and the text. Activities that fall in this category may include (but are not limited to):
	Rhyming activities—The teacher asks children to produce or identify rhyming words. Focus is on the sounds rather than the meaning of language. The teacher may also discuss what rhyming is. Nursery rhymes presented without explicit reference to the text are included here.
	Phoneme segmentation—The teacher directs children to analyze a word by breaking apart the phonemes in the word into discrete segments. For example, breaking the word cat into the sounds /c/ /a/ /t/. This can be done with or without counters of some kind (such as Elkonin boxes).
	Syllable segmentation—The teacher directs children to analyze a word by breaking apart the syllables. This might include clapping, tapping, snapping, jumping, etc. with each syllable of a word.
	Working with initial, middle, or final phoneme—The teacher asks children to identify or say the first, middle, or final phoneme heard in a designated word. For example, the first phoneme in cat is /c/, the last phoneme is /t/. The teacher might also ask students to generate words that begin or end with a certain sound.
	Phoneme deletion/substitution—For deletion, the teacher directs children to say a word and then drop out a particular phoneme in the word (the initial, final, or some medial sounds) and then say the remaining sounds within the word. For example, cat without the /c/ is /at/. For substitution, the teacher asks children to replace a particular phoneme with another phoneme. For example, change the /c/ in cat to /m/ to make mat.
	Blending sounds—The teacher directs children to synthesize separate phonemes to form combinations of sounds or recognizable words without text. This may involve synthesizing two final phonemes, then adding an initial phoneme to these two sounds, or extending the initial phoneme so that it "runs together" with the next phoneme, and so on to the end of the word. For example, a teacher may break a word into two parts /c/ — /at/ and then ask students what the word is (cat).
Alphabetic letter recognition and reproduction	Instruction to teach children the alphabet. This includes reciting the alphabet, writing letters, and recognizing letters. The focus here is NOT on sounds. Activities that fall in this category may include (but are not limited to):
	Reciting/singing the alphabet—The teacher leads children in singing the alphabet song, or in naming all of the letters in the alphabet while showing the written letters.
	Identification/practice of letter names—The teacher identifies a written letter by name, or asks students to identify a written letter by name.
	Letter formation/letter shapes—The teacher demonstrates how to form a letter and asks children to practice. This can be done in many ways including finger in the air, or pencil on paper.
	Alphabetizing/alphabetical order—The teacher works on the concept of alphabetical order by focusing on the order and position of the letters of the alphabet in relation to each other.

COACH'S CLASSROOM OBSERVATION TOOL DEFINITIONS AND INDICATORS	
Alphabetic instruction, grapho-phonemic correspondences	Any instruction that is intended to help students understand that written letters of the alphabet are used to represent the sounds in words (grapho-phonemic correspondences). It includes instruction in which letters make which sounds. Activities that fall in this category may include (but are not limited to): *Rules for letter/sound correspondence*—The teacher is teaching about rules, helping children learn the combinations of letters and the sounds they make. *Vowels/consonants*—Identifying vowels and consonants and discussing their use in writing/speech. *Blends/digraphs*—The teacher works with consonant blends (cl, st, bl, fr, etc.); consonant digraphs, which are two consonants together that make one sound (/sh/, /ch/, /th/, etc.); or vowel digraphs (/ea/, /oa/, /ou/, etc.).
Word work/study phonics (with text)	This refers to any instruction with text that works purely at the word level. This can include words on the board, on a word wall, or in a book or other print medium. This can be with connected text where the teacher has students blend and segment words while in the process of reading a text. This can also occur without connected text such as when a teacher has students working with words in isolation such as on the board. Activities that fall in this category may include (but are not limited to): *Blending/sounding out/reading words*—The teacher presents words and asks children to blend the sounds represented by the letters to form a word. *Sight words/outlaw words*—The teacher presents words that cannot be blended; in other words, they do not follow the usual rules of pronunciation. *Segmentation (syllable or phoneme)*—The teacher works with students to break written words down into syllables or phonemes. *Rhyming (with text)*—The teacher uses words in print to demonstrate rhyming words.
Structural analysis morphology	Instruction about meaningful parts (morphemic units) of words. Examples: activities involving plurals, possessives, prefixes, suffixes, verb tenses, root words, derivations, and etymology. Activities that fall in this category may include (but are not limited to): *Morphology*—The teacher discusses meaningful parts of words, like suffixes, which always convey the same meaning or parts of the word which convey the same meaning. *Etymology/derivation*—The teacher focuses on the origins or derivation of a word by analyzing the elements of the word.
Spelling	Spelling is the specific focus. This is a lesson, or segment of the lesson, specifically focused on spelling. Often a separate speller or lists of words will be used. Examples include grapho-phonemic correspondences as the instructional unit (short vowels, long vowels, consonant blends, diagraphs, and vowel teams) or sight/outlaw words (teacher illustrates words which do not follow phonetic spelling rules and are not easily sounded out).
Spelling in the context of reading	Refers to spelling covered in the context of reading. This does NOT include an extended spelling activity. This usually occurs in passing where a teacher might ask students to spell a word that was just read.
Writing, composition more than single words	These are activities in which children are asked to create stories, journal entries, recipes, essays, etc. and express these ideas by writing them on paper. Activities that fall in this category may include (but are not limited to): *Student composition*—These are activities in which children are asked to create original stories, journal entries, recipes, essays, etc. and express these ideas by writing them on paper. *Journal writing*—The teacher asks children to write down their thoughts in a daily journal, or the teacher may be more direct and give the children a topic to write about in their journals. *Student dictation*—The teacher says words or sentences aloud and students write them. *Teacher-led instruction*—The teacher is modeling/instructing students in writing forms/processes, i.e., the "how to composition." *Sentences (copying and composing)*—The teacher writes words or sentences and asks students to write them, or students work on a worksheet, guided by teacher, and write words and sentences. *Handwriting instruction*—Students practice the proper formation of letters; focus is on correctness of formation and not on identification, and could involve cursive handwriting. Practice or instruction in the proper size, spacing, posture, and strokes of letters. *Copying*—students copy letters, words, or text from a printed stimulus for the purpose of recording the information.
Grammar, capitalization, punctuation, or mechanics	Formal instruction in which the teacher focuses on and gives students opportunities to practice a grammatical concept or a point about mechanics. In kindergarten, this is different from print awareness which is instruction that makes the student aware of what appears in print and how print works. It is a focused lesson or segment of a lesson on grammar or mechanics.

COACH'S CLASSROOM OBSERVATION TOOL DEFINITIONS AND INDICATORS	
Students reading their own writing	Students are involved in peer/self-editing or reading their writing to the class, group, or an individual.
Oral language	The focus is on listening and speaking to communicate meaning. Discussion is academic and not logistic or disciplinary. The teacher and students engage in discussions about words, books, songs, or relevant topics. Activities that fall in this category may include (but are not limited to): *Sharing/discussion*—(Usually without a written stimulus. In kindergarten, there may be a written stimulus such as a calendar, picture cards, etc.). The teacher is leading an activity that is intended to develop students' verbal skills. This includes a variety of activities that help to develop oral skills in the area of vocabulary, grammar, and syntax. *Listening comprehension*—The teacher is focused on the children's understanding of the text they are hearing. Discussion may include structural details (e.g., plot, characters, etc.) and knowledge related to what was read. This is all oral work. Children are not reading the text. *Syntax*—The teacher is focused on the children's understanding of how words and phrases are/should be arranged to create well-formed sentences.
Vocabulary	The teacher focuses on developing the children's knowledge of the meaning of words within the context of teaching reading as the vocabulary is related to print. Activities that fall in this category may include (but are not limited to): *Definitions*—The teacher or student presents the meaning of the word, either directly or by using it in a sentence. *Antonyms/synonyms*—Students/teacher uses antonyms/synonyms to define a word. *Application/use in context*—Students use the words in a sentence.
Fluency	Students read aloud to develop speed, accuracy, or intonation. The intent is improving how quickly and accurately students read words; the intent is not necessarily understanding what is read. Reading aloud is not necessarily fluency. Activities that fall in this category may include (but are not limited to): *Letter or sound naming fluency*—Students name letters or sounds presented in list format (or on flashcards) for the purpose of developing speed and accuracy. *Word fluency*—Students read a list of words for the purpose of developing speed and accuracy (can be presented on flashcards). *Repeated reading of text*—Students engage in repeated reading either with the class, in a small group, or one on one for the purpose of developing speed, accuracy, and/or intonation. Students may echo read with a partner or teacher. Students engage in partner reading with the purpose of developing speed and accuracy (partners should be re-reading a text previously read or reading the text multiple times). *Other*—Students listen to books read aloud with the intent of modeling speed, accuracy, and intonation (could be teacher, computer, or books on tape). Students engage in silent reading with the stated purpose of developing speed or accuracy. This may also include incidental instruction (e.g., comments made by teacher during reading about reading with more speed).
Reading text and/or books beyond the word level	This includes reading or being read to from all types of text. Students may or may not see the text. The teacher may read aloud (without students reading aloud), students may read aloud (with or without teacher), students may read silently, or students may sing/chant a known pattern or song related to the text. This may include supported oral reading, choral reading, independent silent reading, independent oral reading, or listening to books on tape or on a computer.
Preview to prepare for reading	This refers to all discussions of topics/issues related to something the teacher/students will read or have read before that was related to the upcoming reading. For example, if teacher/students read a book about going to the beach, the teacher might ask who has gone to the beach. It differs from the discussions for oral language development because the emphasis is on knowledge that is linked to what is going to be read, or has been read. This also includes picture walking and previewing a book through illustrations. This may also include prior knowledge or predicting where students preview the materials before reading, predicting outcomes based on prior knowledge, or participating in activities designed to measure their level of knowledge before reading a book.

COACH'S CLASSROOM OBSERVATION TOOL DEFINITIONS AND INDICATORS	
Reading comprehension (during or after reading)	The teacher is focused on the children's understanding of the text they are reading. This may include students writing responses to comprehension questions. Activities that fall in this category may include (but are not limited to):
	Strategy instruction—May include any of the following: structural details (character, plot, etc.) and things like summarization, predicting, main idea, etc., direct instruction on comprehension strategies such as story mapping, discussing knowledge related to what was read, or students drawing a picture about what was read or creating a craft related to what was read (e.g., reading about frogs and then making a frog puppet, etc.). This does not include listening comprehension. Additional examples include using comprehension strategies to check decoding accuracy, structural elements, process, predictable text, and teaching comprehension strategies for mastering structure and process.
	Literal questions—The answers to this type of question are right there in the book. This type of question DOES NOT require the students to make an inference.
	Inferential questions—The answers to this type of question require the students to make an inference. The answer is NOT right there in the book.
	Vocabulary—Students have the opportunity to develop their print or oral vocabulary in the context of reading or discussion (e.g., the teacher asks or tells what a word encountered in reading means; identifying things such as body parts, colors, days).
	Reading comprehension monitoring—Monitoring may occur during or after reading. Students learn to be aware of their understanding of text. Tends to be discussion-oriented with little focus on a product or goal. The teacher and students summarize the story as the intent of discussion or activities.
	Listening comprehension monitoring—Monitoring of comprehension occurs during or after reading done by the teacher or other students.
Giving directions, passing out materials, referring to Teacher Edition	This activity is relevant to reading instruction. There is an instructional purpose in giving the directions.
Feedback	This code can only be given if the feedback is specific to a reading relevant comment/activity. This includes corrective feedback or praise.
Productive non-reading instruction	The teacher is not interacting with any students; however, students are still engaged in educational tasks/ activities. The teacher may be checking papers, doing clerical work, speaking to a parent/classroom visitor, etc.
Unproductive non-reading instruction	The teacher is not instructing and students are not engaged in educational tasks/activities. This also includes punitive feedback where a teacher responds negatively to students' work or responses. Examples include chaos in the classroom/disruptive student behavior, other/non-reading activities such as transitioning between activities, discipline, and classroom interruptions.
Grouping	
For each instructional event, select a grouping code that best describes the grouping pattern of students associated with this instructional event. Instructional activities are coded into one of five grouping patterns. Code only formal structures arranged by the teacher, not informal or incidental grouping. Although the seating arrangement of the classroom may be affected by group activities, this item relates to student interaction in a group, not seating arrangement.	
Whole class	The entire class is involved in the same activity or assignment.
Small groups	Class is working in 2 or more groups, with 3 or more students per group. It could be teacher working with a group of 2 or more students.
Pairs	Class is working in groups of 2. One child acts as a peer tutor to another student. Most of the students are working in pairs. Students are in groups of two to share notes, tutor, or work on an assignment/activity.
Independent	Students are engaged individually in an activity/assignment like others in the class (help-seeking behaviors may be observed between students but they are not working in a group).
Differentiated, Individual	Students work on differentiated assignments. Students are not involved in pairing or group activities and are working individually on differentiated assignments. The teacher works individually with a student for 5 minutes or more.

COACH'S CLASSROOM OBSERVATION TOOL DEFINITIONS AND INDICATORS

Student engagement

For each instructional event, rate the level of student engagement. Students following along, but not necessarily vocally participating, are considered engaged. Base your rating on the majority of students in the instructional event. For example, if a few students were off task, but generally all students were actively engaged, code the instructional event as high engagement.

1. Low engagement	More than half staring out the window, engaging in idle chatter, fiddling with materials, inappropriately moving about the classroom
2. Medium engagement	Most students are actively engaged in learning activity (reading, writing, listening, talking about a relevant topic)
3. High engagement	Almost all students are actively engaged in learning activity (reading, writing, listening, talking about a relevant topic)

Quality

For each instructional event observed, rate the quality of instruction. Select the rating that best describes the type of instruction observed. Only observed indicators should be considered when assigning the rating. One way to determine instructional quality is to look for the rating under which a supermajority of indicators applies to the observed instruction. Use the following guidelines for assigning quality indicators for each instructional event or activity.

1. The majority determines the quality rating
   a. Rating should be based on observable behavior using professional judgment, not inferences.
   b. The framework for thinking about teacher quality is based on the assumption that a teacher who falls into the Excellent category is one who addresses the needs of a struggling reader.
   c. A rating of High average, Low average, or Weak represents the degree to which a teacher deviates from this standard. For example, a teacher who is rated Low Average may be an effective teacher for most students, but is not addressing the needs of struggling readers.
2. Assignment of Low average or High average
   a. Low average: Some indicators under Weak are present, but the majority fall under Average.
   b. High average: Some indicators under Excellent are present, but the majority fall under Average.
   c. Special consideration: If a teacher meets a majority (5) of indicators under Weak and all others under Excellent, the teacher's rating would be Low Average for that event.
3. Assignment of Weak or Excellent
   a. To clearly assign either of these extreme ratings, almost all (or supermajority) of indicators must fall within the Excellent or Weak range.
   b. Distinguish between Excellent and High average by considering how closely the teacher meets the needs of a struggling reader.
4. Situation: All indicators fall within Average column
   a. Professional judgment should be used to determine whether to rate as Low or High average.
   b. Remember to keep the struggling reader in mind.
   c. If the teacher has farther to go to meet the needs of the struggling reader, rate as Low average.

4—Excellent	3—High average	2—Low average	1—Weak
Uses language that is direct and explicit.	Inconsistently uses language that is direct and explicit.		Uses language that is indirect and implicit.
Models many examples.	Provides some examples.		Provides no models or demonstrations.
Provides sufficient and varied opportunities for practice.	Provides many opportunities for practice with little variation. Practice opportunities do not seem to be based on student need.		Provides insufficient opportunities for practice with no variation.
Provides immediate and corrective and descriptive feedback.	Provides inconsistent feedback.		Provides little feedback that is nonspecific or no feedback.
Adjusts time to meet student needs.	Uses time appropriately, but use does not seem based on student need, yet still seems adequate for given activity.		Demonstrates poor use of time that is not differentiated and unrelated to student need or task difficulty.
Constantly monitors student performance.	Monitors some students or monitors all students from some activities.		Demonstrates lack of monitoring or monitoring very few students.
Encourages high student engagement and time on task.	Encourages student engagement and time on task varies.		Does not encourage student engagement and time on task.
Scaffolds tasks and materials to meet student needs.	Uses scaffolding inconsistently and does not always tailor it to student needs.		Scaffolds inappropriately or insufficiently.
Uses appropriate pacing, including wait time.			Demonstrates poor pacing, either too slow or too fast with no wait time provided.

Appendix D. Supplemental tables of teacher knowledge scores and instructional practices ratings

This appendix provides descriptive statistics for all variables included in this report. Table D1 provides means and standard deviations at each administration of the Teacher Knowledge of Early Literacy Skills survey by professional development program completion, and table D2 provides means and standard deviations at each observation window of the Coach's Classroom Observation Tool by professional development program completion.

Table D1. Teacher Knowledge of Early Literacy Skills scores by administration window and educators' progress in the professional development program

Observation point and progress in the professional development program	Number	Percent	Teacher Knowledge of Early Literacy Skills score	
			Mean	Standard deviation
Spring 2014 (*n* = 2,421)				
Not started	2,166	89.5	50.3	12.0
In progress	250	10.3	52.7	12.4
Completed	5	0.2	48.0	13.2
Fall 2014 (*n* = 4,683)				
Not started	3,506	74.9	49.4	13.6
In progress	525	11.2	49.3	14.1
Completed	652	13.9	53.3	14.4
Spring 2015 (*n* = 1,788)				
Not started	957	53.5	51.3	14.0
In progress	332	18.6	53.5	13.7
Completed	499	27.9	55.7	14.8
Fall 2015 (*n* = 3,192)				
Not started	1,641	51.4	51.1	14.2
In progress	633	19.8	53.8	14.5
Completed	918	28.8	55.5	14.6

Source: Authors' analysis of data from Mississippi Department of Education (2015b).

Table D2. Coach's Classroom Observation Tool instructional ratings by observation window and progress in the professional development program

Observation point and progress in the professional development program	Number	Percent[a]	Quality of instruction		Student engagement		Teaching competencies	
			Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Winter 2014 (*n* = 121)								
Not started	43	35.5	2.0	0.6	2.0	0.5	3.0	0.7
In progress	72	59.5	2.4	0.6	2.3	0.5	3.5	0.7
Completed	6	5.0	2.4	0.6	2.5	0.4	4.0	0.7
Spring 2014 (*n* = 124)								
Not started	13	10.5	2.3	0.8	2.2	0.5	3.0	0.7
In progress	83	66.9	2.4	0.7	2.3	0.6	3.6	0.8
Completed	28	22.6	2.5	0.7	2.4	0.5	3.4	0.9
Fall 2014 (*n* = 165)								
Not started	10	6.1	2.5	0.6	2.2	0.5	3.5	1.0
In progress	88	53.3	2.4	0.6	2.2	0.6	3.5	0.8
Completed	67	40.6	2.6	0.6	2.3	0.6	3.9	0.9
Spring 2015 (*n* = 151)								
Not started	3	2.0	2.0	0.2	1.9	0.2	2.9	0.5
In progress	35	23.2	2.5	0.7	2.4	0.6	3.8	0.8
Completed	113	74.8	2.7	0.6	2.5	0.5	3.9	0.8

Source: Authors' analysis of from Mississippi Department of Education (2015a).

Notes

1. Throughout the report "educator" refers to the inclusive group of classroom teachers, literacy coaches, and school administrators; "teacher" refers exclusively to grades K–3 classroom teachers. "Teacher knowledge" refers to the construct of teacher knowledge of early literacy skills.
2. The correlation between the number of surveys a respondent completed and the mean TKELS score across surveys was small ($r = .15$).
3. The Schools and Staffing Survey provides descriptive data on the context of elementary and secondary education using stratified probability sample design. The most recent relevant data were collected in 2011/12.
4. For each school served, and for each grade served, the coaches wrote teachers' names on cards. In the presence of the teachers, the coach blindly drew one card for each grade out of a bucket. The coach then arranged a time to observe the selected teacher.
5. To maintain confidentiality, teachers reported only district names, not school names. Therefore, nesting could not account for schools, only districts.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1–7. Retrieved November 4, 2015, from http://CRAN.R-project.org/package=lme4.

Binks-Cantrell, E., Joshi, R. M., & Washburn, E. K. (2012). Validation of an instrument for assessing teacher knowledge of basic language constructs of literacy. *Annals of Dyslexia*, *62*(3), 153–171. http://eric.ed.gov/?id=EJ978830

Bos, C., Mather, N., Dickson, S., Podhajski, B., & Chard, D. (2001). Perceptions and knowledge of preservice and inservice educators about early reading instruction. *Annals of Dyslexia, 51*(1), 97–120. http://eric.ed.gov/?id=EJ637272

Brady, S., Gillis, M., Smith, T., Lavalette, M., Liss-Bronstein, L., Lowe, E., et al. (2009). First grade teachers' knowledge of phonological awareness and code concepts: Examining gains from an intensive form of professional development and corresponding teacher attitudes. *Reading and Writing*, *22*(4), 425–455. http://eric.ed.gov/?id=EJ830780

Carlisle, J. F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading and Writing, 22*(4), 457–486. http://eric.ed.gov/?id=EJ830779

Carlisle, J. F., Kelcey, B., Rowan, B., & Phelps, G. (2011). Teachers' knowledge about early reading: Effects on students' gains in reading achievement. *Journal of Research on Educational Effectiveness, 4*(4), 289–321. http://eric.ed.gov/?id=ED543664

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Stanovich, P. J. (2004). Disciplinary knowledge of K–3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia, 54*(1), 139–167. http://eric.ed.gov/?id=EJ683414

Cunningham, A. E., Zibulsky, J., & Callahan, M. D. (2009). Starting small: Building preschool knowledge that supports early literacy development. *Reading and Writing: An Interdisciplinary Journal, 22*(4), 487–510. http://eric.ed.gov/?id=EJ830783

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus, 2*(222).

Edmonds, M. S., & Briggs, K. L. (2003). Instructional Content Emphasis Instrument. In S. R. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for observing teaching and learning* (pp. 31–52). Baltimore, MD: Brookes.

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., et al. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE No. 4008). National Center for Education Evaluation and Regional Assistance Working Paper. Washington, DC: U.S. Department of Education. http://eric.ed.gov/?id=ED566956

Foorman, B. R., Goldenberg, C., Carlson, C., Saunders, W., & Pollard-Durodola, S. D. (2004). How teachers allocate time during literacy instruction in primary-grade

English language learner classrooms. In M. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 289–328). Baltimore, MD: Brookes.

Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practices during reading/language arts instruction and its relationship to student achievement. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for observing teaching and learning* (pp. 1–30). Baltimore, MD: Brookes.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology, 31*(1), 1–29. http://eric.ed.gov/?id=EJ724459

Karahalios, A., Baglietto, L., Lee, K. J., English, D. R., Carlin, J. B., & Simpson, J. A. (2013). The impact of missing data on analyses of a time-dependent exposure in a longitudinal cohort: A simulation study. *Emerging Themes in Epidemiology, 10*(6), 1–11.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices*, 2nd ed. New York, NY: Springer-Verlag.

Mather, N., Bos, C., & Babur, N. (2001). Perceptions and knowledge of preservice and inservice teachers about early literacy instruction. *Journal of Learning Disabilities, 34*(5), 472–482. http://eric.ed.gov/?id= EJ632661

McCoach, D. B., & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology, 1*(17), 1–12.

Mississippi Department of Education. (2013). *Cambium Learning's proposal for Mississippi Department of Education RFP for Mississippi's K-3 Literacy Professional Development Training System*. Jackson, MS: Author.

Mississippi Department of Education. (2014). *2013–2014 Public school district K-03 grade teachers*. Jackson, MS: Author.

Mississippi Department of Education. (2015a). *Coaches classroom observation tool database*. Jackson, MS: Author.

Mississippi Department of Education. (2015b). *Teacher knowledge of early literacy skills database*. Jackson, MS: Author.

Moats, L. C., & Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Annals of Dyslexia, 53*(1), 23–45. http://eric.ed.gov/?id= EJ679548

Moats, L. C., & Tolman, C. (2009). *Language essentials for teachers of reading and spelling*, 2nd ed. Longmont, CO: Sopris West.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide*, 7th ed. Los Angeles, CA: Muthén & Muthén.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods,* 2nd ed. Thousand Oaks, CA: Sage.

Reutzel, D. R., Dole, J. A., Read, S., Fawson, P., Herman, K., Jones, C. D., et al. (2011). Conceptually and methodologically vexing issues in teacher knowledge assessment. *Reading & Writing Quarterly, 27*(3), 183–211. http://eric.ed.gov/?id= EJ923015

Salinger, T., Mueller, L., Song, M., Jin, Y., Zmach, C., Toplitz, M., et al. (2010). *Study of teacher preparation in early reading instruction* (NCEE No. 2010–4036). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://eric.ed.gov/?id= ED512150

Spear-Swerling, L., & Cheesman, E. (2012). Teachers' knowledge base for implementing response-to-intervention models in reading. *Reading & Writing: An Interdisciplinary Journal, 25*(7), 1691–1723. http://eric.ed.gov/?id=EJ973090

Swars, S., Hart, L. C., Smith, S. Z., Smith, M. E., & Tolar, T. (2007). A longitudinal study of elementary pre-service teachers' mathematics beliefs and content knowledge. *School Science and Mathematics, 107*(8), 325–335. http://eric.ed.gov/?id=EJ790443

U.S. Department of Education. (2013). *Reading 2013 state snapshot report, Mississippi grade 4 public schools.* Washington, DC: Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress. Retrieved November 4, 2015, from https://nces.ed.gov/nationsreportcard/subject/publications/stt2013/pdf/2014464MS4.pdf

U.S. Department of Education. (2016a). *Does content-focused teacher professional development work? Findings from three institute of education sciences studies.* Washington, DC: Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved December 15, 2016, from http://ies.ed.gov/ncee/pubs/20174010/pdf/20174010.pdf

U.S. Department of Education. (2016b). *Public teacher's data file 2011–12.* Washington, DC: National Center for Education Statistics, Schools and Staffing Survey. Retrieved November 4, 2015, from http://nces.ed.gov/datalab/sass/index.aspx.

Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). *BILOG-MG 3: Item analysis and test scoring with binary logistic models.* Chicago, IL: Scientific Software International.

Zaidman-Zait, A., & Zumbo, B. D. (2013). Can multilevel (HLM) models of change over time adequately handle missing data? *The Journal of Educational Research & Policy Studies, 13*(1), 18–31.

The Regional Educational Laboratory Program produces 7 types of reports

Making Connections
Studies of correlational relationships

Making an Impact
Studies of cause and effect

What's Happening
Descriptions of policies, programs, implementation status, or data trends

What's Known
Summaries of previous research

Stated Briefly
Summaries of research findings for specific audiences

Applied Research Methods
Research methods for educational settings

Tools
Help for planning, gathering, analyzing, or reporting data or research