Strengthening Comprehension Instruction for Higher Achievement in Second and Third Grades

Based on Findings and Lessons Learned from Reading First Programs in Florida, Michigan, Pennsylvania, and Utah
Strengthening Comprehension Instruction for Higher Achievement in Second and Third Grades

Based on Findings and Lessons Learned from Reading First Programs in Florida, Michigan, Pennsylvania, and Utah

Alice Furry
Kathleen Cooper
RMC Research Corporation

Acknowledgments

The authors are grateful to the researchers who provided information on reading comprehension achievement and focused classroom observations from their external evaluations and technical reports of the Reading First programs featured in this report:

Florida / Barbara R. Foorman, Professor of Education and Director of the Florida Center for Reading Research, along with Jane Meadows, Project Coordinator for the 2004–2008 Florida Reading First site visits at Florida State University

Michigan / Joanne F. Carlisle, Professor of Education at the University of Michigan and principal investigator for the evaluation of the Michigan Reading First Program, 2003–2008

Pennsylvania / Rita M. Bean, Professor Emerita and Naomi Zigmond, Professor at the University of Pittsburgh School of Education, and co-principal directors for the Pennsylvania Reading First external evaluation, 2004–2009

Utah / Janice A. Dole, Professor and Director of the Center for Reading and Literacy at the University of Utah and co-principal investigator with J.K. Hosp and M.L Hosp on the Utah Reading First Evaluation Team, 2003–2008

The National Reading Technical Assistance Center also appreciates the contributions of Cari Miller, state Reading First Director in Florida, and Rebecca S. Donaldson, state Reading First Director in Utah. The Center is especially thankful for the views expressed by all of the individuals named here in email exchanges and conference calls held in December 2010.

June 2011
Executive summary

Research has shown that many children who read at third-grade level in grade 3 will not automatically become proficient comprehenders in later grades. Therefore, teachers must teach comprehension explicitly, beginning in the primary grades and continuing through high school (Snow, 2002, xii).

Teachers often assume that students will learn to comprehend merely by reading. Although some will, many others will not. Teaching students to comprehend is challenging because reading is complex (Snow, 2002, 5).

A key provision of the Reading First (RF) Program under the No Child Left Behind Act of 2001 (NCLB) requires improving the quality of reading instruction through scientifically based reading programs and instructional methods (aligned with scientifically-based reading research, or SBRR) in five essential components. Of these five, the quintessential and culminating component is reading comprehension. As defined in the Reading First Impact Study Final Report (Gamse et al., 2008, 3),

Comprehension is the understanding of what is being or has been read. Students will not understand text if they can read individual words, but not understand what sentences, paragraphs, and longer passages mean. Proficient readers elicit meaning from—or comprehend—text, rather than simply identifying a series of words. Instruction in comprehension strategies provides specific tools for readers to use to make sense of the text they read. Comprehension strategies are vital to the development of competent readers because they aid in understanding the collective significance of words, sentences, and passages.

This study looks at how four states—Florida, Michigan, Pennsylvania, and Utah—set out to increase the number of grade-level, proficient readers through the implementation of Reading First guidelines addressing critical elements of comprehension instruction. These practices include:

• routine use of assessments for targeting instruction and monitoring student progress,
• on-going opportunities for teachers to receive professional development,
• systematic and explicit instruction based on SBRR-aligned core reading programs, and
• evidence-based pedagogy for teaching comprehension strategies.

This report begins with a brief examination of current research on these four practices, followed by a summary of findings and analyses based on the four selected states’ external evaluations and technical reports, including actual student achievement results on comprehension measures and observed strengths and weaknesses of the teaching practices.

The report concludes with four recommendations for strengthening comprehension instruction in order to increase the proportions of second and third graders performing at or above grade level by year’s end:

• Improve the trustworthiness of assessment data on comprehension achievement
• Improve the comprehensiveness and long-term utility of teacher and coach professional development on comprehension instruction

• Continue usage of comprehensive core reading programs (SBRR aligned) and increase instructional time for daily comprehension lessons

• Acknowledge and communicate the current research on effective pedagogy related to comprehension instruction and increase research on classroom practices that lead to gains in students’ comprehension achievement.

The report has four sections:

Section 1, Introduction, sets the context for the study and identifies comprehension as an issue of increasing concern in the educational community and nation at large and provides the reader with the organizational structure of the full study.

Section 2, Evidence-based practices for comprehension instruction, offers a brief review of the research on effective teaching practices that lead to improved student achievement: monitoring assessments, professional development for teachers, core reading programs and instructional time, and pedagogy.

Section 3, Comprehension achievement linked to classroom practices in selected Reading First states, provides a summary of key findings with analyses for each of the teacher practices introduced in Section 2.

Section 4, Recommendations for strengthening comprehension instruction and achievement, concludes with four recommendations, with subsets associated with the four teaching practices, that address large-scale literacy initiatives with stronger comprehension instruction.
The Reading First policy to practice initiative, under the No Child Left Behind (NCLB) Act of 2001, Part B, § 1201, places instructional responsibility on state education agencies (SEAs) and local education agencies (LEAs) to provide teachers with K–3 reading programs aligned with scientifically-based reading research (SBRR). The law requires that comprehension strategies be taught explicitly and systematically (Part B, § 1208). This study focuses on teaching practices that strengthen reading comprehension for K–3 students with an emphasis on lessons learned in four states (Florida, Michigan, Pennsylvania, and Utah) at the conclusion of Reading First (2008 or 2009).

Research shows that comprehension strategy instruction assists K–3 learners to understand, remember, and communicate with each other about what they have read (National Reading Panel, 2000). State technical assistance networks have communicated this finding to LEAs, with the related findings that explicit instruction has been researched as a successful way to help readers use specific comprehension strategies and students can be taught to use these strategies.1

National and state evaluation reports indicate consensus that Reading First did not have the same impact on students’ comprehension achievement as it did on three other components: phonemic awareness, phonics, and oral fluency (Gamse et al., 2008). A simple explanation may be that these three components are finite—they can be sequenced, taught systematically, and assessed easily. Another reason for disappointing student achievement may be the way comprehension was likely taught (Beck, 2010, 96).

The instructional cycle of Reading First was based on a model to plan, teach, assess, and adjust, with the cycle repeated frequently. The cycle required monitoring assessments to know how to differentiate instruction, on-going professional development to increase teacher knowledge, adequate time to conduct lessons aligned with SBRR, and effective pedagogy to guarantee achievement.

In Section 2, the current research on comprehension instruction is organized around monitoring assessments, professional development, core reading program content and instructional time, and pedagogy. In Section 3, the four states’ key practices are summarized for relevant strengths and weaknesses found in implementing comprehension instruction. The report concludes with four categories of recommendations for strengthening comprehension instruction and student achievement based on the research and the findings from these four states.

---

1 The most important strategies listed in the research resource document Put Reading First (CIERA, 2003, Second Edition) are monitoring comprehension by identifying where difficulty occurs, restating a sentence or passage, or looking back or forward through the text; using graphic and semantic organizers; answering questions; generating questions, recognizing story structures or informational text structures, and summarizing. Additional strategies include relating to prior knowledge, making and confirming predictions, and visualizing. Higher-order, critical thinking skills such as inferring from text, synthesizing knowledge across narrative and informational texts, or critiquing an author’s point of view are reinforced as pedagogical approaches for ensuring students are able to interpret the meaning of text.
Scientifically based research findings and confirmed practices described by knowledgeable experts shape what is known about effective comprehension instruction. The purpose of this section of the study is to isolate those practices that are found to significantly affect student comprehension achievement. The specifics of these practices are grouped under general headings and are described in this section of the study, to be followed in the subsequent section with an examination on how well these practices were implemented and evaluated in four selected Reading First funded states.

Assessments for monitoring comprehension

Regular assessment of students’ academic knowledge and skills can offer teachers a clear picture of a student’s strengths and weaknesses. Data on progress can benefit both student and teacher, instructionally. Monitoring students is important and so is monitoring instruction (Dewitz et al., 2010, 251). There are four types of purposeful assessments. The first three need to be administered regularly to monitor student progress and instructional effectiveness. The fourth needs to be administered annually to monitor overall program success. All four need to meet high standards for reliability and validity.

- **Screening reading assessments** are a first step in identifying students who may be at risk for delayed development or academic failure and who need further diagnosis for special services or additional reading instruction.

- **Diagnostic reading assessments** identify a student’s specific areas of strength and weakness to move the student in a timely way toward proficient reading.

- **Progress monitoring assessments** (classroom-based) evaluate an individual student’s and whole class learning based on systematic observations and curriculum-embedded tests to check progress and determine instructional effectiveness.

- **Program outcome assessments** make summative judgments at end-of-unit or end-of-year intervals.

A purposeful system that assesses students’ comprehension abilities builds the teacher’s capacity to identify good and poor comprehenders (individually and in subgroups), differentiate instruction; plan daily lessons, and monitor students’ on-going progress toward mastery. (Snow, 2003, 193). It follows, then, that such a system should be easily accessible and widely used in state and local educational systems. Many appropriate, valid, and reliable grade 1–3 comprehension assessments exist. Publishers usually include curriculum-embedded comprehension tests in their comprehensive core reading programs at the end of weekly lessons and/or units, at mid-year or trimester intervals.

2 The Carnegie Corporation of New York’s Council on Advancing Adolescent Literacy published Measure for Measure: A Critical Consumer’s Guide to Reading Comprehension Assessments for Adolescents (Morsy et al., 2010). The Guide suggests widely-used and commercially-available assessments. The assessments can be used in grades 1-3 as well for grades 4-12. For example, for screening: Degrees of Reading Power (DRP), Scholastic Reading Inventory (SRI); for diagnostics: Qualitative Reading Inventory (QRI), Diagnostic Assessment of Reading (DAR); hybrids with multiple purposes: Gates-MacGinitie (GMRT), Group Reading Assessment and Diagnostic Evaluation (GRADE), Stanford Diagnostic Reading Test (SDRT), Gray Oral Reading Test (GORT 4).
terms, and at the end of the year. Rarely, however, are screening and diagnostic tests for comprehension found in
the comprehensive core programs. If they are offered, they often lack evidence of reliability and validity.

**Professional development on comprehension instruction**

Professional development for teachers often drives instructional reform. Reform depends heavily on professional
development to support teachers in bringing scientifically-based reading instruction practices into the classroom.
Following the guidelines of the National Reading Panel (2000), professional development in comprehension should
cover content, scope, and efficacy to deepen instruction and produce students who can understand and respond to
text as critical thinkers.

Building student proficiency in comprehension can be characterized as the next horizon in professional
development. Teachers need to know how to best teach comprehension and measure student outcomes.

> Teachers explain comprehension more often than they did 30 years ago, but their explanations are not
effective enough for all students to reach independent transfer to real reading. Of primary importance in this
area is the need to approach professional development from the perspective of developing professional
thinkers as opposed to developing technicians who follow procedures (Block et al., 2008, 32).

Comprehension instruction has moved far beyond the simple posing of teacher questions to elicit student answers.
The content of professional development needs to cover the research and instructional practice on comprehension.
Some examples:

- The direct teaching of comprehension strategies has been shown to be effective. Block and Duffy suggest
a list of nine researched and validated strategies for improving reading comprehension: (1) prediction,
(2) monitoring, (3) questioning, (4) imagining, (5) looking back, rereading, and fixing, (6) inferring,
(7) finding main ideas, summarizing, and drawing conclusions, (8) evaluating, and (9) synthesizing. (Block
et al., 2008, 22).

- The literature supports direct explanation of strategies coupled with modeling and coaching (guided practice)
to transfer responsibility from teacher to student (Duffy, 2002; Taylor et al., 2003, 24).

- Research supports integrated, multiple strategy instruction to teach students to adapt and use strategies
flexibly, according to the reading task. Examples of models include:
  - experience-text-relationship methods that stress building background knowledge before introducing a text
to students,
  - KWL (know, want to know, and learned), a method that encourages active thinking about a text,
  - reciprocal teaching builds on using the four comprehension strategies of predicting, questioning, clarifying,
and summarizing to uncover meaning, and
  - QAR (question/answer relationships) categorizes questions by how they can be answered–from expressed
statements in the text, from implicit information in the text, or from the reader’s background knowledge.

Recent data suggest that teachers need intensive, continued training over four to nine months and ample
explanations and time to practice before they can become adept at explaining and teaching comprehension
strategies to their students, especially struggling readers (Collins, 1991; Duffy, 1993). A recent analysis of more
than 1,300 studies on the relationship between teacher professional development and student achievement found
five studies in which teachers who received an average of 49 hours of professional development annually boosted
their students’ achievement about 21 percentile points (Yoon et al., 2007, iii).
Several studies confirm that teachers move through stages in learning to teach comprehension (Block et al., 2008, 28). Many teachers begin by naming strategies and telling students to use them. They may progress to modeling and integrating multiple strategies with text. Some teachers reach an innovative stage where they can articulate why strategies are important and understand how to teach them. Supporting a teacher’s progress toward the highest level of comprehension instruction—invention—is the goal of sustained, intensive, and content-focused professional development. Unfortunately, few studies exist that confirm the claim that teacher knowledge affects student achievement (Hill, 2007, 117). Most studies have found only a weak relationship between teachers’ general knowledge of reading and student outcomes (Carlisle et al., 2008; and Carlisle et al., 2009). We need new studies that investigate and confirm whether teacher knowledge assessed through observed classroom practices influences student outcomes.

Comprehension lessons in core reading programs

Through NCLB, the label “basal reading program” (for forming base, fundamental instruction) changed to “core reading program” (for providing critical, primary, essential instruction). Now educators generally believe that core reading programs provide an important instructional tool “for organizing instruction by time, texts, and teaching strategies” (Dewitz et al., 2010, 281). When a local educational agency adopts a core reading program today, it normally prepares and monitors a system-wide pacing guide to guarantee that all or most of the lessons are delivered by the end of the school year. Since the National Reading Panel (2000) issued its report, core reading programs carry the banner “based on scientifically based reading research.” However, these programs have not been comprehensively subjected to rigorous experimental research with students and teachers randomly assigned to treatment and control groups. Therefore, one cannot assume that they will succeed with all readers. Rather, “at best, the label of SBRR means that authors, editors, and freelance writers created instructional lessons and incorporated skills and strategies that reflected their best understanding of the reading research” (Dewitz, et al., 2010, 175).

Success with a core reading program usually means that users maintain fidelity in program content and delivery. A high level of fidelity occurs when “you get a program with an internal design and you follow that design. This would include using the materials in a particular sequence, adhering to the amount of time and practice called for by the program and following the recommendations for grouping or re-teaching students. It would mean using all of the essential components as they are designed.”

Many studies, not solely focused on the use of a comprehensive core reading program, have attempted to find a relationship between the degree of fidelity to program implementation and achievement gains (Dane & Schneider, 1998; Ruiz-Primo, 2006). Some researchers have found that the duration and intensity of program implementation are significant predictors of achievement (Schiller, 2001) and significantly correlate with student outcomes (Leinhardt, Zigmond & Cooley, 1981). Some recent program evaluations have addressed whether there is a relationship of fidelity to a program (including comprehensive core reading programs) to significant gains in student achievement.6

3 Refer to the Five Criteria for High-Quality Teacher Professional Development (No Child Left Behind Act of 2001, Section 9101 under Part A of Title IX, Definition 34).
4 The reference to “we” throughout this report represents the authors’ considered opinion.
5 Credit for this definition is attributed to Linda Diamond of the Consortium on Reading Excellence referred in Dewitz et al., 2010, 311.
6 For example, the California finding (Haager et al., 2008, 72): “California examined the fidelity of implementation as defined by the ‘degree to which an intervention (program) is implemented as planned (Greshman, Gansle, & Noell, 1993).’ Based on the state’s Reading First Implementation Index survey measure, which pooled items over a four-year period (2004-2008), the authors found that the largest effect size and the highest correlation between schools’ student achievement scores and implementation factors were attributed to the dimension of school implementation/instruction. Referred to as the ‘most powerful implementation dimension’ with a statistically significant (p<.05) standard Beta effect of 0.087, the 28 response items covered the use of the core reading program — pacing schedule, pre-planning, monitoring assessments, and management of fidelity to the program.”
The teaching of comprehension strategies with fidelity to the core reading program may or may not fully cover all major comprehension strategies or support the required pedagogy. Recently, the International Reading Association commissioned curriculum analyses of comprehension strategy instruction in five of the most used core reading programs published either in 2003 or 2005: McGraw-Hill Reading, SRA Open Court, Harcourt Trophies, Houghton Mifflin Reading, and Scott Foresman Reading (Dewitz et al., 2010,127).

The analysis revealed strengths and weaknesses in the design and introduction of comprehension strategy lessons through one year of instruction. The findings for grades 3 through 5 programs suggest that not all core reading programs are equal though all have some relative strengths and weaknesses. A caveat of the study was that teachers need to “know thy program” by grade level. They need to be vigilant about the adequacy of coverage for specific evidence-based instructional actions (for example, asking questions and modeling answers, explaining and modeling strategies, engaging in direct explanation, engaging in discussion, and providing guided and independent practice); and the number of key comprehension strategies covered in the lesson or unit (for example, predicting, self-questioning, comprehension monitoring, summarizing, determining importance, making inferences, and focusing on text structure).

Publishers rarely recommend how much instructional time should be dedicated to comprehension lessons in their core reading programs, leaving it to the LEA to make that decision in terms of minutes of comprehension instruction per day. While research shows that the amount of reading instruction time in elementary school strongly predicts later reading achievement (Snow et al., 1998,129), no evidence exists confirming that the amount of daily instruction time in comprehension instruction matters.

One expert in early reading and an author for a major K–6 core reading program provides one explanation for the lack of recommendations about how much time should be dedicated to comprehension strategy instruction:

Even if specified times were provided by grade level for comprehension instruction in a publisher’s program, minutes by themselves without consideration of the quality of instruction, will not improve reading proficiency.
(M. Roit, personal communication, December, 2010).

The amount of time for comprehension instruction appears to increase as the grades progress upward. In kindergarten and grade 1, more instructional minutes are given to phonemic awareness, phonics, and oral fluency instruction, and fewer minutes for vocabulary and comprehension. In grades 2 and 3, less emphasis is given to these reading components, allowing more time for comprehension and vocabulary instruction. By grades 4 through 6, establishing an appropriate amount of time for comprehension strategies becomes even more difficult as instructional time is shared with history, social studies, science, and math.

A recent study on reading volume (or text reading) found that the grade 3 editions of the major core reading programs recommended about 15 minutes per day for using student reading materials in reading instruction, or about 17% of a 90-minute reading block (Brenner & Hiebert, 2010, 357). However, it is not likely that this represents the full number of instructional minutes for teaching comprehension.

Given the critical need for teaching and practicing comprehension, LEAs and schools should establish a minimum number of daily minutes dedicated to these skills. Comprehension is complicated. It involves word, sentence, paragraph, and extended paragraph levels of understanding. It must result in the transfer of responsibility from teacher to student for discovering and thinking critically about meaning by using comprehension strategies. This necessitates considerable time for instruction, practice, and reflection. We offer a reasonable criterion for teaching comprehension skills—50% of reading/language arts instructional time as early as second grade.
Suggested minimum daily minutes for grades 2 through 6 comprehension instruction per reading block

90-minute reading block = 45 minutes
120-minute reading block = 60 minutes
150-minute reading block = 75 minutes
180-minute reading block = 90 minutes

LEAs and schools should reach an agreement on the allocation of time for comprehension instruction by grade level. The agreement should include classroom observations to confirm compliance with the agreed-upon times. They should also review evidence of student progress with valid, reliable school-wide assessments at least three times a year.

We encourage and anticipate further research to test the intuitive belief that the amount of quality instructional time will result in increased student achievement gains. When comprehension instruction is rigorous and the time is sufficient, increased student achievement should follow.

Pedagogy for comprehension instruction
What really matters is what happens in the classroom. “The ‘how’ of instruction may be as important as the ‘what’” (Taylor et al., 2003, 5). The best pedagogy for comprehension strategies includes:

- **Direct instruction** with complete, clear explanations of how and why strategies are used
- **Modeling and coaching** students to solve problems and make important inferences from what they read
- **Guided practice** so that students can assume more responsibility for using the strategies
- **Guided structuring and monitoring of classroom discussion** to help students develop a deeper understanding of what they read
- **Asking questions that prompt** students to elaborate and explain their thinking and improve the quality of their responses
- **Assessing student work** and providing specific and timely feedback.

No one disagrees that well-delivered instruction makes the difference in student achievement. The *how* can trump the *what*. The question is: *what* promotes effective pedagogy. This is really a holistic systems question. The effective teacher uses valid, reliable, and accessible assessments, participates in effective professional development and transfers new knowledge to classroom practice, employs the best comprehensive core reading program materials available, and protects sufficient instructional time. Improved pedagogy depends on the support the teachers receive while honing their craft. And, “the extent to which teachers are flexible and adaptive in designing and carrying out reading lessons depends on their understanding of reading and reading processes” (Carlisle et al., 2011, 3).
Comprehension achievement linked to classroom practices in selected Reading First states

Answers to the following questions on comprehension achievement and classroom practices for students in grades 2 and 3 appear in the Reading First annual external evaluation reports and technical studies commissioned by states.

- How well did Reading First students achieve on comprehending what they read by the end of third grade?
- What supporting practices (assessment monitoring, professional development, core reading programs, and instructional time) enhanced effective comprehension instruction?
- What pedagogical practices (instructional actions) were found to be significantly associated with student achievement?

This section details findings from four states: Florida, Michigan, Pennsylvania, and Utah. These states were featured in the Journal of Literacy Research (21(1), 2010) on the extent to which Reading First affected the comprehension achievement of primary grade students. The authors of these reports were also the principal investigators commissioned to conduct annual Reading First evaluations and technical reports in their respective states. Over the years, these reports documented achievement data, programs’ compliance status, and most importantly, observation findings on instructional time and pedagogical practices.

As a preface to the discussion of these practices, we offer a brief summary of the achievement outcomes for second and third graders on comprehension achievement tests for the selected states.

**Student achievement results for comprehension**

Reading on grade level by the end of third grade is the achievement goal for Reading First schools. Federal guidelines required annual state reports on whether schools had “significantly increased” the number of students reading proficiently at grade level or above (U.S. Education, 2002, 21). Table 1 shows the states’ first groups of Reading First schools (at the end of either the fifth or sixth year of Reading First funding), along with the proportion of students who met state-determined cut-off scores for proficiency in reading comprehension. The table also shows gain scores from base to final year and the achievement tests administered by each state by grade level.

---

7 The policy statement on student achievement: The purpose of Reading First is to ensure that all children in America learn to read well by the end of third grade (U.S. Department of Education, 2002, 1). Both state and local educational agencies must report schools with the largest achievement gains and progress the state and local educational agencies are making in reducing the number of students in grades 1 through 3 who are reading below grade level based on data that are valid and reliable (U.S. Department of Education, 2002, 20-21).
Table 1

Summary of comprehension achievement results by state
proportion of students at/above proficiency and percentage gains
base year to final reported year

<table>
<thead>
<tr>
<th>State (years) Program years</th>
<th>Grade levels</th>
<th>Proportion at/above proficiency years Base Final</th>
<th>% Gains</th>
<th>Achievement test used to measure comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida (2003-04 &amp; 2007-08) 5 years</td>
<td>2</td>
<td>.55 .64</td>
<td>+9</td>
<td>Stanford Achievement Test 10th Ed. Comprehension Subtest, Cut-off at 40th percentile Florida Comprehensive Assessment Test, Reading Comprehension, Cut-off at 284 for Level 3 proficiency</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.57 .64</td>
<td>+7</td>
<td>Florida Comprehensive Assessment Test, Reading Comprehension, Cut-off at 284 for Level 3 proficiency</td>
</tr>
<tr>
<td>Michigan (2003-04 &amp; 2008-09) 6 years</td>
<td>2</td>
<td>.28 .39</td>
<td>+11</td>
<td>Iowa Tests of Basic Skills, Comprehension Subtest, Form B Cut-off at 50th percentile Iowa Tests of Basic Skills, Comprehension Subtest, Form B Cut-off at 50th percentile</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.24 .29</td>
<td>+5</td>
<td>Iowa Tests of Basic Skills, Comprehension Subtest, Form B Cut-off at 50th percentile</td>
</tr>
<tr>
<td>Pennsylvania (2003-04 &amp; 2008-09) 6 years</td>
<td>2</td>
<td>.49 .49</td>
<td>+0</td>
<td>CTB Terra Nova 3, Comprehension Subtest, Cut-off at 40th percentile Pennsylvania System of School Assessment, Comprehension Subtest, Cut-off score at 1235</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.33 .56</td>
<td>+23</td>
<td>Pennsylvania System of School Assessment, Comprehension Subtest, Cut-off score at 1235</td>
</tr>
<tr>
<td>Utah (2003-04 &amp; 2007-08) 5 years</td>
<td>2</td>
<td>.61 .66</td>
<td>+5</td>
<td>Utah Criterion-Referenced Test, Total Reading (including Comprehension), Cut-off changes Utah Criterion-Referenced Test, Total Reading (including Comprehension), Cut-off changes</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.51 .61</td>
<td>+10</td>
<td>Iowa Tests of Basic Skills, Total Reading, Form B, Cut-off at 50th percentile Iowa Tests of Basic Skills, Total Reading, Form B, Cut-off at 50th percentile</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.45* .38</td>
<td>-7</td>
<td>* ITBS introduced in 2004-05</td>
</tr>
</tbody>
</table>

Note: See References for evaluation report citations: Florida (Foorman et al., 2010, 77); Michigan (Carlisle, 2008, 8); Pennsylvania (Bean, et al., 2010, 15; and Zigmond et al., 2009, 7); Utah (Dole et al., 2005, 34; and 2009, 216).

Student achievement results in comprehension—findings and analysis

Comments from the states’ evaluators confirm their views on whether progress had been made in the proportion of students who were proficient readers by year 5 or year 6 in the program:

- **Florida.** The results, from the first to the fifth year, “revealed a tendency for the percentage of students reading on grade level to increase – ranging from seven to nine percentage points” (Foorman et al., 2010, 88).

- **Michigan.** The third-grade students showed overall relatively little progress (Carlisle et al., 2008, 7). However, there was “only modest improvement in the percentage of students on grade level” (Carlisle et al., 2010, 66). One explanation for this pattern is – “Instruction in the five components of reading required by RF was not sufficiently infused with cognitively challenging instruction of the kind that is thought to contribute to academic achievement” (Carlisle et al., 2010a, 66).
• **Pennsylvania.** The first group of third graders had “the most dramatic change seen in the statewide assessment (PSSA) . . . after six years of implementation . . . (with) an increase of approximately 23% in the percent of students performing at proficient or advanced over the baseline year” (Zigmond et al., 2009, 8).

• **Utah.** “Third grade students in Cohort 1 schools performed about the same in Years 4 and 5. In the fifth year of the program, the Cohort 1, 2, and 3 students demonstrated mixed performance on the tests (CRT for Total Reading) at the different grade levels. In general, more students achieved proficiency at first and second grade levels; however, fewer third graders demonstrated increased performance. These data would seem overall positive, however, it is difficult to interpret these data” (Dole et al., 2009, 5).

Because no federal or consistent state criterion for declaring “significant increase” exists, we believe a proxy gain score of ten percent or more, from the Base Year 1 through Final Year 5 or 6, can be used to interpret whether a state’s reading comprehension program can explain increased proportions of students at/above proficiency. A question remains as to what proportion of students needs to achieve proficiency for the state to be “significantly” successful.

At first glance, the numbers for grade 2 students indicate that only Michigan met the criterion of successfully increasing the proportion of students at/above proficiency (gain score of 11 percent). The numbers for grade 3 students show that both Pennsylvania and Utah met the criterion successfully (gain score of 23 percent and 10 percent, respectively). After deeper examination, specific interpretations of the achievement results are difficult to ascertain. All of the reading tests are valid and reliable. However, there are many variables. The constructs of the tests (content), the type of test (criterion-referenced or norm-referenced), and the determinations for cut-off scores for proficiency vary among states of 40th percentile or 50th percentile (see Appendix A for descriptions of states’ measures on comprehension).

• Michigan’s ITBS norm-referenced test offers a subtest with substantial number of items on reading comprehension skills and practices. While Michigan’s grade 2 students met the gain score criterion of 10 or more percentage points, the actual proportion of students meeting the 50th percentile cut-off score is 39 percent.

The question remains for the state to determine if this percentage represents a “significant” performance.

• Pennsylvania’s PSSA criterion-referenced test offers a subtest with a vague and limited set of reading comprehension skills and practices. The state’s third grade students exceeded the gain score criterion with more than a 13 point gain. With 56 percent of the third grade students performing at grade level there appear to be some successful program effects.

• Utah’s CRT-Total Reading criterion-referenced test is a test that covers all of the reading domains of oral language, phonics, spelling, vocabulary, comprehension, and writing for both second and third grades. Thus, this test, with limited comprehension-focused items, cannot be cited as a measure focused solely on reading comprehension. Therefore, the Utah gain score of 10 percentage points on this test should not be considered as meeting the significant increase criterion. However, the ITBS, for which grade three scores decreased, seems to confirm that the CRT-TR measured much more than reading comprehension.

We note that Florida’s second grade test (SAT-10) and Pennsylvania’s second grade test (CTB, TN-3) used the norm-referenced tests’ 40th percentile to determine grade level proficiency, rather than the publisher recommended 50th percentile. Again, the state-determined cut-off scores for grade-level proficiency on criterion-referenced tests are not comparable across states.
Also worthy of consideration is the extent to which some states’ standards-based criterion-referenced tests for comprehension are comparable to norm-referenced measures. The lack of comparability stems from the different numbers of items and the resulting lack of comprehensive coverage of multiple dimensions of comprehension. Michigan state evaluators provided their rationale for using a norm-referenced test: “The ITBS Reading Comprehension subtest can be considered a better and more challenging assessment than Michigan Educational Assessment Program English Language Arts Reading (MEAP ELA), as it can help provide more accurate and comprehensive information about students’ reading comprehension ability. It thus makes more sense if teachers in Reading First schools can make instructional decisions based on ITBS Reading Total (a composite score of Reading Comprehension and Vocabulary subtests) rather than the MEAP ELA Reading.” (Zeng et al., 2008, 37)

So, how well were Reading First students able to comprehend what they read at the end of third grade? Without federal guidelines that define “significant increases” and, in some cases, no common trustworthy test for measuring reading comprehension, there can be no national standard nor test applied for judging how well Reading First students achieved in reading comprehension. While states confirm progress, we acknowledge that states have made progress, but conclude that a broad generalization based on available achievement data cannot be substantiated.

**Practices that affect comprehension instruction and achievement**

Federal guidelines for Reading First encourage states to include four critical practices that support the teaching of the five fundamental reading components, including comprehension:

- use valid and reliable classroom assessments for screening, diagnosis, and monitoring student progress and instructional effectiveness,
- provide quality teacher professional development programs,
- implement comprehensive core reading programs based on scientifically based reading research (SBRR) and allocate the necessary instruction time, and
- apply proven pedagogy.

This study reviewed how the four Reading First states carried out these key practices that support comprehension instruction in 2nd and 3rd grades. The review revealed the strengths and weaknesses of these practices.

**Monitoring assessments**

Monitoring assessment tools inform teachers about precisely where their students’ comprehension skills are breaking down so instruction can be better targeted. Table 2 indicates the two states of the four studied, that used comprehension monitoring assessments in 2nd and 3rd grades.

---

8 The Reading First guideline for assessment monitoring: A high-quality, effective reading program must include rigorous assessments with proven validity and reliability. These assessments must measure progress in the five essential components of reading instruction (phonemic awareness, phonics, oral fluency, vocabulary and comprehension) . . . Diagnostic assessments provide more in-depth information on students’ skills and instructional needs that forms the basis of the ideal instructional plan. Classroom-based instructional assessments determine whether students are making adequate progress or need more support to achieve grade-level reading outcomes (U.S. Department of Education, 2002, 7).
Table 2

<table>
<thead>
<tr>
<th>State</th>
<th>Assessment</th>
<th>Screening and diagnostics</th>
<th>Progress monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>Diagnostic Assessments of Reading (DAR)</td>
<td>10%</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Degrees of Reading Power (DRP)</td>
<td>10%</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Core Reading Program Lesson/Unit Tests</td>
<td>33%</td>
<td>25%</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>Diagnostic Assessments of Reading (DAR)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Gray Oral Reading Test-4 (GORT-4)</td>
<td>✓</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Core Reading Program Lesson/Unit Tests</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note. The Florida information in the table was found in two site visit reports (FCRR, 2005, 5; FCRR, 2009, 11). The Pennsylvania report did not report the percentage of classroom use, but did classify assessments in terms of their use (Zigmond et al., 2005, 19). The Michigan and Utah state evaluators made no mention of the use of comprehension assessments for diagnostic or progress monitoring in their annual reports.

Monitoring assessments—findings and analysis

Three published norm-referenced tests (DAR, DRP, and GORT-4) have grade level norms and are valid and reliable. Diagnostic Assessments of Reading diagnoses why a student struggles with comprehension. Degrees of Reading Power and Gray Oral Reading Test-4 screen and monitor comprehension progress over time. Core reading program tests produced by publishers offer less reliable and valid profiles of a student’s ability to comprehend.

A case can be made that very few classrooms in these four states use norm-referenced screening and diagnostic tests to determine how far students lag behind, and which comprehension skills are causing substantial difficulties. Progress monitoring tools to periodically assess student progress and the effectiveness of instruction were not routinely used. Given that these tools are not used regularly, measuring instructional effectiveness becomes more difficult. Of note: All states use Dynamic Indicators of Basic Early Literacy Skills (DIBELS) for screening and progress monitoring; however, this test does not measure reading comprehension, which is why it is not listed in Table 2 and why states that only use DIBELS are not presented in Table 2.

The information in Table 2 suggests that comprehension tests for screening, diagnostic and progress monitoring were rarely used in grades 2 and 3 Reading First classrooms across the states studied. One evaluator suggested that “only in the most latter years [of Reading First] did comprehension instruction become a focus of implementation. Phonemic awareness, phonics, and oral fluency mastery were thought, at the time, to be the significant predictors for comprehension achievement, and thus the emphasis on these components, and particularly on fluency, was foremost in the classroom teacher’s repertoire of assessments (state evaluator, personal communication, December 2010).”

---

9 Refer to Morsy et al., 2010 for full descriptions of these assessments for comprehension: DAR, 14-16; DRP, 11-13, and GORT-4, 21-23.

10 Stahl (2003) confirms the researched finding that oral reading accuracy is related to comprehension only in first and second grade, with the correlations in grade 3 and beyond dropping to nearly zero. Thus, oral reading accuracy may be important only in the early grades, while vocabulary and comprehension strategy use become important later on.
All states reported regular use of checklists, observations, and benchmark tests to give teachers daily feedback and prepare teachers to become expert through more data driven practices (FCRR, 2008, 11; Dole et al., 2009, 206). However, if these states failed to use norm-referenced assessments of reading comprehension, teachers would not have the vital information to reflect on how to support students in mastering key comprehension skills and strategies.

Professional development

All four featured states conducted professional development on the five instructional components. For example, leadership agencies in Florida created and delivered professional development. Florida offered summer reading academies and regional and school-based trainings throughout the year (Linder et al., 2007, 4, 22). The Pennsylvania Department of Education sponsored statewide offerings as well. Utah was unique in requiring all K–3 Reading First teachers to complete 21 semester hours (7 courses, 3-hours each) of graduate-level university coursework to earn the Utah Basic Level 1 Reading Endorsement Certificate. Districts contracted with local universities for the courses (R. Donaldson, personal communication, December 2010). Table 3 names the professional development topics on comprehension as documented in annual evaluation reports.

Table 3

<table>
<thead>
<tr>
<th>State</th>
<th>Year</th>
<th>Professional development program on comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>2005–06</td>
<td>Comprehension Instruction</td>
</tr>
<tr>
<td></td>
<td>2006–07</td>
<td>Reading Strategies/Techniques</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creating Independence through Student Owned Strategies (CRISS)</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>2006–07</td>
<td>Meaning of Text, Part I and Part II</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Research-based Approaches to Comprehension Instruction (Text Talk)</td>
</tr>
<tr>
<td></td>
<td>2007–08</td>
<td>Language Essentials for Teachers of Reading and Spelling: Module on Comprehension</td>
</tr>
<tr>
<td>Utah</td>
<td>2003–08</td>
<td>Advanced Reading Comprehension (Level 1 Reading Endorsement Graduate Course)</td>
</tr>
</tbody>
</table>

Note: In focus groups, Florida teachers expressed preference for the 2006-07 offerings over the state-sponsored regional Summer Reading Academies (FCRR, 2008, 15). Michigan teachers had access to a series on Foundations, Texas Reading Academies On-line, and publisher trainings. No specific programs were specified (Carlisle et al., 2007, 27). We can assume some general coverage on comprehension, however, we cannot confirm extensive coverage. Pennsylvania reported considerable numbers of professional development participants in its annual reports. However, only three trainings on comprehension were offered during state Technical Assistance meetings. It is unclear how many teachers took part. (Zigmond et al., 2007 and 2008). Utah’s State Office of Education required all K–3 Reading First teachers to take this three-hour course within the first three years the teacher served as a Reading First teacher.

11 The Reading First guideline for providing professional development: Professional development must prepare all teachers to teach all of the essential components of reading instruction, and to know how they are related, the progression in which they should be taught, and the underlying structure of the English language. Teachers also must understand why some children have difficulty learning to read well (U.S. Department of Education, 2002, 7).
**Professional development—findings and analysis**

Table 3 contains very few state level professional development offerings on the topic of comprehension instruction as reported in annual evaluation reports. Table 3 does not include, however, LEA or school-sponsored professional development programs. On-line programs, publisher trainings, and other options may have featured comprehension-related topics.

By all accounts, there were few professional development offerings in vocabulary and comprehension in the six-year lifespan of Reading First, compared to offerings in phonemic awareness, phonics, and oral fluency. By 2006-07, the evidence was growing that schools felt the need to increase professional development in comprehension. For example, Florida survey data showed requests for "more ongoing professional development (e.g., how to effectively teach vocabulary and comprehension) . . . (with a) move away from recall knowledge and . . . deeper and more explicit . . . teaching, particularly in higher-order thinking skills (principal, FCRR, 2008, 27).

Of the four states’ teacher surveys, only one had a survey question on the value of their training on comprehension, yet the surveys had many questions on phonemic awareness, phonics, oral fluency, and assessments. This and other evidence strongly suggests that the states’ professional development plans were not robust enough to move teachers from the foundational stage to the innovative stage of understanding and teaching comprehension strategies (see Block et al., 2008, 28).

**Core reading programs**

Federal Reading First guidelines required scientifically based reading research-aligned reading programs in every funded school. In practice, this guideline translated into states allowing LEAs to select core reading programs that best served their students. Table 4 identifies the core reading programs used in the states.

---

12 The Reading First guideline for reading program and instructional time: “A high-quality reading program that is based on scientifically based research must include instructional content based on the five essential components of reading. A coherent design includes explicit instructional strategies that address students’ specific strengths and weaknesses, coordinated instructional sequence, ample practice opportunities, and aligned student materials, and may include the use of targeted, scientifically based instructional strategies as appropriate. The design should also consider the allocation of time, including a protected, uninterrupted block of time for reading instruction of more than 90 minutes per day” (U.S. Department of Education, 2002, 6).
### Table 4

<table>
<thead>
<tr>
<th>States</th>
<th>Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>Harcourt Trophies (2003, 2005)</td>
</tr>
<tr>
<td></td>
<td>Houghton Mifflin Reading: A Legacy of Literacy (2003, 2005)</td>
</tr>
<tr>
<td></td>
<td>Reading Mastery Plus (2002, 2005)</td>
</tr>
<tr>
<td></td>
<td>Success for All</td>
</tr>
<tr>
<td></td>
<td>Houghton Mifflin Reading: A Legacy of Literacy (2003, 2005)</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>Harcourt Trophies (2003, 2005)</td>
</tr>
<tr>
<td></td>
<td>Reading Mastery Plus (2002, 2005)</td>
</tr>
<tr>
<td></td>
<td>Success for All</td>
</tr>
<tr>
<td>Utah</td>
<td>Harcourt Trophies (2003, 2005)</td>
</tr>
</tbody>
</table>

### Value of core reading programs — findings and analysis

The reading programs in Table 4 represent those among the top six used across all Reading First states (U.S. Department of Education, 2006, 43). Evaluation reports, in a variety of ways, assert the core programs played a primary role in teaching comprehension:

- Florida. Teachers said that their core reading program played a major role in comprehension instruction: 2004: 91% (FCRR, 2005, 5) and 2008: 75% (FCRR, 2009, 10).

- Michigan. Classroom observations indicated that the core reading program was used for comprehension instruction about 65% of the time in second grade and 77% of the time in third grade (Carlisle et al., Technical Report 8.1, 2010, 16). 13

- Pennsylvania. Evaluators noted that “the core program selected was not a guarantee that a school would improve in the percent of students at proficiency” on the end-of-year reading comprehension measure, PSSA (Zigmond et al., 2008, 35).

Reading First is credited for promoting the use of comprehensive core reading programs school-wide to teach beginning reading. In Florida and Michigan, the core reading programs’ comprehension lessons were widely used. 14 No studies were done by the states to determine whether the programs were equally successful in increasing student achievement in comprehension. The term “fidelity” related to the core reading program was neither defined nor measured by the states.

---

13 Appendix B provides information on the Teacher’s Instructional Practices (TIP) observation tool used to confirm use of the core reading program.

14 For example, based on classroom observations, three to four times annually in 2006-07, Michigan evaluators found that the core reading program’s comprehension lessons were used for instruction about 65% of the time for grade 2 and 77% of the time for grade 3 (Carlisle et al., 2010b, 16).
Studies analyzing LEA practices have found that the most improved schools have fully utilized a common core reading program in every classroom (e.g., Williams et al., 2006). (However, we have limited data as to whether comprehension lessons were extensively used in Reading First schools, or whether the schools that did use them were among the most improved.) Educators generally agree that core program lessons serve to set forth explicit instructional expectations, reduce the fragmentation of instruction, and offer opportunities for teachers to collaborate on lesson planning by grade level. State evaluation reports usually gave only a state-level view of the featured program. LEAs often show data on monitoring results on the degree to which the program was fully and effectively taught, whether all program resources were used, and whether teachers attended training on program implementation and differentiating instruction.

**Instructional time**

Evaluators in Florida and Utah reported instructional time dedicated to comprehension instruction. Table 5 displays the observation tools used and the average number of minutes over multiple years. The descriptors used to measure comprehension instructional time on the Instructional Content Emphasis observation tool (forms ICE-R and ICE-R2) included vocabulary in the context of reading or discussion, prior knowledge/predicting, students monitoring their comprehension through activities, students or teacher monitoring through listening comprehension, instruction or use of comprehension strategy, and other instruction involving getting meaning from the text (see Appendix B for more details on the ICE-R observation tool).

**Table 5**

<table>
<thead>
<tr>
<th>State</th>
<th>Year</th>
<th>Observation tool</th>
<th>Grade level</th>
<th>Total random classrooms</th>
<th>Average minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL(^{15})</td>
<td>2004-07</td>
<td>ICE-R</td>
<td>2</td>
<td>103/105</td>
<td>Comprehension: 49.3, Comprehension: 58.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UT(^{16})</td>
<td>2005-07</td>
<td>ICE-R2</td>
<td>2</td>
<td>91/90</td>
<td>Comprehension: 61.2, Comprehension: 63.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While the high end of Florida’s reading instructional block was around 100 minutes daily, its percentage of total block minutes was higher than Utah’s (49% for 2nd grade and 57% for 3rd grade), yet the difference in total minutes between the two states is about 80 minutes. Assuming that 50% of the block is recommended for comprehension, Utah should spend half of its reading/language arts block on comprehension—an average of 90 minutes per day.

Pennsylvania’s 2007-08 observation study, displayed in Table 6, used the CIERA School Change Classroom Scheme (renamed Reading First Observation Checklist and Rubric), to measure the proportion of time and

---

\(^{15}\) Florida evaluators reported a range of instructional time for the full reading block (grade 2: 93.4 minutes to 100.4 minutes and grade 3: 91.7 minutes to 102.0 minutes) (FCRR, 2005, 2006, 2007, and 2008). The minutes were calculated based on an average of these minutes and percentage of time engaged in comprehension content over five years with 103 2nd grade and 105 3rd grade randomized observations. Comprehension instruction accounts for approximately 49% and 57% for 2nd and 3rd grade respectively (using the highest range of minutes).

\(^{16}\) In Utah, most classrooms posted a daily 180 minute reading block. Minutes were calculated based on a reading block of 180 minutes in 2nd and 3rd grades and the percentage of time spent on comprehension content over three years with 91 2nd grade and 90 3rd grade randomized observations. Comprehension accounts for 34% of total reading instructional time for 2nd grade and 35% for 3rd grade (Dole et al., 2006, 2007, and 2008).
associated minutes attributed to comprehension skill instruction (e.g., teaching the main idea, sequencing) and strategies (e.g., teaching how to summarize, predicting) in the 45-minute period within the 90-minute reading block (Zigmond, 2008, 23).  

Table 6

<table>
<thead>
<tr>
<th>Grade level</th>
<th>Comprehension lesson content</th>
<th>% of time within 90-minutes</th>
<th>Actual minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Skills and Strategies</td>
<td>6.6</td>
<td>6.0</td>
</tr>
<tr>
<td>3</td>
<td>Skills and Strategies</td>
<td>12.2</td>
<td>11.0</td>
</tr>
</tbody>
</table>

The comparison between percentage of time for comprehension instruction as a whole (in Table 5) to time for specific comprehension lessons for skills and strategies (in Table 6) raises two issues when interpreting the differences in actual minutes: 1) the observation instruments used are not comparable, and 2) the relative value of time focused on comprehension as a whole and its specific components of skills and strategies.

Instructional time analysis

The quality of instruction, balanced with the amount of instructional time (at least 50% of the reading instruction block), may affect achievement results. Unfortunately, there is no comparative achievement data or 50% instruction time data to test this assumption in the context of these four states’ results. A more accurate statement might be that instructional time is important but insufficient; and high quality instruction without adequate instructional time is also insufficient. In our opinion, it stands to reason, because teaching comprehension is both scientific and artful, teachers need enough time each day to extend guided and independent practice, provide constant attention to students’ learning needs, and offer instruction that leads to higher levels of thinking in response to text.

Pedagogy for comprehension instruction

From the four featured states, two observation studies exist on teachers’ instructional actions related to comprehension. The evaluators for Pennsylvania and Michigan sought to determine the proportion of practices that took place during comprehension instruction in 2nd and 3rd grade classrooms. The observation tools, methods of study, and the profiles of the schools were varied. Both states offer considerable insight into the proportion of instructional actions dedicated to comprehension.

Pennsylvania

Table 7 reports Pennsylvania’s results on observed practices in the category of content emphasis for instruction for comprehension skills and strategies, level of student discourse on meaning of text, and the nature of teacher instructional actions when teaching the core reading program. The evaluators used the CIERA Classroom Observation Scheme (Taylor et al., 2003, 25-26) to classify the proportion of five-minute observation segments

---

17 See Appendix B for a description of the CIERA School Change Classroom Scheme.

18 The Reading First guidelines for instructional practices: Certain elements should be visible in any Reading First classroom . . . regardless of which specific program is in use . . . Expectations are clear . . . in-class grouping strategies are in use, including small group instruction as appropriate to meet student needs . . . There is active student engagement in a variety of reading-based activities, which connect to the five essential components of reading and to overall, clearly articulated academic goals . . . high levels of time on task are also evident (U.S. Department of Education, 2002, 6).

19 Appendix B contains descriptions of the observation tools. Appendix C contains background profiles on the characteristics of the school populations.
during 45-minute observations of instructional content (comprehension skills and strategies), student discourse (on meaning of text), and the nature of instruction (e.g., recitation). Note that in Table 7, Group 1 schools means those schools that were observed for the first time in the first two Reading First years (2003-2005); Group 2 schools were observed for the second time during the evaluation years 2004-2006; and Group 3 schools were observed for the third time during the evaluation years 2005-2007. The Group 3 outcomes possibly indicate the cumulative impact of more years in the program on changing practices.

Table 7

<table>
<thead>
<tr>
<th>Mean proportion of five-minute observation segments during 45-minute observations of comprehension lessons in combined 2nd and 3rd grade classrooms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CIERA adapted dimensions</strong></td>
</tr>
<tr>
<td><strong>Content emphasis of instruction</strong></td>
</tr>
<tr>
<td>Comprehension skills and strategies</td>
</tr>
<tr>
<td>Skills (e.g., main idea, cause-effect, fact-opinion)</td>
</tr>
<tr>
<td>Strategies (e.g., predicting, summarizing)</td>
</tr>
<tr>
<td><strong>Student discourse</strong></td>
</tr>
<tr>
<td>Meaning of text, lower level</td>
</tr>
<tr>
<td>(Students talking about meaning of text that is at a lower level of thinking)</td>
</tr>
<tr>
<td>Meaning of text, higher level</td>
</tr>
<tr>
<td>(Students talking about meaning of text that is engaging them in higher-level thinking)</td>
</tr>
<tr>
<td><strong>Nature of instruction</strong></td>
</tr>
<tr>
<td>Recitation/questioning and answering</td>
</tr>
<tr>
<td>(Teacher is engaging students in answering questions, or responding to low-level question and answer format)</td>
</tr>
<tr>
<td>Modeling</td>
</tr>
<tr>
<td>(Teacher is showing/demonstrating how to do something or how to do a process as opposed to simply explaining it)</td>
</tr>
<tr>
<td>Coaching</td>
</tr>
<tr>
<td>(Teacher is prompting/providing support that can transfer to other situations as students attempt to answer a question or perform a strategy or activity)</td>
</tr>
<tr>
<td>Telling/explaining</td>
</tr>
<tr>
<td>(Teacher is telling or giving students information or explaining how to do something)</td>
</tr>
</tbody>
</table>

Note: This table is adapted from the Pennsylvania External Evaluation Report (Zigmond et al., 2007, 28, 49).

**Pedagogy for comprehension instruction—Pennsylvania’s findings and analysis**

The Pennsylvania studies of three groups of schools, with varying numbers of observations, offer considerable observation evidence on the frequency of instruction in comprehension skills and strategies (e.g., in Group 3, 27 percent of observations); how classroom student discussions emulated higher-level thinking discussions on text
meaning (e.g., in Group 3, 28 percent of observations); and what instructional actions were more commonly used (e.g., in Group 3, 62 percent of observations for recitation through low-level questioning and answering; and 60 percent of observations for telling or explaining information).

The performance of Group 3 schools, with 132 classroom observations, demonstrates the cumulative effects of maintaining or changing teacher practices during the second through fourth year in the Reading First program. In summary, comparisons between Group 1 schools (the first or base group of schools with only one observation between 2004 and 2005) and Group 3 schools (the final group of schools with three observations between 2005 and 2007) show improved emphasis of comprehension instruction. Observations of Group 3 in the final group of schools show increased attention to skill and strategy instruction (.19 to .27); a significant increase in teacher coaching of students (.33 to .41); and giving information and explanations to students (.43 to .60); and a decrease in low-level classroom discourse (.38 to .28).

The findings demonstrate that in the latter years of Reading First, teachers (in Group 3) were focusing more instructional time on comprehension instruction—providing more explicit explanation and more prompting for student responses and as such were becoming more skillful teachers of reading.

**Michigan**

Michigan conducted an observational study of teachers’ instructional methods in the context of comprehension lessons. The evaluators stated, “unless teaching methods or instructional actions are viewed within the context of the lesson in which they are used, the emerging picture of instruction is not interpretable – and certainly would not be useful to reading teachers.” (Carlisle et al., Technical Report, 8.1, 2010, 25). Based on the Michigan Reading First observation data, using the Automated Classroom System for Reading (ACOS-R), three dimensions of instructional actions were studied: pedagogical structure, teacher-directed instruction, and support for students’ learning. Observations were gathered four times during 2007-08 from 287 comprehension lessons with third graders in 19 schools in six Reading First districts.20 The research question, associated with the findings illustrated in Table 8, asked whether third-grade teachers’ instructional actions during reading comprehension lessons contributed to their students’ reading comprehension achievement (Carlisle et al., 2011, 1).

20 Appendix B contains the description of the observation tool. Appendix C contains a background profile of the school population’s characteristics.
Table 8

<table>
<thead>
<tr>
<th>Lesson ACOS-R dimensions</th>
<th>Instructional action</th>
<th>Proportion of action used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogical structure (PS)</td>
<td>PS1: Explaining the purpose of the lesson</td>
<td>.36</td>
</tr>
<tr>
<td>(Actions that teachers take to help students understand the purpose and structure of a given lesson)</td>
<td>PS2: Explaining the value/relevance of the lesson</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>PS3: Giving directions for activity</td>
<td>.78</td>
</tr>
<tr>
<td></td>
<td>PS4: Providing a wrap up or summary of what has been accomplished</td>
<td>.13</td>
</tr>
<tr>
<td>Teacher-directed instruction (TDI)</td>
<td>TDI1: Telling/Explaining</td>
<td>.77</td>
</tr>
<tr>
<td>(Actions that teachers take to ensure effective learning and practice of literacy skills and knowledge)</td>
<td>TDI2: Modeling/Coaching</td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>TDI3: Asking questions for evaluation</td>
<td>.85</td>
</tr>
<tr>
<td></td>
<td>TDI4: Providing practice or review activities</td>
<td>.69</td>
</tr>
<tr>
<td>Support for student learning (SSL)</td>
<td>SSL1: Fostering discussion</td>
<td>.29</td>
</tr>
<tr>
<td>(Actions on part of the teacher to engage students in the lessons, assess their responses to the content and activity of a lesson, and make use of students’ skills, strategies, and knowledge)</td>
<td>SSL2: Assessing students’ work; providing feedback</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>SSL3: Giving students an opportunity to ask questions</td>
<td>.10</td>
</tr>
</tbody>
</table>

Note: Table 8 is adapted from the Scientific Studies of Reading article by Carlisle et al., 2011,10: “Embracing the complexity of instruction: A study of the effects of teachers’ instruction on students’ reading comprehension.”

Proportion of instructional actions used during comprehension lessons—Michigan’s findings and analysis

The unique and well-designed approach to examine teachers’ instructional actions within and across third grade comprehension lessons offer some preliminary findings about whether there is a relationship between teachers’ instructional actions and student achievement gains.21 There were several noteworthy findings22:

- Based on the three instructional actions offer evidence that:
  - The teacher-directed instruction (TDI) was found more frequently in lessons than the pedagogical structure (PS) and the support for student learning (SSL).
  - The length of the comprehension lesson positively influenced the use of instructional actions (PS, TDI, SSL).

- Based on student achievement performance (ITBS) offer convincing evidence that:
  - Teachers’ instructional actions, TDI and SSL, were significantly and positively related with student reading comprehension gains (Carlisle et al., 2011, 24).
  - Teachers’ self-reporting on their instructional practices showed two positive associations with students’ achievement: frequency of use of the comprehensive program and of differentiated instruction (Carlisle et al., 2011,19).
  - Teachers’ “frequent use of TDI and SSL tended to increase reading comprehension achievement more for students eligible for FRL than for students not eligible for FRL” (Carlisle et al., 2011,19).

21 In a preceding study (2006-07), the Michigan evaluators observed that 77 percent of comprehension lessons were based on the use of the comprehensive core reading program (Carlisle et al., Technical Report 8.1, 2010b, 16). See Appendix B for a description of the Teachers’ Instructional Practices (TIP) used to identify core reading program for teaching comprehension lessons.

22 The statistical results summarize the basic findings reported in the Scientific Studies of Reading article by Carlisle et al., 2011, 16-22.
The evaluators concluded that they found a way to identify effective teachers based on their use of the three dimensions of instructional actions within and across comprehension lessons. The findings reinforce the importance of the three dimensions, which are distinctive from each other. More specifically, the findings demonstrate what has been thought to make a difference in practice and outcomes, but now there is some evidence to suggest the following:

- Pedagogical structure (PS) finds teachers more often help students understand the directions for an activity (PS3, .78), and rarely explain the value/relevance of the lesson (PS2, .09) or provide a summing up of what has been accomplished during the day’s lesson (PS4, .13).

- Teacher-directed instruction (TDI) finds teachers using all of the sub-features fairly evenly, but with more frequent use of asking questions for evaluation (PDI3, .85).

- Support for students (SSL) finds teachers modestly helping students engage in the lessons (SSL1, .29 and SSL2, .22), but very rarely handing-off to students the opportunity, or responsibility, to ask questions (SSL3, .10)

The evaluators concluded with an important caveat: “We do not think that there could ever be such a thing as a set of instructional actions that all teachers should use in all comprehension lessons for all students — and further research is needed to test this perspective” (Carlisle et al., 2011, 26).
Recommendations for strengthening comprehension instruction and achievement

Data from the four states featured in this report provide insights into the overall success of Reading First. In many ways, this reform initiative changed the face of early reading instruction across the nation. Many elements of Reading First made it successful. It was not an instructional experiment; it stood on a solid foundation of confirmed scientific research. By focusing on just four grade levels, Reading First concentrated on the specific learning needs of a select segment of students. It fostered partnerships between state and local educational agencies (LEAs) and universities. It required state-level external evaluations and LEA-level internal evaluations. It was a prescriptive, but voluntary reform. It had clear and well-defined guidelines, with flexibility for implementation. It served those most willing to accept its tenets.

Much of what we have learned from these four states echoes the best intentions and similar outcomes of their peer states. Though each state is unique, some common issues, even barriers to success, emerged. Here we state these issues and connect some key recommendations for policy-makers to address them in future national or state literacy initiatives.

**Issue: There is a lack of comparability of tests, consistency in standards, and accessibility to teachers of comprehension assessments used within and across states**

**Recommendation:**
Improve the trustworthiness of assessment data on comprehension achievement

- Require a common, norm-referenced comprehension achievement test to measure annual student outcomes on the many dimensions of comprehension skills and strategies; use publishers’ 50th percentiles for determining on-grade level or above performance and confirm the presence of technical psychometric features (reliability, evidence of relationship with other comprehension measures, validity, and quality of scaling).
- Require access to and use of norm-referenced screening, diagnostic, and progress monitoring assessments at the classroom level, and use publishers’ curriculum-embedded, progress monitoring assessments.
- Require classroom teachers to receive professional development in the use of screening, diagnostic, and progress monitoring assessments.

**Issue: Comprehension instruction is complex and requires on-going and substantial teacher professional development and careful monitoring of classroom practice**

**Recommendation:**
Improve the comprehensiveness and long-term utility of professional development on comprehension instruction for teachers and coaches

- Require a comprehensive, multi-staged professional development plan that prepares both new and experienced teachers and coaches to master the foundations of comprehension instruction, including understanding, teaching, and integrating strategies; advancing the development of critical thinking skills; and monitoring assessment.
• Require a minimum of 50 hours of professional development every year for every teacher, organized by grade level, with 30 hours on comprehension instruction. Graduate the professional development offerings to accommodate teachers’ increasing knowledge and skills.

• Require monitoring of classrooms using proven observation systems for recording teaching practices during comprehension lessons to guarantee that training translates to more skillful teaching and gains in students’ achievement in comprehension.

Issue: Effective teaching relies on high quality instructional materials. Teachers value comprehensive core reading programs and need adequate instructional time to ensure coverage

Recommendation:
Continue comprehensive core reading programs and increase daily instructional time for comprehension instruction

• Require high-quality comprehensive core reading programs as the essential tools for teaching comprehension.

• Set the expectation that implementation should include all of the materials in the core reading program as designed and sequenced.

• Require a study of the comprehension component of the core reading program to determine if more intensive, research-confirmed pedagogical practices are needed.

• Require an adequate block of instructional time for comprehension instruction, at a minimum of 50 percent of the total reading/language arts time (recommend 120 minutes total reading block) in 1st through 3rd grade.

• Require all teachers to receive professional development on the specific adopted, comprehensive core reading program (content and methods).

Issue: More attention has been given to research findings about the foundational skills of early reading than findings about comprehension. More focused research is needed to guide instruction and evaluation of effective practices.

Recommendation:
Acknowledge and communicate current research on comprehension instruction pedagogy and increase research on classroom practices that lead to gains in comprehension achievement

• Acknowledge the scientific research studies that support instructional practices that build proficiency in comprehension.

• Acknowledge the scientific research studies that link teacher knowledge and teacher comprehension instruction to gains in student achievement.

• Encourage additional research that links teachers’ knowledge of comprehension instruction, assessed through observations, to students’ comprehension achievement outcomes.

There is nothing easy about leading a large-scale reform. It is encouraging, though, that each iteration of reform efforts informs the next one. We have learned a great deal about the importance of leadership in general and about leadership actions that yield results. In the big picture, vision, collaboration, and passionate commitment to the goal are essential leadership elements. Leaders who put into action these recommendations will improve the odds of students’ success.


Descriptions of the features of the achievement tests used by states for assessment of comprehension

**Florida**
Stanford Achievement Test, Edition 10 is a nationally normed test with established reliability and validity credentials. It features a comprehension subtest measuring the following skills: initial understanding, interpretation, critical analysis, and awareness and usage of reading strategies in response to reading literary, informational, and functional text passages. Grade-level proficiency is set at the 40th percentile (ten percentiles lower than recommended by the test publisher).

Florida Comprehensive Achievement Test (FCAT), is a criterion-referenced test with a reading comprehension subtest with standards set for Levels 3–5 for grade 3. It features items for interpretation of words and phrases; identification of main idea, plot and purpose of reading selection; interpretation of comparisons and cause and effect relationships; similarities and differences among characters, setting and events; and validity and accuracy of information. Level 3 indicates that the student has partial success in the challenging reading content and the performance is inconsistent. “Reliability has been shown to be at .90; and test score content and concurrent validity have been established through a series of expert panel reviews and analysis” (Foorman et al., 2010, 75).

**Michigan**
The Iowa Tests of Basic Skills (ITBS, 2003, Form B) is a nationally-normed test with established reliability and validity credentials. The construct for the reading comprehension subtest for third-grade requires students to draw inferences or to generalize about what they have read for two-thirds of the test questions; other questions require students to derive word meaning from context clues (though vocabulary is a separate subtest construct). The 50th percentile is used to indicate performance at grade level.

**Pennsylvania**
The CTB Terra Nova 3 (2007) is a nationally-normed test with reliability and validity credentials. Its reading comprehension subtest measures the following skills: understanding literal meaning of text by identifying stated information, indicating sequence of events, and defining grade-level vocabulary; analyzing text by drawing conclusions, inferring relationships, identifying theme and story elements; evaluating and extending meaning by making predictions; distinguishing between fact and opinion, and reality and fantasy; and transferring ideas to other situations, judging author’s purpose, point of view, and effectiveness. Grade level proficiency is set at the 40th percentile (ten percentiles lower than recommended by the test publisher).

The Pennsylvania System of School Assessment (PSSA) is a criterion-referenced test that measures comprehension and reading skills related to understanding fiction and non-fiction and interpreting components within and between texts, literary devices in fiction and non-fiction texts, and concepts and organization of non-fiction texts. “According to traditional reliability statistics, PSSA is a reliable measurement instrument and validity
studies have shown PSSA reading scores might be considered analogous to reading comprehension scores from nationally recognized standardized (normed) achievement tests with correlations in the moderate range (.6 to .7)” (Bean et al., 2010, 14). Four cut-scores are used to designate levels of performance: below basic, basic, proficient, and advanced.

**Utah**

The Utah State Office of Education publishes the Criterion-Referenced Test (CRT) for Total Reading that covers oral language, phonics, spelling, vocabulary, comprehension, and writing. The criterion for grade-level performance is set at the proficient level based on a cut-off score determined for each grade level. “Reliability coefficients for the CRT, using the Kuder-Richardson 20 estimates of internal consistency, were reported in the low .90s . . . The cut scores for determining below basic, basic, proficient, and advanced were determined each year by a Utah State Department of Education panel. They may change by one or two percentage points, but are considered fairly stable” (Dole et al., 2010, 34).

The Iowa Tests of Basic Skills (ITBS, Form B, 2003) is a nationally-normed test with established reliability and validity credentials. The Total Reading test includes two subtests: vocabulary and comprehension (factual understanding, inference and interpretation, and analysis and generalization at the level of sentence and short passages). The proficiency level on the ITBS is set at the 50th percentile.
Descriptions of the classroom observation tools used for identifying instructional actions during comprehension lessons

Automated Classroom Observations System for Reading (ACOS-R)

**Source:** Requests for information can be made to Dr. Joanne F. Carlisle: jfcarlisle@umich.edu

**User:** Michigan used ACOS-R in 2007–08 after recruiting 6 Reading First LEAs to carry-out four observations during periods when comprehension lessons were taught. Forty-four third grade teachers participated from 19 schools.

**Comprehension lessons:** The study included those observations for comprehension instruction, including lessons that used text and represented all groupings of students (whole class, small group, or individual).

**Design:** The ACOS-R tool used a PC tablet for recording codes and handling data management. Every 5-minute interval was coded; and during this interval, other features were noted: purpose of lesson, the grouping arrangements (e.g. whole class), materials used in the lesson, instructional actions, word meaning actions, and average number of students actively engaged in the lesson. The observers were given text boxes to record notes about the content of instruction. Details on the training of observers and inter-rater reliability are found in Appendix A of the technical report (Carlisle et al., 2010, 38).

**Data analysis:** The proportions of time compiled at the lesson level were computed by summing over intervals and subintervals within the reading block (ranging between 90 minutes and two hours) that was identified by purpose (comprehension) and then averaged by dimension and sub-dimension. Note: If a lesson was taught more than once, each was considered a separate lesson.

CIERA School Change Classroom Observation Scheme


**User:** Pennsylvania used a modified form of the CIERA Scheme, renamed Reading First Observation Checklist and Rubric. From 2004 through 2007, three rounds of schools participated in the classroom observations: Group 1 by 2005 had 24 school sites and had collected 264 observations; Group 2 by 2006 had 24 school sites and had collected 273 classroom observations; and Group 3 by 2007 had 12 school sites and had collected 132 observations.

**Design:** In five-minute segments within a 45-minute instructional period during the 90-minute block for reading, it was possible to code: who is in the classroom, instructional types of groups of students, general focus (e.g., reading), specific focus (e.g., meaning of text, higher-level text comprehension, comprehension skill, comprehension strategy), material (e.g., textbook, narrative, informational, board/chart), teacher interaction (e.g., tell/give information, recitation, modeling), expected student response (e.g., reading, reading turn-taking, listening). Also observation notes could be taken within each five-minute segment.
Theoretical base: This observation system was designed to investigate multiple aspects of literacy lessons. Variables in the observation scheme consider explicit phonics skill instruction, instruction in applying word-recognition strategies to test, comprehension skill instruction, comprehension strategies instruction, lower and higher-level thinking related to text, teachers’ use of various interaction strategies (e.g., coaching, modeling, telling, and recitation), student time on-task, and active (as opposed to passive) pupil response to a reading lesson. Each variable is coded and based on current research findings.

Main purpose: Observations focused on the types of routine instructional actions of teachers. For example: recitation, when the teacher engages the students in answering questions, or responding, usually with low-level question-answer-question-answer sequence; modeling, when the teacher shows or demonstrates how to do something or how to do a process as opposed to simply explaining it; coaching, when the teacher prompts/provides support that will transfer to other situations as students are attempting to answer a question or to perform a strategy or activity (see Taylor et al., 2003, Appendix B).

Data analysis: Pre-data collection, observers participated in training sessions that included video, classroom observations, and systematic discussion and feedback about their performance. They achieved the reliability criterion of a minimum of 80 percent agreement. The average proportion of time observed across all five-minute segments within the 45-minute observation, specific to content of instruction, was calculated.

Instructional Content Emphasis (ICE-R and ICE-R2)


Users: Florida and Utah conducted the classroom observations in the morning during the reading block period in grades 2 and 3.

Definition of an observed reading instructional event: An instructional event is a distinct or unique activity in which the content, grouping, and materials are coordinated around a certain instructional component. The primary intent of the activity determines the focus of the event (i.e., the main objective of the lesson rather than the method or strategy).

The purpose: Both Florida and Utah used the ICE-R or modified ICE-R2 to determine and report the percentage of instructional time devoted to the main instructional categories, including comprehension. The other dimensions available, but not reported, are: instructional subcategory, grouping, and materials.

The descriptors for the main category of comprehension: The instructional event categorized as comprehension, included: vocabulary in the context of reading or discussion, prior knowledge/predicting, students monitoring their comprehension through activities, students or teacher monitoring through listening comprehension, instruction or use of comprehension strategy, and other instruction involving getting meaning from the text.

Observation data collected over years: Florida. In the spring of 2004 through 2007, observation data were collected during 45-minute sessions in randomly selected 103 grade 2 classrooms and 105 grade 3 classrooms, representing 408 Reading First schools. The average number of students in these classrooms was between 16 and 18. The average percentage of instructional time dedicated to comprehension was 49% for grade 2 and 57% for grade 3 (see FCRR, 2004–2007). Utah. Observation data were collected during February and March of 2005 through 2007 during three-hour reading/language arts block sessions for stratified random samples of 91 second grade classrooms and 92 third grade teachers. The numbers of schools involved were not reported. The percentage of instructional time dedicated to comprehension was 34% for grade 2 and 35% for grade 3.
Theoretical base: Instrument developers credit the Reform Up Close study for the multi-dimensional and taxonomical design of the instruction. Content validity was established by a thorough literature review and consultation with reading experts. The instructional categories and subcategories were based on numerous documented sources on instructional practice found in beginning reading instruction, including scientific research on reading instruction.

Teacher’s Instructional Practices (TIP)


User: Michigan conducted classroom observations with the use of TIP in five LEAs with 16 schools in 2005-06 and in eight LEAs with 15 schools in 2006–07. Six classrooms per grade level, K–3, were observed three to four times a year during the fall, winter, and spring (Carlisle et al., 2010, 6). The 2005–06 observations focused on the organizational structure of centers and what activities occurred in these centers. In 2006–07, the focus was on how struggling readers engaged with normal achieving readers in center activities.

Design of instrument: TIP was designed for evaluating instructional practices for the five components of effective reading instruction (phonemic awareness, phonics, fluency, vocabulary, and comprehension) in Michigan’s Reading First classrooms. TIP has the capability to observe classroom activities every five minutes for a 15 second recording of what is happening in the classroom based on the use of a code sheet. The codes focused on (a) purpose or focus of instruction (which component), (b) modality (teacher engagement with students in activities that are cognitively demanding – coded throughout the 5 minute period), (c) grouping of teacher and students, (d) engagement of students, (e) individuals present in classroom, and (f) materials used. Also the TIP Supplement can be used to observe instruction in centers (e.g., literacy centers or literacy work stations with its codes focused on: (a) the student’s activity, (b) the student’s appropriate use of materials, (c) what accountability for student work is applied, (d) the student’s focus of attention, and (e) the student’s interactions. There is also a Post Classroom Observation Teacher Feedback Form. Teachers are able to give input on what their focus of instruction was when observed. Information on the validity and reliability of TIP Observation and details on the training of observers is provided in the Manual.

Definition of comprehension: TIP definition of comprehension “understanding the text that is being read” (Manual, 11). The comprehension activities thought to be evident in classrooms are listed as follows: (1) identifying the main idea, (2) story-mapping, (3) completing a graphic organizer, (4) what-questions, (5) beginning/middle/end, (6) modeling comprehension strategies, (7) generating questions, (8) summarizing, (9) predicting, (10) making inferences, and (11) literacy discussion groups (e.g., questioning the author).

Data analysis: The central measure of TIP is time spent on reading activities. Every three five-minute segments are averaged together to create 15-minute intervals for reporting: how time is spent on different literacy purposes (e.g., comprehension); organizational structures for grouping students (whole group, small group, flexible grouping); use of specific program material; and ways teacher and students interact.
Profiles of the state participants in the pedagogical observation studies of instructional actions

Pennsylvania

Observation Tool:
   Reading First Observation Checklist and Rubric of the Modified CIERA Scheme

Duration of Observation During Comprehension Instruction Time:
   five-minute segments in 45 minute observation session

Classrooms Observed:
   2004–05: 27-Group 1 schools in 14 districts (17% of all schools), a total of 264 classroom observations (grades 2 and 3 combined) – 1 observation over two years
   2005–06: 24-Group 2 schools in 14 districts (15% of all schools), a total of 273 classroom observations (grades 2 and 3 combined) – 2 observations over two years
   2006–07: 12-Group 3 schools in 7 districts (8% of all schools), a total of 132 classroom observations (grades 2 and 3 combined) – 3 observations over three years

Student Characteristics for Grades 2 and 3:
   Group 1 schools — 75% minority; 83.6% (average) free or reduced price lunch; 11.9% limited English proficient
   Group 2 schools —75% minority; 68.6% (average) free or reduced price lunch; 11.8% limited English proficient
   Group 3 schools –77% minority; 69.8% (average) free or reduced price lunch; 11.8% limited English proficient

Michigan

Observation Tool:
   Automated Classroom Observation System for Reading (ACOS-R)

Duration of Observation During Comprehension Instruction Time:
   five minute segments in identified purpose category (coded comprehension) during the total time of the reading block (ranging from 90 minutes to 2 hours) per observation session

Classrooms Observed:
   44 third grade classrooms in 19 school districts in 2007-08

Characteristics of Teachers:
   91% female, 21% non-White, and 52% with Master’s degrees

Teaching Experience:
   Averaged 13 years of teaching experience (1-39 years)

Classroom Size:
   Averaged 22 students (range of 13-28)
Student Characteristics:
  36% minority
  73% (average) received free or reduced price lunch
  18% limited English proficient

Coverage of Comprehension Lessons (compared to all reading lessons):
  27% or 287 of the total of third grade lessons

Average Duration of Comprehension Lesson:
  19.3 minutes (17.07 SD)