
A RESEARCH AGENDA FOR IMPROVING READING COMPREHENSION

In this report, the RRSg characterizes reading comprehension in a way that the group believes will help organize research and development activities in the domain of reading comprehension. In Chapter Two, we provided a working definition for reading comprehension and outlined a framework including three core elements—reader, text, and activity—which are situated in a larger sociocultural context. Chapter Three elaborated on the elements by describing what we know about variation within them. In this chapter, the RRSg proposes a research agenda that prioritizes three specific domains of reading comprehension for future research: instruction, teacher preparation, and assessment. In making these proposals, the RRSg emphasizes the need for research that builds on what is already known, that will contribute to better theories of reading development, and that will produce knowledge that is usable both in classrooms and in policymaking arenas. To that end, this chapter describes what is already known within each of these three domains and describes areas for future work.

COMPREHENSION INSTRUCTION

Good instruction is the most powerful means of developing proficient comprehenders and preventing reading comprehension problems. Narrowly defined, comprehension instruction promotes the ability to *learn from text*. More broadly, comprehension instruction gives students access to culturally important domains of knowledge and provides a means of pursuing affective and intellectual goals. A major goal for the research agenda we propose is improving classroom instruction in comprehension, both by exploring how to ensure the broader implementation of instructional strategies known to work and by building a research base to inform the design of new instructional paradigms.

Effective teachers of comprehension enact practices that reflect the orchestration of knowledge about readers, texts, purposeful activity, and contexts for the

purpose of advancing students' thoughtful, competent, and motivated reading. Instructional decisionmaking is a dynamic and highly interactive process. To illustrate, Chapter Three described the many reader variables that are integral to proficient reading comprehension. Drawing on this literature, we characterize students along a continuum from "low need" to "high need" in terms of the instructional support they will require to become proficient comprehenders. However, this characterization of the reader must also take into account the nature of the text that the student is reading and the nature of the task that is motivating the reader. We argue that any reader can be considered high-need depending on how challenging the text is (i.e., the text is poorly written, dense, or contains a number of unfamiliar ideas) or depending on the way the reader is to demonstrate his or her understanding of the text (e.g., recall, reasoning, application, or evaluation). Finally, the teacher must consider the broad range of contextual factors that influence instructional opportunities for particular learners.

These contextual factors include, but are not limited to, community- and schoolwide factors, the culture of the classroom, the specific curriculum and instructional activities in which students are engaged, and the nature of the interaction between teacher and students as well as among students. Similarly, a student who appears to be a high-need reader when the reader variables are considered in isolation may, in fact, be very successful in an instructional setting in which the teacher attends to this student's needs while selecting texts, designing tasks for him or her, and deciding how to structure the context to best support the student's participation and learning.

To maximize the possibility that research will yield *usable knowledge*, instructional research, regardless of the method employed, needs to attend to each of these elements of reading comprehension. Careful descriptions of both the texts used in the research and the specific nature of the task(s) for which students are using reading in the specific context of instruction need to accompany careful descriptions of the participants. The context includes, but is not limited to (in the case of classroom-based research), general classroom conditions (reported in Pressley et al., 2001) that set the stage for effective instruction, the specific nature of the instructional activity or activities in which the learner is engaged, and the specific nature of the support that teachers, peers, and instructional tools (e.g., computers) provide.

What We Already Know About Comprehension Instruction

The RAND Reading Study Group's prioritization of comprehension instruction set forth in the agenda presented in this chapter is based upon a fairly well-articulated knowledge base.

1. Instruction that is designed to enhance reading fluency leads to fairly significant gains in word recognition and fluency and to moderate gains in comprehension.

A substantial amount of practice over an extended period of time is required for a reader to acquire fluency. Most fluency instruction consists of the repeated reading of the same text and uses many techniques. Sometimes the repeated reading practice is done independently; sometimes the reader is assisted by a teacher who provides corrective feedback; sometimes the reader listens to the text before practicing or reads along with a teacher or a tape. Some studies have incorporated partner reading in which peers, not a teacher, give feedback.

The National Reading Panel (NRP) (2000) examined the wide-ranging literature on repeated reading. A meta-analysis of 14 studies indicated that the mean weighted effect size of comparisons of one or another of these techniques versus a no-instruction control varied depending on what type of outcome measure was examined. It was largest (.55) when the outcome measure was word recognition, next largest (.44) with a fluency measure, and smallest (.35) with a comprehension outcome measure. The NRP found that repeated reading was effective for normal readers through grade 4 (there were no studies of normal readers beyond grade 4) and for students with reading problems throughout high school.

The NRP also examined three other sets of studies: studies looking at the immediate effect of different programs of repetition and feedback during oral reading on the reading performance of a specific passage (these studies did not attempt to assess transfer to uninstructed passages); studies using small groups of students; and studies that compared the efficacy of two different oral reading procedures. All three sets of studies corroborated the findings of the meta-analysis, indicating the value of repeated reading. No conclusions could be drawn about the relative effectiveness of independent repeated reading and guided oral reading practice or of any other two procedures, such as reading with or without feedback. One exception to this conclusion of no differences comes from a study by Rashotte and Torgesen (1985). They compared passages that either shared or did not share many words with the outcome measures. They noted gains when the passages shared words but no gains when the passages did not share words. This result suggests that very poor readers probably at least learn words from repeated reading (Faulkner & Levy, 1999). Most studies have found that reading interconnected text is necessary for effective fluency instruction, but one recent study (Tan & Nicholson, 1997) has indicated that reading of isolated word lists also leads to increased fluency.

Several studies have indicated that these repeated-reading techniques are feasible for classroom use (Dixon-Krauss, 1995; Rasinski, 1990). No extensive

preparation is needed to use these techniques successfully (Reutzel & Hollingsworth, 1993). Studies dealing with readers with learning disabilities have found that peer tutoring can be successfully incorporated into the instruction (Mathes & Fuchs, 1993; Simmons et al., 1994).

Other studies have assessed the effect of simple practice in reading, such as Sustained Silent Reading. However, merely encouraging students to read extensively did not result in improved reading, according to the findings of a meta-analysis (NRP, 2000). It is thus not clear whether there are conditions under which practice in reading would promote fluency and comprehension.

Another approach to promoting fluency involves ensuring that proficiency and fluency are acquired during instruction in all components of reading, starting with letter knowledge and phonemic awareness and moving to decoding and word recognition (Berninger, Abbott, Billingsley, & Nagy, in press; Wolf & Katzir-Cohen, 2001). Berninger, Abbott, Brooksher, Lemos, Ogier, Zook, & Mostafapour (in press); and Wolf & Katzir-Cohen (2001) have developed intervention programs that address specific component skills, foster linkages among all relevant systems—orthographic, phonological, semantic, and morphological—and emphasize fluency at each step. These programs are very new, and no data on their success in promoting fluency are currently available.

2. Instruction can be effective in providing students with a repertoire of strategies that promote comprehension monitoring and foster comprehension.

Because meaning does not exist in text, but rather must be actively constructed, instruction in how to employ strategies is necessary to improve comprehension. To construct meaning, students must monitor their understanding and apply strategic effort. We know that students who are good comprehenders read for a purpose and actively monitor whether that purpose is being met. They notice when something they are reading is incongruous with their background knowledge or is unclear, then they take action to clarify their understanding, such as rereading or reading ahead. They may also stop periodically when reading to summarize what they have read as a way to check their understanding.

To further enhance comprehension, good comprehenders also use strategies that help them retain, organize, and evaluate the information they are reading. Among these strategies is a well-defined set that we know, as a result of rigorous investigation and replication, leads to improved comprehension when employed by readers. This set of strategies includes concept mapping, question generating, question answering, summarizing, and story mapping as delineated in the NRP report (2000). Additional strategies investigated in non-experimental

studies that may also prove beneficial to students include mental imagery, knowledge activation, mnemonics, and expository pattern identification.

Judging by the experimental studies reviewed by the NRP (2000), we know that engaging students in identifying the big ideas in a text and in graphically depicting the relationships among these ideas improves their recall and comprehension of text. We also know that in grades 3–5, engaging students in elaborative questioning improves their comprehension of text read during instruction and their comprehension of new text read independently. Similarly, teaching students in grades 3–9 to self-question while reading text enhances their understanding of the text used in the instruction and improves their comprehension of new text. Studies conducted in the upper elementary grades indicate that learning to paraphrase text, identify the gist of a text, and identify and integrate the big ideas in a text enhance the recall of text and the capacity to understand new text. Teaching students in grades 3–6 to identify and represent story structure improves their comprehension of the story they have read. In the case of this strategy, there was no evidence that the strategy transferred to the reading of new stories and improvement was more marked for low-achieving readers.

3. The explicitness with which teachers teach comprehension strategies makes a difference in learner outcomes, especially for low-achieving students.

Understanding the nature of the reading comprehension problems experienced by many students who are low achieving has helped in developing instructional approaches that enhance the comprehension abilities of these students. An important instructional strategy for these learners consists of making instruction very explicit. Explicit instruction provides a clear explanation of the criterion task, encourages students to pay attention, activates prior knowledge, breaks the task into small steps, provides sufficient practice at every step, and incorporates teacher feedback. It is particularly important for the teacher to model the comprehension strategies being taught. Careful and slow fading of the scaffolding is important.

Sometimes this explicit instruction is helpful for low-achieving students but is superfluous for normal readers (Wong & Jones, 1982). Sometimes improvement occurs not because of the specific strategies being taught but because students have been actively interacting with the texts. This active interaction triggers the use of strategies that inactive learners possess but do not normally use.

Explicit instruction generates the immediate use of comprehension strategies, but there is less evidence that students continue to use the strategies in the classroom and outside of school after instruction ends (Keeny, Cannizzo &

Flavell, 1967; Ringel & Springer, 1980) or that they transfer the strategies to new situations.

Recent studies have underscored the importance of teacher preparation when the goal is to deliver effective instruction in reading comprehension strategies (Duffy et al., 1987; Brown et al., 1996). This is especially important when the students are low performing. Implementing a direct approach to cognitive strategy instruction in the context of the actual classroom has proven problematic. Proficient reading involves much more than using individual strategies. It involves a constant, ongoing adaptation of many cognitive processes. Successful teachers of reading comprehension must respond flexibly and opportunistically to students' needs for instructive feedback as they read. Lengthy, intensive teacher preparation is effective in helping teachers deliver successful strategy instruction that has improved student outcomes on reading comprehension tests.

4. There are a number of working hypotheses about the role of instruction in explaining and addressing the problems of poor comprehenders.

One of the most vexing problems facing middle and secondary school teachers today is that many students come into their classrooms without the requisite knowledge, skills, and dispositions to read the materials placed before them. These students are, for one reason or another, poor comprehenders. Poor comprehenders are students who can neither read nor demonstrate satisfactory understanding of texts appropriate for their grade level. Many teachers are frustrated by what they see as an ever-increasing number of students who are poor comprehenders.

Instructional research with poor comprehenders has been motivated by a particular set of hypotheses about impediments to comprehension. Some of these hypotheses suggest that the problems of poor comprehenders are an outgrowth of differential instruction; that is, these students have been denied the kinds of instruction that advance reading comprehension. This hypothesis is particularly relevant for students who have a history of reading problems (e.g., decoding problems in grades 1 and 2). For example, McDermott and Varenne (1995) documented that teachers working with high-achieving students focused on higher-order thinking with text and communicated clearly that the purpose of reading was understanding. In contrast, these same teachers, when working with low-achieving students, focused on low-level factual reading, interrupted children's reading more frequently than their errors would justify (see also Shake, 1986), and communicated little about comprehension as the goal of reading. A corollary to this hypothesis is that students with a history of reading challenges read less text; hence, they accrue less background knowledge to bring to the reading of new text.

Research has indicated, however, that specific instruction, for example, pre-reading, can improve poor comprehenders' understanding of a difficult text. Researchers have used instructional scripts that provide students with essential background knowledge, key concepts, and vocabulary (Graves, Cooke, & LaBerge, 1983) or have activated students' background knowledge through extended discussions (Langer, 1984). Researchers have also used such activities as story structures or graphic organizers to provide scaffolding for improved comprehension of a selected text (NRP, 2000). Pre- and post-writing activities have also been used as effective instructional activities to promote comprehension for low-achieving readers. These instructional activities effectively address the problem of poor comprehension by providing this sort of instructional scaffolding to help low-achieving readers comprehend texts above their independent reading level.

In addition, poor comprehenders can be guided to effectively employ a number of strategies to improve their understanding of text. For example, researchers have helped poor comprehenders draw inferences by using a pre-reading strategy in which they activate attention and prior knowledge or by using particular strategies in the course of reading, such as restating information from the text (Chan et al., 1987; Idol-Maestas, 1985; Schumaker et al., 1982).

The nature of the strategy taught seems less significant than the role that strategy instruction plays in engaging the reader in active interaction with the text (Chan & Cole, 1986). A synthesis of the research literature on teaching comprehension strategies to students with learning problems (Gersten, Fuchs, Williams, & Baker, in press) indicates that successful comprehension instruction for the poor comprehender is characterized by explicit modeling by the teacher, additional opportunities for practice with feedback, skillful adjustments to the learner's level, and the reader's mindful engagement with the purposes for reading.

5. The role of vocabulary instruction in enhancing comprehension is complex.

As we described earlier in this report, vocabulary knowledge is strongly linked to reading comprehension (Freebody & Anderson, 1983), and there is reason to believe that vocabulary knowledge is an especially important factor in understanding the reading problems experienced by second-language learners (García, 1991; Laufer & Sim, 1985). However, this relationship between vocabulary knowledge and comprehension is extremely complex, confounded, as it is, by the complexity of relationships among vocabulary knowledge, conceptual and cultural knowledge, and instructional opportunities.

These complexities speak to the unique and significant role that instructional research can play in enhancing the education field's understanding of the role

of vocabulary knowledge in comprehension. The NRP (2000) found that direct instruction of vocabulary improved reading comprehension. The effects of extensive reading on vocabulary growth are, however, debatable. The NRP did not find compelling evidence that programs that are designed to increase independent reading, such as Sustained Silent Reading, promoted vocabulary growth. Nevertheless, there is a powerful correlational relationship between the volume of reading and vocabulary growth among first-language learners (Stanovich & Cunningham, 1992), and “book-flood” studies (in which children are provided with numerous books for use at school or at home) with second-language learners have shown powerful effects (Elley, 1991). Further, a wealth of evidence relates children’s oral language experiences to subsequent vocabulary growth (Dickinson & Tabors, 2001). Much of this evidence comes from studies of the effects of homes and preschools on language development. Less is known about the effects of school-based oral language activities and vocabulary learning and growth, although Meichenbaum and Biemiller (1998), among others, have argued that the fourth-grade slump cited earlier in this report is caused, at least in part, by the failure of schools to promote oral language development while children are still working on the mechanics of reading.

Much of the instructional research in vocabulary has been designed to document, or compare, the effectiveness of different methods of teaching individual words. Although some generalizations can be made about the characteristics of effective vocabulary instruction (Stahl & Fairbanks, 1986), the number of studies that have directly examined the effects of vocabulary instruction on reading comprehension is still relatively small. Some of the strongest demonstrations of the effects of vocabulary instruction on reading comprehension—the work of Beck and her colleagues (e.g., Beck, Perfetti, & McKeown, 1982; McKeown, Beck, Omanson, & Pople, 1985)—used rather artificial texts heavily loaded with unfamiliar words. Little, if any, research addresses the question of which conditions—the types of texts, words, readers, and outcomes—can actually improve comprehension.

Effective vocabulary instruction presupposes choosing the right words to teach. This is another area in which more research is needed. How does a teacher choose which words to teach? What are the instructionally relevant subcategories of words? Graves (2000) and others have suggested some distinctions that must be considered, such as the difference between teaching new concepts and teaching new labels for familiar concepts, or the difference between teaching students to recognize in print words already in their oral vocabularies and teaching them words not yet in their reading or oral vocabularies. Nation (1989; Laufer & Nation, 1999) has offered another instructionally relevant way to categorize words—as high-frequency words, domain-specific technical vocabulary,

low-frequency words, or high-utility academic vocabulary. Although such distinctions are undoubtedly crucial in making instructional decisions, there is still little documentation of how well teachers can use such categories or of the actual effect of such categories on the effectiveness of vocabulary instruction.

Some vocabulary researchers (e.g., Laufer & Sim, 1985) have stressed the importance of high-frequency words for learners of English, because a relatively small number of words constitute the bulk of words encountered in text. However, the most effective methodology for teaching high-frequency words still needs to be explored, given that such words are also the most likely to have multiple meanings. Others have stressed the importance of focusing on words intermediate in frequency—not so frequent that they are already known by almost everyone, yet frequent enough to be worth teaching. Much remains to be learned about identifying these words and about the effectiveness of instructional approaches that focus on such words. Another dimension of choosing words for instruction has to do with the relationships among instructed words. Materials for learners of English as a second language often group words on the basis of meaning. However, some evidence suggests that teaching words in groups that are highly similar in meaning is a hindrance, rather than an aid, to learning (Tinkham, 1993; Waring, 1997).

Teaching individual words presupposes some sort of explanation of their meanings, which is most likely to be in the form of a definition. Although some research has explored the effectiveness of different types of definitions (Fischer, 1994; McKeown, 1993; Scott & Nagy, 1997), relatively little is known about this area. To our knowledge, for example, no one has explored the question of whether different types of definitions are appropriate for different types of words or for different stages of word learning (e.g., initial exposure versus consolidation and refinement of word knowledge). Research could help illuminate what knowledge, skills, and abilities best allow learners to benefit from definitions or, more generally, from vocabulary instruction (e.g., dictionary skills, metalinguistic abilities, language proficiency levels). There is little question that one component of proficient comprehension is the ability to cope with any unfamiliar words encountered during reading. Readers need to be able to use the information provided by context, by morphology (word parts), and by dictionaries or other reference materials and to coordinate information from these sources.

In a recent meta-analysis, Fukkink and de Glopper (1998) found that instruction in the use of context improved students' ability to use contextual clues to figure out word meanings. However, on the basis of a similar meta-analysis, Kuhn and Stahl (1998) argued that such instruction was not demonstrably more effective than simple practice. Instruction in the use of morphology and definitions has been less thoroughly investigated than instruction in the use of context. The

possibility of online dictionaries and other word-learning aids opens up additional areas for research.

The effectiveness of context for second-language learners is still a matter of debate. A variety of evidence indicates that second-language learners have more difficulty using context than do native-language learners (e.g., Nagy, McClure, & Montserrat, 1997). However, second-language learners who face the task of simply learning new labels rather than learning new concepts may be at a relative advantage.

Research is also needed on what makes some students more effective independent word learners than others. Some of the contributing factors, such as language proficiency and existing vocabulary and background knowledge, are obvious. Phonological processing ability contributes to vocabulary learning, especially for second-language learners (Eviatar & Ibrahim, 2000; Muter & Diethelm, 2001). It also seems likely that a variety of metalinguistic abilities contribute to vocabulary learning (Nagy & Scott, 2000).

A number of vocabulary researchers have expressed the opinion that “word consciousness” or “word awareness” may be an important element in promoting vocabulary growth (Graves, Watts-Taffe, & Graves, 1998). As yet, no research has measured such a construct, let alone documented its effect on vocabulary learning. One reason that word consciousness and its effects on vocabulary growth are not well understood is that various constructs could fall under this heading but they are not all necessarily related to one another. For example, the concept of words (Roberts, 1992), morphological awareness (Anglin, 1993; Carlisle, 1995), word schemas (Nagy & Scott, 1990), word play, and an appreciation for effective word choice (Scott & Nagy, 1997) could all fall under the term *word consciousness*. Little is known, however, about how these constructs relate to one another or to vocabulary growth.

Various aspects of word consciousness may be crucial to strategies for independent word learning. Morphological awareness is undoubtedly involved in using word parts to make inferences about the meanings of new words. Word schemas—knowledge of what might constitute a possible meaning for a word—could be an important part of making inferences about new words encountered in context (Nagy & Scott, 1990) and may also contribute to the effective use of definitions. For example, Fischer (1994) speculates that one factor limiting the effectiveness of second-language learners’ use of bilingual dictionaries is the expectation that there will be one-to-one mappings between the meanings of words in two languages.

For speakers of Spanish who are learning English (or vice versa), a specific type of word awareness—awareness of cognate relationships—may be especially important. Many words in the vocabulary of literate or academic English are

similar in both form and meaning to everyday Spanish words (e.g., *tranquil/tranquilo* and *pensive/pensivo*). Bilingual students differ in their ability to recognize such relationships (Nagy, García, Durgunoglu, & Hancin-Bhatt, 1993; García & Nagy, 1993), and the ability to recognize such relationships appears to be associated with more effective reading strategies (Jiménez, García, & Pearson, 1996).

Each of the four components of a vocabulary curriculum outlined by Graves (2000)—teaching individual words, encouraging wide reading, teaching word-learning strategies, and promoting word consciousness—is likely to make an important contribution to students’ long-term vocabulary growth and, hence, to their reading comprehension. However, in addition to our incomplete knowledge about each component, we know extremely little about their relative contribution and how they interact with one another.

6. Teachers who provide comprehension strategy instruction that is deeply connected within the context of subject matter learning, such as history and science, foster comprehension development.

As we described earlier in this chapter, the NRP evidence suggests that teaching such reading strategies as questioning, summarizing, comprehension monitoring, and using graphic organizers facilitates reading comprehension. Several quasi-experimental investigations show that when the strategy instruction is fully embedded in in-depth learning of content, the strategies are learned to a high level of competence (Guthrie, Van Meter, Hancock, Alao, Anderson, & McCann, 1998). If students learn that strategies are tools for understanding the conceptual content of text, then the strategies become purposeful and integral to reading activities. Connecting cognitive strategies to students’ growing knowledge of a content area enables students to both increase their awareness of and deliberately use the strategies as means for learning (Brown, 1997) in microgenetic analyses of instruction. Unless the strategies are closely linked with knowledge and understanding in a content area, students are unlikely to learn the strategies fully, may not perceive the strategies as valuable tools, and are less likely to use them in new learning situations with new text.

Integrating strategy instruction into content domains requires a balance. The priority of instructing for reading comprehension must be balanced with the priority of teaching the content area itself. Teachers can help students learn that gaining new ideas, increased understanding, and literary experience is an aim of reading and that strategies are a powerful way to accomplish that aim. This information helps students use strategies reliably when they are appropriate. If comprehension strategies are taught with an array of content and a range of texts that are too wide, then students will not fully learn them. If strategies are taught with too narrow a base of content or text, then students do not have a

chance to learn how to transfer them to new reading situations (Rosenshine & Meister, 1994). The optimal balance enables students to learn that strategies are an important means for understanding but are not the main point of reading activities. The main purposes for reading are gaining meaning and gaining knowledge.

An important aspect of strategy development is to enable students to become self-initiating (Alexander & Murphy, 1998), according to several reviews of empirical literature. Students who spontaneously apply a strategy, such as questioning, when it is sensible will improve their comprehension. Thus, to be effective comprehenders, students must have motivation, self-efficacy, and ownership regarding their purposes for reading and their strategies. Teaching strategies integrated with content enables students to become proficient, self-regulating strategy users.

7. Using various genres of text (i.e., narrative and informational text) diversifies instructional opportunities, as assessed by teacher and student discourse.

A knowledge of text structure is an important factor in fostering comprehension. Students with some knowledge of text structure expect texts to unfold in certain ways. Even before they enter school, children have a rudimentary sense of narrative structure. The first texts they are introduced to in school are narrative in structure, which allows an easy transition from oral to written language (Van Dongen & Westby, 1986). In school, children are also introduced to expository text, which is more complex, diverse, and challenging.

Readers who are unaware of structure do not approach a text with any particular plan of action (Meyer, Brandt, & Bluth, 1980). Consequently, they tend to retrieve information in a seemingly random way. Students who are aware of text structure organize the text as they read, and they recognize and retain the important information it contains.

Simple exposure to stories is helpful, but explicit instruction is valuable. Children are taught to ask themselves generic questions that focus on the principal components of a story, which helps them identify the relevant and important information in stories (Mandler & Johnson, 1977; Stein & Glenn, 1979; Williams, 1993). In addition to their value as an organizational guide to the text structure, the questions enhance the active processing of the text, thus qualifying the generic questions as comprehension-monitoring instruction. Such instruction improves students' ability to see relationships in stories, answer comprehension questions, and retell the stories in a focused fashion. The positive effects of an intervention are most likely to accrue on measures closely aligned with the specific instruction provided. The effect of interventions that teach the use of text structure is not as strong on transfer measures. Although stories constitute

the bulk of reading material for instruction in the early grades, a case for greater inclusion of other text genres has been made (Duke, 2000; Pappas & Barry, 1997). Such inclusion will allow instruction that more closely matches the demands of reading in later grades.

As readers progress through school, the demands placed on them change. At about grade 4, they are expected to read expository material in content instruction. Because expository text is often dense with information and unfamiliar technical vocabulary, students must perform complex cognitive tasks to extract and synthesize its content (Lapp, Flood, & Ranck-Buhr, 1995). Expository text involves relatively long passages, less familiar content, and more complex and varied structures (Armbruster & Anderson, 1984). Explicit teaching about structure enables students to differentiate among common structures and to identify the important information in a text in a coherent, organized way (Armbruster & Armstrong, 1993).

Various instructional techniques have been used to help students comprehend expository text, including teaching them to use generic questions to self-question (Wong & Jones, 1982), to use mapping to analyze the text (Swanson, Kozleski, & Stegink, 1987; Boyle & Weishaar, 1997), to summarize (Nelson, Smith, & Dodd, 1992), and to employ other simple strategies. These interventions were effective.

A body of research exists on methods for adapting or modifying texts (e.g., Beck, McKeown, Sinatra, & Loxterman, 1991) to make them easier to comprehend. This literature is important, but it does not address the issue of helping students understand the texts they may encounter in their content area classes and on high-stakes tests.

8. Teachers who give students choices, challenging tasks, and collaborative learning structures increase their motivation to read and comprehend text.

For students from grade 1 to grade 12, classroom activities that enable and encourage them to take responsibility for their reading increase their reading achievement. For example, extensive observations of classroom instruction for primary students show that when teachers provide challenging passages for reading, students exert effort and persistence. And when students have a limited, but meaningful, choice about the learning activity, such as which part of a text to read, they invest greater energy in learning than when the tasks are always prescribed by the teacher (Turner, 1995).

With elementary and middle school students, quasi-experimental and structural equation modeling studies have shown that teachers who provide meaningful choices and autonomy increase students' motivation to read and to expend effort to gain knowledge from text (Reeve, Bolt, & Cai, 1999). The ex-

planation for the benefit of autonomy support for reading comprehension is that students become more-active learners when teachers provide a minimal, but meaningful, choice in the topics, texts, activities, and strategies for learning. For example, when given a choice of two books for a comprehension activity, students will choose the one that interests them. This interest deepens the students' thinking and their use of strategies and background knowledge during reading (see Schiefele, 1999, for a review of experimental evidence). High interest, derived from choice, leads to high comprehension.

The roles of motivation and engagement as links between instruction and achievement have been documented by many investigators (Skinner, Wellborn, & Connell, 1990; see Guthrie & Wigfield, 2000, for a review of empirical research). In brief, the most predictive statistical models show that engagement is a mediator of the effects of instruction on reading achievement. If instruction increases students' engagement, then students' achievement increases. In this literature, engagement refers to a combination of the following: (a) the use of cognitive strategies; (b) the presence of an intrinsic motivation to read; (c) the use of background knowledge to understand text; and (d) the social interchanges in reading, such as discussing the meaning of a paragraph or the theme of a narrative. Therefore, instruction affects reading comprehension outcomes through the avenue of active engagement in frequent, thoughtful reading for understanding.

9. Effective teachers enact a wide range of instructional practices that they use thoughtfully and dynamically.

Most people do not realize how complex teaching is. Effective teachers do more than teach specific strategies or make available to students a wide variety of texts. Indeed, effective teachers of reading engage in a diverse array of instructional practices (NRP, 2000; Pressley et al., 2001; Taylor, Pearson, Clark, & Walpole, 1999). This panoply of practices results in a complex environment in which comprehension can be fostered.

A review of studies of effective teachers reveals some of these important instructional practices and activities. For example, effective teachers establish a complex set of organizational and management routines in their classrooms, which they use to ensure a minimal amount of disruption and a maximal amount of time-on-task. Indeed, almost all of the time in the classrooms of effective teachers is spent on instruction. In addition, effective teachers provide an atmosphere of support and encouragement. In their classrooms, readers feel comfortable taking risks and are expected to achieve.

Effective teachers also use a variety of instructional practices that relate more specifically to reading comprehension. For example, effective teachers ask high-level comprehension questions that require students to make inferences

and to think beyond the text. Effective teachers help readers make connections between texts they read and their personal lives and experiences. Effective teachers use small-group instruction to meet the individual needs of their readers. Effective teachers provide their readers with practice reading materials at their appropriate reading level. Effective teachers of young readers monitor progress in reading by using informal assessments.

One critically important, but thorny, aspect of teaching reading in general and comprehension in particular is the appropriate balance between teaching skills and using literature. Over the last 20 years, the reading field has vacillated between the two—with fierce opposition between those recommending one or the other. However, the choice does not seem to concern most teachers. In a survey of teacher practices, Baumann, Hoffman, Moon, and Duffy-Hester (1998) reported that teachers believed both to be essential for good teaching. In fact, teachers reported that they taught skills *and* extensively used literature.

10. Despite the well-developed knowledge base supporting the value of instruction designed to enhance comprehension, comprehension instruction continues to receive inadequate time and attention in typical classroom instruction across the primary and upper elementary grades.

In the late 1970s, research revealed that teachers devoted only 2 percent of the classroom time designated for reading instruction to actually teaching students how to comprehend what they read (Durkin, 1978–79). Twenty years later not much has changed in the upper elementary (Pressley, 2000) or primary grades (Taylor et al., 1999). For example, Taylor and colleagues documented the limited opportunities that children in grades K–3 had to develop knowledge and thinking even in the context of schools that were effectively “beating the odds”—that is, schools that were realizing higher early reading achievement gains than would be predicted given the demographics of their student populations. Using survey and classroom observation data, they reported that only 16 percent of the teachers in the entire sample emphasized comprehension.

Despite the hypothesized role that inexperience with informational text plays in the fourth-grade slump (Chall, Jacobs, & Baldwin, 1990), and despite evidence that some young children prefer to read informational text (Pappas & Barry, 1997), primary-grade classrooms have a significant dearth of informational texts (Duke, 2000). Beginning in grade 4 and throughout their formal education, students will spend the majority of their time reading expository text, yet instruction in grades 1–3 primarily uses narrative text. Recently a plethora of engaging informational texts, written for primary-grade students, has become available. However, these books are not yet in sufficient supply in primary

classrooms, and primary-grade teachers have not yet balanced teaching reading for informational and narrative texts.

What We Need to Know About Comprehension Instruction

What specific issues of educational urgency exist, and how can we formulate the most promising research directions for addressing them? We start with four problem statements related to low-achieving students and one concerning second-language readers, then turn to issues of instructional design relevant to the entire student population.

1. For poor comprehenders in the general education setting, would focusing more time on comprehension instruction while using currently available curricula and instructional strategies generate adequate gains?

Studies of classroom practice are unanimous in noting the scarcity of time devoted to comprehension instruction. Neither in the primary grades, when the focus of reading instruction is typically word reading, nor in the middle elementary grades do teachers spend much time helping students learn how to approach complex texts strategically. Although the current approaches to teaching comprehension are neither adequately rich nor research-based, the possibility exists that they are adequate to address comprehension problems for some learners, if sufficient time is devoted to instruction.

2. For poor comprehenders in the general education setting, how should time and instructional emphasis be allocated among (a) promoting fluency, (b) teaching vocabulary, (c) instructing students in the use of reading strategies, (d) providing extensive reading of informational and literary text, (e) encouraging writing based on reading, (f) using multimedia to support content learning, and (g) using computer programs to improve reading skills?

Some evidence supports the efficacy of promoting fluency, teaching vocabulary, teaching strategies, promoting wide reading, and encouraging writing based on reading in promoting comprehension. In contrast, little evidence supports the efficacy of using multimedia for content learning or computer programs for skill development, but these practices are widely implemented. Teachers need guidance, which is totally absent in the available research literature, about how to combine and prioritize these various instructional approaches in the classroom.

3. How do teachers identified as effective with low achievers create, administer, and use reading assessments that are related to curricular goals and useful for informing instruction across grade levels and across diverse populations of students? Further, how do effective teachers determine the

knowledge, skills, and dispositions that diverse readers bring to reading activities?

Studies of effective teachers have been informative about aspects of instruction that work well to improve comprehension. We know little, though, about effective teachers' selection, use, and interpretation of assessments to inform their practice. Such practice-based wisdom, if it is indeed available, could be useful if verified and disseminated more widely.

4. For low-achieving students in high-poverty schools, what organization of instructional practices is beneficial: (a) instruction in word recognition and fluency, (b) access to and use of an abundance of content and literary texts, (c) explicit teaching of reading strategies, (d) explicit teaching of vocabulary and the use of vocabulary knowledge in reading, (e) out-of-school literacy pursuits to enhance reading development, (f) writing based on reading, and (g) opportunities for multimedia links to support reading and writing tasks?

As noted above, little evidence supports the efficacy of giving instruction in word recognition and fluency, teaching vocabulary, teaching strategies, promoting wide reading, and encouraging writing based on reading in promoting comprehension. Also, there is little evidence concerning the efficacy of computer programs for skill development or of out-of-school literacy supports, but these practices seem promising. Teachers working in high-poverty schools need guidance on how to combine and prioritize various instructional approaches in the classroom. In particular, they need to learn how to teach comprehension while attending to the often poor word-reading skills their students bring to the middle and later elementary grades.

5. For students who are learning English as a second language, how should time and instructional emphasis be allocated among (a) giving instruction in word recognition and fluency, (b) teaching vocabulary, (c) instructing about strategies, (d) providing extensive reading of informational and literary text, (e) encouraging writing based on reading, (f) using multimedia to support content learning, (g) using of out-of-school literacy pursuits to enhance reading development, and (h) using computer programs to improve reading skills?

Teachers of English-language learners, like teachers of poor comprehenders in the general education setting and teachers working in high-poverty schools, have available a number of instructional techniques and strategies that research has shown to be effective and additional techniques that are endorsed by the wisdom of practice. However, selecting among these various instructional practices for particular students and groups of students and devoting

appropriate amounts of time to using the practices remain a challenge. And research offers little guidance.

6. Under what conditions does instruction about strategies to improve reading comprehension actually lead to students' using the strategic approaches for various texts and tasks in diverse contexts and at different age levels? What specific instructional activities, materials, and practices are related to effective comprehension and to the engagement of students from various cultural and linguistic backgrounds at varying grade levels?

It is well documented (NRP, 2000) that students can be taught to use strategies to advance their ability and inclination to independently learn from text.¹ Furthermore, evidence suggests that a relatively small set of strategies appears to be consistently effective across diverse populations of students, with diverse forms of text, and for diverse tasks that the reader is to accomplish. Finally, and perhaps most important, there is evidence that the power of strategy instruction is the extent to which strategies are taught in the service of interpreting text, not as ends in and of themselves. But this robust knowledge base is still incomplete.

7. How can excellent, direct comprehension instruction be embedded in content instruction that uses inquiry-based methods and authentic reading materials?

Contemporary national benchmarks in science call for instruction to be inquiry-based. The standards in history call for students to learn the practices of historical analysis, including the use of primary documents. Contemporary language arts standards call for students, at all ages, to read authentic literature across genres (e.g., novels, memoirs, interviews) and to write in various genres. Web-based technology affords students the opportunity to access numerous sources of information. All of these opportunities provide potentially powerful contexts in which students can learn to interpret text and can learn how to learn from text. However, with the exception of a few studies (Brown & Campione, 1994; Guthrie, et al., 1998), we know little about how these instructional contexts lead to improved reading comprehension or about how specific teacher practices in these contexts can lead to improved comprehension. Specifically:

- What is the role of direct instruction in specific comprehension-monitoring and comprehension-fostering strategies in an inquiry-focused learning environment?

¹This statement applies to upper elementary through adult education. We have a much leaner knowledge base regarding strategy instruction in the preschool and primary grades.

- How can activities that are designed to promote knowledge-building be extended to enhance self-regulated reading?
 - What role does experience with a diverse array of texts, used in the context of subject-matter learning, play in promoting thoughtful, competent, and motivated readers?
8. How do we ensure that all children know the vocabulary they will encounter in content area and advanced texts?

A number of significant researchable issues are related to the role of vocabulary in enhancing comprehension. We focus here on four subsets: (a) selecting the words to teach, (b) teaching strategies for learning words independently, (c) fostering word consciousness, and (d) examining the interplay between different components of a vocabulary curriculum. (See Appendix A for a specific description of these issues.)

9. How do national, state, and local policies and practices facilitate or impede the efforts of teachers to implement effective comprehension instruction?

The policy literature and teacher journals are filled with examples of how policy changes improved or undermined educational effectiveness. There are notable examples of successfully implemented policies imposed or encouraged by districts and states that changed instructional practices in the domain of word reading. However, a systematic analysis of the effect of these and other policies on comprehension instruction has not been undertaken.

TEACHER EDUCATION AND PROFESSIONAL DEVELOPMENT IN READING COMPREHENSION

An important goal of research on reading comprehension is the larger goal of improving students' reading proficiency. This goal, however, is mediated by at least two critical variables. First, the research must be translated into appropriate *instruction*. Second, *teachers* must enact that instruction. Regardless of the quantity and quality of research-based knowledge about comprehension, unless teachers use that knowledge to improve their instruction, students' reading achievement will not improve. In other words, as Sykes (1999) argued, recent advances in research-based best practices have an effect only to the extent that teachers adopt those practices.

There is reason to question whether teachers use research-based best practices to teach comprehension or other subject areas. Cuban (1993) has argued that, in general, although teachers have made some changes in their classrooms over the last 100 years, the basic forms of instruction have not changed. The recent Third International Math & Science Study (TIMSS) Videotape Classroom Study

(Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999) corroborated Cuban's observations and conclusions. In the TIMSS study, researchers found that most American teachers, even those who say they use reform models, still teach using traditional practices. Hiebert and Martin (2001) showed that teachers distort much knowledge about mathematics reform to make it consistent with their existing practices. These researchers found that true changes in teaching practice based on research were rare among American teachers.

Whereas some researchers have questioned the extent to which teachers use research-based best practices in their instruction, other researchers have pointed to teacher quality as one of the most critical variables in student achievement. Teacher quality is defined in many ways, from advanced degrees to deep subject matter knowledge to deep pedagogical knowledge (Shulman, 1986). Whatever way it is defined, it is clear that the expertise of the teacher matters, and it matters a lot. In an extensive review of the research on teacher quality and student achievement, Darling-Hammond (2000) found that teacher quality and expertise consistently and accurately predicted student achievement. Additionally, Sykes (1999) pointed to the rather weak effects of efforts at systemic reform without adequate professional development. Sykes argued that early systemic reform efforts—focused on new assessments, new curriculum frameworks, and teaching standards—are not enough to improve student achievement. Research has demonstrated that these efforts need to be accompanied by strong professional development. More-recent systemic reform efforts have focused squarely on the teacher as the center of reform.

One particularly puzzling aspect of school reform is that despite the key role ascribed to teachers when explaining why reforms fail (Cohen & Ball, 1990; Cremin, 1965; Darling-Hammond, 1990), we continue to craft fairly minimal roles for teachers in conceptualizing and enacting reform. The minimal role of the teacher is also vexing when we consider the findings on factors affecting student achievement. Although 48 percent of the variance in student achievement is attributable to home and family factors that are largely out of the school system's control, 51 percent of the variance is attributable to controllable factors, 43 percent of which can be attributed to teacher quality (Ferguson, 1991). Despite these findings, we seem to have few ideas about how to enlist the support of teachers in reform efforts, how to enhance their capacity to maximally contribute to the reform effort, and how to engage teachers in reshaping reform efforts in response to their experiences in enacting reform.

Fullan (1992) reported that the time spent in deliberating on and enacting new educational policies has generally been three times greater than the average time allotted for planning the initial implementation. One hypothesis for this finding might be that we know very little about how to structure and support such a planning process.

Many policymakers have identified the critical role of the teacher in the reform process. “Teachers are, in one sense, the problem that policy seeks to correct” (Cohen & Ball, 1990, p. 238). Underinvestment in teacher knowledge has killed many a reform movement in the past, especially those that strove toward child-centered forms of education (Darling-Hammond, 1990). Cremin (1965) attributes the past failures of educational reform efforts to teacher capacity. The landmark research reported in the special issue of *Educational Evaluation and Policy Analysis* devoted to teachers’ responses to the California mathematics reforms was enormously helpful to our getting a finer sense of the role of the teacher in mediating the change process. The direct study of how innovations affected teaching practices across five elementary teachers’ classrooms revealed the varied responses that these teachers made as a function of their knowledge and beliefs. In addition, this research illustrated the ways in which teachers filled in the gaps in their understanding of the policy, creating a melange of practices.

Thus, the teacher must be front and center as we discuss how to improve comprehension instruction in schools today. The question becomes, *How can we bring about increased teacher quality and expertise in teaching reading comprehension?* Teachers who exhibit increased teacher quality and expertise have a deep knowledge about the reading process and reading comprehension. They also have the knowledge and skills to implement research-based instructional strategies in their teaching, ideally while also making their practice-based reflections on those instructional strategies available to researchers. In this report, we identify what we know about the answer to this question and raise new questions for additional research.

To answer this question, we look at two bodies of research: one on teacher education and another on professional development. Teacher education or teacher preparation programs refer to four- and five-year programs (both undergraduate and graduate) whose goal is to prepare individuals for teacher certification. Professional development refers to the ongoing education of certified teachers. We limit our discussion to teacher education and professional development that directly relate to learning how to teach reading comprehension, even though we draw from the larger educational research base in order to answer our question. And we acknowledge in advance that the research base on effective teacher education and professional development is disappointingly thin. Nonetheless, we argue that it is sufficient to support doing a better job than we are now doing, even as we pursue research designed to provide enhanced content about excellent comprehension instruction and about improved models for teacher education and professional development.

What We Already Know About Teacher Preparation

A common belief among many Americans is that teaching is something that people can do without much preparation (Darling-Hammond & Green, 1994). The need for teacher preparation programs has always been suspect. In fact, during the 1990s, many alternative teacher education programs were developed to certify teachers without requiring traditional teacher preparation. These programs were based in the belief that individuals with extensive life experiences and expertise in a particular domain—science, history, physics, math—could certainly teach in that domain with minimal preparation.

Although this trend has continued over the past 10 years, a plethora of literature related to teacher preparation programs has become available (for the most recent review of this work, see Sikula, 1996). Most of this literature, however, consists of descriptions and discussions of existing teacher preparation programs, case study analyses of pre-service teachers' beliefs and experiences, and recommendations for improving teacher preparation programs that are based on theory, logic, or experience. In addition, the literature is largely descriptive and qualitative (for a recent review of this work applied to reading education, see Fisher, Fox, & Paille, 1996). Although this body of work can be helpful in identifying issues and constructs for future study, it cannot, by itself, be used to make legitimate claims about teacher education programs. In fact, the NRP (2000) found no studies that measured student achievement as a result of teacher education. As Anders and her colleagues (Anders, Hoffman, & Duffy, 2000) stated, "Few . . . claims [about teacher education and reading] stand on a solid research base" (p. 727).

Nevertheless, we do know a few things about teacher preparation programs. For example, we know that pre-service teachers often enter teaching programs with firmly held beliefs about the nature of knowledge and the nature of teaching. These beliefs have been acquired through their own experiences as learners in schools. These beliefs shape how they view the teaching and learning processes and their own teaching and learning. We also know that many pre-service teachers enter teaching with the idea that there is "little need to obtain a knowledge base in pedagogy in order to become effective teachers" (Lanier & Little, 1986, p. 11). In the past, these candidate teachers have viewed education courses as weak and easy courses, the "Mickey Mouse" courses of the university. NRP (2000) found, however, that pre-service teacher education programs appear to improve candidate teachers' knowledge about teaching and learning; pre-service teachers, in other words, learn what they are taught. Thus it is reasonable to conclude that well-designed teacher education programs have a positive effect on reading outcomes.

What We Need to Know About Teacher Preparation

These claims leave much work to be done before we can better understand the effect of teacher preparation programs on developing expertise in teaching reading comprehension. Several key questions need to be addressed:

- What knowledge base (e.g., in language development, sociolinguistics, multiculturalism, reading development) do teachers need for effective reading comprehension instruction?
- What is the relative power of various instructional delivery systems (e.g., field-based experiences, video-based cases, demonstration teaching, microteaching) for helping teachers acquire the knowledge and skills they need to successfully teach comprehension to students of different ages and in different contexts?
- What do extant national data sets (e.g., NAEP) show about the extent to which teacher preparation experiences relate to teacher practices and student performances on comprehension measures?

What We Already Know About Teacher Professional Development

Conventional wisdom among teacher educators is that pre-service teachers are easier to work with than practicing teachers. Although pre-service teachers certainly hold prior beliefs about teaching and learning, these teacher educators think that the beliefs of practicing teachers are typically more entrenched. Many believe that practicing teachers, through their teaching experiences and classroom routines, have developed established ways of thinking about and implementing instruction—ways that are often resistant to change. For example, it is very difficult for practicing teachers to learn how to use instructional strategies that are different from the ones with which they are familiar. Joyce and Showers (1996) found that it takes as many as 30 instances of practicing a new routine before teachers can successfully incorporate it into their repertoire of practice.

Other research has corroborated this conventional wisdom. A body of research demonstrates the ineffectiveness of many traditional forms of in-service education for teachers (Cochran-Smith & Lytle, 1999). First, we know that the traditional staff development format is a relatively brief “one shot” workshop in which a presenter presents information to teachers about instructional practices. The effectiveness of these workshops, when evaluated at all, is typically measured through surveys of teacher satisfaction and only rarely by changes in teacher behavior. For the most part, teachers report that they perceive professional development in general to be of little use or value.

What We Need to Know About Teacher Professional Development

But what conditions promote effective professional development experiences? Effective professional development is associated with several characteristics (NRP, 2000). First, effective programs cover longer periods of time than do less-effective programs. Second, extensive investment of both money and time is needed on a continual basis for effective professional development. Third, effective professional development is content-focused and provides teachers with theoretical understandings of subject matter (Darling-Hammond, 2000; Elmore, 1999–2000; Joyce & Showers, 1996). Finally, a wide variety of content, when used for professional development, appears to be successful (NRP, 2000).

Since most of these claims about professional development in general relate to professional development to improve reading instruction as well, we can use the claims to identify what we do not know about effective professional development that supports high-level reading comprehension instruction. Among the things we need to know are the following:

- What content (declarative and procedural knowledge about readers, text, tasks, and contexts) and sequencing of content lead to effective professional development programs?
- How do various instructional delivery systems for professional development (e.g., in-class coaching, participatory learning, video-based cases, demonstration teaching, collaborative planning, lesson studies) influence the acquisition of knowledge and skills that lead teachers to enact effective instructional practices for students of different ages and in different contexts?
- What are the critical components of professional development that lead to effective instruction and sustained change in teachers' practice?
- How do teachers' existing beliefs and instructional practices influence how teachers use new information about teaching reading when that new information conflicts with what they already believe and do?
- What are various ways to support teachers so that they are willing to spend the time and cognitive effort and energy necessary to improve their comprehension instruction?

ASSESSMENT OF READING COMPREHENSION

Understanding the nature of the problem of reading comprehension requires having available good data identifying which readers can successfully undertake which activities with which texts. Such data are not available, in part because the widely used comprehension assessments are inadequate. Further, the

improvement of instruction relies crucially on the availability of information about the effectiveness of instruction. Teachers need reliable and valid assessments tied closely to their curricula so that they can see which students are learning as expected and which need extra help. In addition, schools, districts, and states are increasingly calling for reliable and valid assessments that reflect progress toward general benchmarks of reading, writing, and mathematics ability. For the area of reading comprehension, good assessments that are tied to curriculum as well as good assessments of general comprehension capacity are sorely needed. These assessments need to be constructed in accordance with the many advances in psychometric theory.

What We Already Know About Comprehension Assessments

Currently available assessments in the field of reading comprehension generate persistent complaints that these instruments

- inadequately represent the complexity of the target domain
- conflate comprehension with vocabulary, domain-specific knowledge, word reading ability, and other reader capacities involved in comprehension
- do not rest on an understanding of reading comprehension as a developmental process or as a product of instruction
- do not examine the assumptions underlying the relationship of successful performance to the dominant group's interests and values
- are not useful for teachers
- tend to narrow the curriculum
- are unidimensional and method-dependent, often failing to address even minimal criteria for reliability and validity.

Indeed, most currently used comprehension assessments reflect the purpose for which they were originally developed—to sort children on a single dimension by using a single method. Even more important, though, is that none of the currently available comprehension assessments is based in a viable or articulated theory of comprehension. And none can give us a detailed or convincing picture of how serious the problem of comprehension achievement in the United States is. These considerations, as well as the thinking about the nature of reading comprehension represented in this document, create a demand for new kinds of assessment strategies and instruments that (1) more robustly reflect the dynamic, developmental nature of comprehension; (2) represent

adequately the interactions among the dimensions of reader, activity, text, and context; and (3) satisfy criteria set forth in psychometric theory.

Currently, widely used comprehension assessments are heavily focused on only a few tasks: reading for immediate recall, reading for the gist of the meaning, and reading to infer or disambiguate word meaning. Assessment procedures to evaluate learners' capacities to modify old or build new knowledge structures, to use information acquired while reading to solve a problem, to evaluate texts on particular criteria, or to become absorbed in reading and develop affective or aesthetic responses to text have occasionally been developed for particular research programs but have not influenced standard assessment practices. Because knowledge, application, and engagement are the crucial consequences of reading with comprehension, assessments that reflect all three are needed. Further, the absence of attention to these consequences in widely used reading assessments diminishes the emphasis on them in instructional practices as well.

What We Need in the Area of Comprehension Assessments

The entire research enterprise sketched out in this report depends on having a more adequate system of instrumentation for assessing reading comprehension. A satisfactory assessment system is a prerequisite to making progress with all aspects of the research agenda we propose. Thus we argue that investing in improved assessments has very high priority. It is clear that we cannot even sketch the seriousness of the problem of reading comprehension in the United States or the nature of the decline in comprehension outcomes that is the source of much worry until we have an assessment system that can be used across the developmental range of interest and that assesses the same construct across that range.

Assessing the effect of changes in instruction depends on having valid, reliable, and sensitive assessments. The effect of assessment on instruction is a question that constitutes a research agenda of its own, particularly in this highly accountability-oriented era of education reform. But the power of high-stakes assessments over instruction and curriculum can be somewhat mitigated if teachers have available alternative assessment options that give them more useful information.

Any system of reading assessments should reflect the full array of important reading comprehension consequences. We argue that a research program to establish expectable levels of performance for children of different ages and grades on this full array of consequences is necessary. Such a program is a prerequisite to developing performance criteria at different age and grade levels and to pursuing questions about reader differences associated with instruc-

tional histories, social class, language, and culture in reading comprehension outcomes.

Although the reading comprehension consequences defined above constitute the basis for designing a comprehension assessment that would reflect success, our view suggests that assessments designed to reflect readers' cognitive, motivational, and linguistic resources as they approach a reading activity are also necessary. For instance, when the outcomes assessment identifies children who are performing below par, process assessments could help indicate why their reading comprehension is poor. Further, diagnostic assessments are crucial in dissecting the effect of particular instructional or intervention practices. Ideally, we would move ultimately toward assessment systems that can also reflect the dynamic nature of comprehension, for example, by assessing increments of knowledge about vocabulary and particular target domains that result from interaction with particular texts.

We see the development of an assessment system for reading comprehension as having a very high priority. Such a system should be based in contemporary approaches to test development and evaluation. We recognize that developing a comprehensive, reliable, and valid assessment system is a long-term project. Crucial for such a system are the criteria for judging performance across the developmental span. Nonetheless, a substantial start could be made in the short run, either by targeting the assessment of outcomes and reader resources as a major task of the research agenda or by encouraging the development of prototype assessments for outcomes and reader resources within other research efforts (such as research focused on instructional efficacy). Such an effort is central to pursuing larger research agendas, such as longitudinal work to create a picture of the development of reading comprehension, a large-scale effort to determine how U.S. children are functioning as readers, or a systematic pursuit of differences in reading comprehension performance related to cultural background, social class, and language status.

The approach to assessment proposed here differs from current approaches to reading assessment in that it would both grow out of and contribute to the development of an appropriately rich and elaborated theory of reading comprehension. Assessment procedures generated by this approach are thus more likely to be influenced and changed by theoretically grounded reading research. Our approach also highly values the utility of assessment for instruction. Of course, comprehensive assessment systems can place high demands of time on students and teachers; thus, we have an obligation to develop assessments that are embedded in and supportive of instruction, rather than limited to serving the needs of researchers.

A comprehensive assessment program reflecting the thinking about reading comprehension presented here would have to satisfy many requirements that have not been addressed by any assessment instruments, while also satisfying the standard psychometric criteria. The minimum requirements for such a system follow:

- *Capacity to reflect authentic outcomes.* Although any particular assessment may not reflect the full array of consequences, the inclusion of a wider array than that currently being tested is crucial. For example, students' beliefs about reading and about themselves as readers may support or obstruct their optimal development as comprehenders; teachers may benefit enormously from having ways to elicit and assess such beliefs.
- *Congruence between assessments and the processes involved in comprehension.* Assessments that target particular operations involved in comprehension must be available, in the interest of revealing inter- and intra-individual differences that might inform our understanding of the comprehension process and of outcome differences. The dimensionality of the instruments in relation to theory should be clearly apparent.
- *Developmental sensitivity.* Any assessment system needs to be sensitive across the full developmental range of interest and to reflect developmentally central phenomena related to comprehension. Assessments of young children's reading tend to focus on word reading rather than on comprehension. Assessments of listening comprehension and of oral language production, both of which are highly related to reading comprehension, are rare and tend not to be included in reading assessment systems despite their clear relevance. The available listening comprehension assessments for young children do not reflect children's rich oral language-processing capacities, because they reflect neither the full complexity of their sentence processing nor the domain of discourse skills.
- *Capacity to identify individual children as poor comprehenders.* An effective assessment system should be able to identify individual children as poor comprehenders, not only in terms of prerequisite skills such as fluency in word identification and decoding, but also in terms of cognitive deficits and gaps in relevant knowledge (background, domain specific, etc.) that might adversely affect reading and comprehension, even in children who have adequate word-level skills. It is also critically important that such a system be able to identify early any child who is apt to encounter difficulties in reading comprehension because of limited resources to carry out one or another operation involved in comprehension.
- *Capacity to identify subtypes of poor comprehenders.* Reading comprehension is complexly determined. It therefore follows that comprehension dif-

difficulties could come about because of deficiencies in one or another of the components of comprehension specified in the model. Thus, an effective assessment system should be able to identify subtypes of poor comprehenders in terms of the components and desired outcomes of comprehension. It should also be capable of identifying both intra- and inter-individual differences in acquiring the knowledge and skills necessary for becoming a good comprehender.

- *Instructional sensitivity.* Two major purposes for assessments are to inform instruction and to reflect the effect of instruction or intervention. Thus, an effective assessment system should provide not only important information about a child's relative standing in appropriate normative populations (school, state, and national norms groups), but also important information about the child's relative strengths and weaknesses for purposes of educational planning.
- *Openness to intra-individual differences.* Understanding the performance of an individual often requires attending to differences in performance across activities with varying purposes and with a variety of texts and text types.
- *Usefulness for instructional decisionmaking.* Assessments can inform instructional practice if they are designed to identify domains that instruction might target, rather than to provide summary scores useful only for comparison with other learners' scores. Another aspect of utility for instructional decisionmaking is the transparency of the information provided by the test given to teachers without technical training.
- *Adaptability with respect to individual, social, linguistic, and cultural variation.* Good tests of reading comprehension, of listening comprehension, and of oral language production target authentic outcomes and reflect key component processes. If performance on a task reflects differences owing to individual, social, linguistic, or cultural variations that are not directly related to reading comprehension performance, the tests are inadequate for the purposes of the research agenda we propose here.
- *A basis in measurement theory and psychometrics.* This basis should address reliability within scales and over time, as well as multiple components of validity at the item level, concurrently with other measures and predictively relative to the longer-term development of reading proficiency. Studies of the dimensionality of the instruments in relationship to the theory underpinning their construction are particularly important. Test construction and evaluation of instruments are important areas of investigation and are highly relevant to our proposed research agenda.

Clearly, no single assessment would meet all these criteria. Instead, we propose an integrated system of assessments, some of which may be particularly appropriate for particular groups (e.g., emergent or beginning readers, older struggling readers, second-language readers, or readers with a particular interest in dinosaurs). Further, the various assessments included in the system would address different purposes, such as a portmanteau assessment for accountability or screening purposes, diagnostic assessments for guiding intervention, curriculum-linked assessments for guiding instruction, and so on. Given that we are proposing multiple assessments, we believe that studies of their dimensionality and of the interrelations of these dimensions across measures are especially critical.

A sample of issues that would certainly arise in the process of developing a comprehensive assessment system for reading comprehension follows:

- The effect of various response formats on performance.
- Variation in performance across types of text.
- The effect of nonprint information.
- The effect of various formats and accommodations on the test performance of learners of English as a second language.
- Variation in performance across a variety of types of discourse and genres, including hypertext.
- The effect on performance of specifying different purposes for reading.
- The capacity to differentiate domain-specific and reading-general operations.
- The need to reflect performance on literacy tasks typical of electronic reading, such as retrieval.
- The capacity to explore issues that go outside the traditional rubric of comprehension, such as scanning, intertextuality, domain-specific strategies, and consulting illustrations.
- The reliability, validity, and dimensionality of different assessment instruments and approaches.

Key Issues the Research Agenda Should Address

The key questions and issues that a research agenda on reading assessment needs to address and that are closely connected to the RRSB's proposed areas for future instruction research, include the following:

- How can the education community measure strategic, self-regulated reading, including a student's use of such strategies as questioning, comprehension monitoring, and organizing the knowledge gained from text?
- To what extent are performance-based assessments of reading sensitive to a student's competencies in such processes as vocabulary, cognitive strategies, writing ability, oral language (syntax), reading fluency, domain content knowledge of the texts, and such dispositions as motivation and self-efficacy for reading?
- How do we design valid and reliable measures of self-regulated, strategic reading that teachers can administer in the classroom to inform their instructional decisions?
- What informal assessments should teachers use to identify children who may need additional or modified instruction within the classroom to prevent a referral to special education services?
- How do we construct informal assessments to assist teachers in identifying how to help students who have low reading comprehension? For example, how could teachers identify which children need to be taught specific reading strategies or supported in domain knowledge acquisition or motivational development?
- What reading comprehension assessment could be both administered efficiently by all teachers in a school and used across grades to document student growth and guide teacher decisions about the appropriate texts, tasks, contexts, and learning activities for students?
- What available measures of motivation and engagement in reading can be linked to reading competencies, related to growth over time, and used to guide classroom learning activities?
- What measures of reading fluency can be used at the levels of the individual student, the classroom, and the school and can be related to reading comprehension and reading motivation?
- Which measures of reading comprehension are sensitive to specific forms of reading instruction and intervention for all readers?
- What are the dimensions evaluated by different assessments in relation to more traditional assessments and the proposed new approaches to assessment? How well does the dimensionality map onto the theories behind the development of the assessments?