6

Taking Stock of the Reading for Understanding Initiative

P. David Pearson, *University of California, Berkeley*
Annemarie Sullivan Palincsar, *University of Michigan*
Peter Afflerbach, *University of Maryland*
Gina N. Cervetti, *University of Michigan*
Panayiota Kendeou, *University of Minnesota*
Gina Biancarosa, *University of Oregon*
Jennifer Higgs, *University of California, Davis*
Miranda Fitzgerald, *University of North Carolina at Charlotte*
Amy I. Berman, *National Academy of Education*

CONTENTS

INTRODUCTION

In this chapter, we address the central issue of this volume, "What rewards did we reap from the substantial investment made by the Institute of Education Sciences in this focused effort to better understand and improve reading comprehension on the pathway to improved student achievement?" We try to answer this grand question by deconstructing it into more specific questions that, taken together, move us toward a grand answer:

1. Looking across the work of the six teams within the three domains detailed in the earlier chapters of this volume (the strands of nature and development, assessment, and curriculum and instruction), what findings and common features of the work stand out as markers of progress?
2. Recognizing that not all of the salutary features of the Reading for Understanding (RfU) initiative fit neatly into our three-strand structure, what other noteworthy outcomes of the initiative contribute to the narrative about what was learned?
3. Using as benchmarks the contextual issues and movements that were influential when this work began almost 10 years ago, how do we think differently about reading comprehension as a result of the RfU?
4. What legacy, in terms of an agenda for future reading comprehension research, has the RfU left for the field?

Our approach to answering the grand question and its four facets is straightforward. First, we look across Chapters 2–5 to offer a concise summary of what we have already covered in greater detail. Second, we add a section on additional lessons learned, mainly about the affordances of the process employed in carrying out this work. Third, we look back at the influential contextual factors (policies, theories about reading development, or pedagogical movements) we introduced in Chapter 1 and use them as benchmarks for assessing how we think differently about reading comprehension a decade after the RfU began. Finally, we look beyond the RfU to assess its legacy in terms of future work by responding to "What would the Request for Application (RfA) for the next RfU look like?"

QUESTION 1: SUMMARIZING THE CONTRIBUTIONS OF THE RFU INITIATIVE

Nature and Development of Comprehension

With respect to the *nature and development of comprehension*, the RfU studies revealed the number and range of skills, knowledge, and dispositional characteristics that support reading comprehension, as well as the relative importance of these skills and knowledge as students matriculate through the grades. These studies informed the field's understanding of the *linguistic* and *cognitive* skills associated with successful listening and reading comprehension and called our attention to the role of both *word* and *world* knowledge in comprehension activity.

At the preschool and elementary levels, linguistic, cognitive, and behavioral skills were particularly significant predictors of comprehension, while at the adolescent level,

other reader- specific factors (e.g., background knowledge, vocabulary knowledge, discourse expertise, strategy use, and inferencing) emerged as explanatory factors regarding performance and development. The factors included in the adolescent portfolio are largely discipline-specific language and reasoning skills: academic language (both vocabulary and discourse) that characterizes advanced discussion about text and language, as well as epistemological (how we know what we know and do not know) and perspective-taking dispositions (how we learn to decenter in examining text-based narratives, arguments, and explanations).

Specific to language skills, the evidence from the developmental studies suggested that language might most productively be regarded as a single entity, or perhaps as a cluster or assemblage of closely related skills, thus calling into question both assessments and interventions that privilege discrete language skills. Specific to cognitive skills, the RfU development studies revealed that some skills, namely, attentional control and self-regulation, made small but significant contributions to comprehension, while comprehension monitoring and inferencing made more substantial contributions to comprehension, especially in distinguishing between stronger and weaker comprehension. The RfU studies, while confirming the role of word and world knowledge in listening and reading comprehension, extended our understanding of this relationship by illustrating the role that word and world knowledge play in supporting readers' ability to engage in inferencing and comprehension monitoring.

Finally, research on the nature and development of comprehension raised new questions for future research regarding the significance and malleability of different knowledge sources and skills at different points in development and in relation to different text genres and text characteristics. These questions constitute one of the research legacies of the RfU initiative.

Assessment

The RfU initiative had a profound influence on the development and validation of *reading comprehension assessments*, giving rise to a new generation of relevant measures of the construct.

First, we learned that authenticity, complexity, and psychometric adequacy can all be achieved, even in a single assessment. The Global Integrated Scenario-Based Assessment (GISA) system demonstrated the feasibility of assessing the broad, multidimensional comprehension construct in an authentic way while also achieving technical adequacy.

Second, assessment does not have to be distinct from learning; in fact, the approach to scenario-based assessment utilized in GISA entails learning as a fundamental premise of the assessment process.

Third, to reflect and capture the dynamic nature of reading comprehension, we need multiple assessment systems that vary in construct coverage and tasks. Within the RfU portfolio, both the Reading Inventory and Scholastic Evaluation (RISE) and Florida Center for Reading Research (FCRR) Reading Assessment (FRA) assessment systems ensure broad coverage of the components that are either a part of reading comprehension or on the developmental pathway to it. GISA, by contrast, represents the orchestration of reading comprehension and other variables in using the fruits of reading comprehension to perform related comprehension tasks.

Fourth, knowledge is an integral component of reading comprehension, and as such it should be integrated, not simply treated as a nuisance variable and controlled. GISA has integrated knowledge in the design of the assessment as an integral component and provided evidence for the feasibility and efficacy of this approach.

Fifth, looking across the entire range of the RfU assessments, including those developed for particular studies by the RfU teams, the consortium addressed not only prior knowledge, but also metacognitive and self-regulatory strategies, reading strategies, and motivation and engagement. The integration of these variables represented a significant advance in comprehension assessment design and provided a set of tools sensitive enough to inform and evaluate the effects of high-quality instruction. A number of these assessments were theoretically robust and reflected a reconceptualization of comprehension consistent with advances in theory and research as well as numerous national and international standards and movements. The stage is set for future work in which these tools can replace traditional standardized reading comprehension assessments. Furthermore, these assessments can be investigated for their potential to support new approaches to curriculum and pedagogy, particularly those that privilege differentiated instruction and the application of text-supported knowledge to new problems and situations.

Sixth, the connection between assessment and instruction is underscored by an inherent reciprocity that has been realized by the full scope of assessment efforts of the RfU. This reciprocity is identified by two important characteristics of the assessment itself, instructional value and instructional sensitivity. Instructional value refers to an assessment's capacity to provide information about a student's strengths and weaknesses in particular skills or processes that might become candidates for instruction. Instructional sensitivity refers to an assessment's capacity to reflect the effects of instruction or intervention; this is a key attribute in pedagogical research designed to promote particular comprehension processes. Because the RfU assessments acknowledged and reflected a broader and more authentic conceptualization of reading comprehension that emphasized instructional value and sensitivity, significant progress has been made in this area. Specifically, interventions that moved the needle on GISA, the broader index of reading comprehension, were primarily multicomponent, suggesting good coverage between the aspects of the construct being assessed and those being trained. Even though these effects were generally small, they underscore the importance of reflecting the multicomponent nature of reading comprehension, including the capacity to orchestrate those components to produce new knowledge or learning, in a summative assessment. Similarly, reflecting the multicomponent nature of reading comprehension also increased instructional value. Both the FRA and RISE assessment systems evaluate multiple reading components and thus provide information regarding students' strengths and weaknesses that can inform instructional decisions.

Compared to other recent initiatives, such as the Partnership for Assessment of Readiness for College and Careers (PARRC) and Smarter Balanced Assessment Consortium efforts promoted by various state consortia to be responsive to curricula consistent with the Common Core State Standards (CCSS), GISA bears a stronger resemblance to these other assessments than does either RISE or FRA. Both the GISA and CCSS-aligned assessment privilege some approximation to authentic literacy activities. However, neither the PARRC nor the Smarter Balanced assessments come close to GISA in privileging

purpose (knowing from the beginning what the culminating product will be) or context (completing the assessment in a virtual community setting).

In conclusion, the assessment systems developed by the RfU are broader and more authentic than those typically available in the educational marketplace. They are also developmentally sensitive, and emphasize instructional sensitivity and value. These assessments have a strong theoretical basis and defensible psychometric properties. The calibration and validation studies were extensive and iterative, and were undertaken across the United States. The result is a set of forward-thinking assessments that not only meet the standards of educational and psychological testing, but also promise to advance both research and practice in reading comprehension for years to come.

Curriculum and Instruction

With respect to *curriculum and instruction*, the RfU initiative produced a range of positive, but often inconsistent, results on a wide range of measures across the K–12 continuum. Effects were greater and more consistent for assessments that represented curriculum-aligned and researcher-developed measures than for those that were curriculum-independent and published measures of key outcomes. The strongest effects were observed for measures of vocabulary, morphology, comprehension monitoring, and knowledge acquisition. Those interventions that most consistently "moved the needle" on reading comprehension and a host of related measures (such as vocabulary, knowledge acquisition, application, and enabling skills) were characterized by well-orchestrated, multicomponent instruction.

Consistent with those studies that focused on development, instructional studies revealed that the relationships among enabling skills, knowledge, and reading comprehension are dynamic (changing across grades) and synergistic (improvement in one process can enhance performance in another). Across the RfU teams, researchers catalogued the roles of different types of knowledge in reading comprehension: declarative, procedural, conditional, disciplinary, and epistemic. This portfolio represents an *advancement in our conceptualization of the role of knowledge in comprehension*. In previous eras, we would have suggested that declarative knowledge of the topics of the texts was the most important resource for comprehension, followed closely by procedural knowledge of how to engage in strategies or practices to demonstrate comprehension, with a nod to conditional knowledge (when, why, and where to use it). Twenty years ago, that trio would have told the story about the role of knowledge in comprehension. The insights emerging from the RfU, hard on the heels of developments in adolescent literacy in the 2010s, demonstrates that the knowledge repertoire needs to be extended to include how disciplinary knowledge (how we talk, write, think, explain, and argue about key ideas in the major academic fields of study) and epistemic knowledge (how knowledge is generated and evaluated within the disciplines) change the relationship between knowledge and comprehension in substantive ways. The RfU work suggests that these five types of knowledge are variously cause, consequence, and covariate of reading comprehension.

Learning to read and *reading to learn* surfaced in the RfU portfolio as complementary goals rather than separate stages of development. These two complex processes were revealed to be interwoven fruitfully across students' school careers, and the RfU research

> The strategies enabled the students to see different nuances, discuss, and pull information together.
>
> —*RfU Participating Teacher*

described their ongoing development and interaction with one another. Granted, we learned, particularly from the adolescent teams (Catalyzing Comprehension through Discussion and Debate [CCDD], Promoting Adolescents' Comprehension of Text [PACT], and Reading, Evidence, and Argumentation in Disciplinary Instruction [READI]), that both learning to read and reading to learn are different enterprises in history, literature, and science; however, within each discipline, we also learned that reading to learn and learning to read were complementary. Conversely, we learned that even in the primary grades, as early as pre-kindergarten (pre-K), there is also a complementary relationship between reading to learn and learning to read (see Let's Know! [LK] and Content Area Literacy Instruction [CALI] for examples in Chapter 4).

The instructional research provided further evidence that student *engagement* with texts and tasks supports comprehension. The RfU researchers provided details on the instruction and classroom environments that contribute to students' engagement, and on the particular aspects of reading comprehension that benefit from engagement.

QUESTION 2: ADDITIONAL LESSONS LEARNED FROM THE RFU WORK

As suggested earlier, not all of the lessons learned emerged from the substance of the three strands of development, assessment, and pedagogy. Some were implicit in the processes used to carry out the work and in the constellation of findings across teams. Three such lessons stand out: affordances of the consortium model, methodology to account for professional learning issues, and insights about why it is hard to "move the needle" on (particularly distal) measures of reading comprehension.

The Affordances of the RfU Consortium Model

The research model enacted in the RfU consortium provides a demonstration of what is possible in the design, implementation, and analysis of lines of inquiry with the affordances of adequate funding, extended time frames, and a diverse array of expertise to carry out the work. When there is a sufficiently long runway, scholars have the opportunity to exploit the complementarity of research methods, scholarly traditions, and academic disciplines. As a result, the RfU scholars were able to engage in:

* Fine-grained, longitudinal, largely theory-driven examinations of the infrastructure of processes and knowledge developed by readers over time and in response to instruction;
* Examinations of exemplary and/or typical practice in school settings to establish the contexts in which pedagogical work might be situated;
* Design-based research and development in which teachers and schools are partners, not subjects, and, through interaction and consensus building with researchers, contribute to the creation and implementation of curriculum, assessments, and teaching practices;

* Extensive professional development, often in sustained, long-term professional learning communities;
* Well-designed, pedagogically complex, adequately powered efficacy studies and randomized controlled trials (RCTs) in which both learning and teaching are measured;
* Efforts that built upon each team's prior, less well-funded, efforts to design curricular and instructional interventions in the same pedagogical family;
* Secondary analyses of their own data, with an eye toward determining whether particular approaches were effective with particular subgroups (e.g., low-performing readers, emergent bilingual learners (EBs), or students from low-income schools); and
* Regular interaction with the researchers from the other teams in meetings brokered by the Institute of Education Sciences (IES) staff overseeing the RfU.

Even though the six projects were very different from one another, there were productive opportunities to interact with one another about the entire spectrum of research concerns—including models of reading, useful measures, sampling designs, new statistical tools, professional development formats, and tools for promoting or monitoring treatment fidelity.

At its best, the RfU research reflected a broad view of the science of reading, drawing from diverse traditions, paradigms, and theories. The research was also informed by the relevant scientific perspectives of researchers from affiliated fields, including psychology, sociology, learning sciences, linguistics, and literary criticism. This broad perspective was evident in how the RfU research was designed, the nature of data collection and analysis, and the breadth of examined outcomes. In terms of the "science" of reading, the RfU teams were informed by scientific literature that ran the gamut from phonemic awareness to phonics to morphological awareness to vocabulary to metacognition to motivation, and to comprehension itself.

Related science also made important contributions to the conceptualization and execution of the RfU research. For instance, FCRR employed a common set of measures across a wide range of language-related interventions to allow for the possibility of highly precise comparisons across the entire family of interventions. The Language and Reading Research Consortium (LARRC) used a design study to build understanding of local school needs and dovetail them with research intent, to enlist and gain support of key participants in the research work, and to determine the "place" of research in existing school contexts. READI conducted theoretical analyses of history, science, and literary criticism to determine common and unique features of the culture of each discipline, which then informed comprehension curriculum and instruction, and the development of evidence-based argumentation and of extensive scaffolds—heuristics for sense making, evaluating the validity of information, developing arguments, and reflecting on personal progress along the way. PACT and CCDD designed comprehension instruction that attempted to boost students' motivation and engagement, while LARRC sought to enhance young children's metacognition related to language comprehension.

In short, the affordances of the RfU grant mechanism enabled multiple, complementary strands of research (on the development, assessment, and instruction of reading

comprehension) to be conducted simultaneously. The coordination of multiple research goals, along with time to allow for iterations of design and refinement and of longitudinal data, helped to increase the "yield" of the RfU. Moreover, the RfU researchers will continue to publish results from this investment for years to come because of the richness of the data gathered. To continue to foster this kind of breadth, depth, and coordination of research inquiries, funders such as IES should reconsider whether the separation of project types (i.e., project types 1 through 4 are labeled exploration, development and innovation, initial efficacy and follow-up, and measurement) and investment in narrowly focused inquiries is ever likely to provide the kind of nuanced results that the RfU has produced. In fact, the very structure of the RfU acted as something of a guarantee that, even if an intervention or hypothesis was not fruitful, researchers were still left with data that improved our understanding of reading comprehension.

Teacher Professional Learning

One objective of the RfU initiative was to encourage changes in reading comprehension instruction as a means of improving student reading comprehension performance. The RfU studies explicitly set out to disrupt "business as usual" by enacting a reconceptualization of comprehension teaching routines, reading comprehension curricula, expectations of student performance, and learning outcomes. Underlying this goal for student learning was an expectation that teacher expertise, as indexed by teacher knowledge of how to enact more challenging classroom practices, would be the focus of professional development activities, in a form that was developmentally appropriate for a given grade-level band. In such instances, teaching basic reading strategies (e.g., using context to help determine word meaning, or identifying an explicit main idea statement) is important and necessary, but not sufficient, for developing students' capacity to engage with texts and tasks in more complex and challenging ways, enacting what we have recently come to call deeper learning (NRC, 2012). If the practices that students engage in are to be disrupted, so too must the practices that teachers are asked to enact in prescribed commercially available or mandated curricula. The centrality of teacher professional development in support of helping students move to enhanced levels of comprehension and performance is evident across many of the RfU teams. The nature of this professional development—ongoing, detailed, collaborative, and tailored to individual teachers' needs—is critical.

> One of the largest shifts in my practice was due to unpacking my own moves, first with "Flowers" [used in the professional development session] and then in my classroom texts. Shifting the purpose of classroom discussion—from asking questions about the text that students would answer or asking students to only make personal connections—to apprenticing students into the kind of thinking that I do as an expert reader in the discipline. [It] shifted the way I planned for discussions.
>
> *—RfU Participating Teacher*

Many of the resulting RfU approaches to reading comprehension instruction necessitated new approaches to teaching that placed new demands on teachers. The

RfU comprehension curricula required the creation of means to educate and evaluate teachers, which was accomplished through professional development (for the education) and measuring fidelity of implementation (for the evaluation). A common medium for professional development in the RfU consortium was the design study, which, according to the NRC (2002), focuses on "the evolution of learning. The learning might be that of a student, a teacher, or an organization" (p. 28). This approach to research and development demanded collaboration among varied stakeholders, particularly researchers and teachers, and the iterative construction of their goals and needs along the pathway to new and challenging curricula. The approach also encouraged broad participation in framing research questions, designing interventions, and professional development. Here, teachers' professional development was tied to learning the particulars of comprehension instruction within a specific research project, and was a necessary means of promoting growth in teachers' declarative, procedural, conditional, disciplinary and epistemic knowledge for a particular curriculum, such as those reviewed in Chapter 4 (i.e., LK, Comprehension Tools for Teachers [CTT], Word Generation [WG], Strategic Adolescent Reading Intervention [STARI], PACT, Comprehension Circuit Training, or READI).

In terms of innovative RfU comprehension curricula, professional development was necessary to ensure that teachers were able to teach new curricula that variously promoted young children's metacognitive strategies (LARRC, Johansson, & Arthur, 2016), middle schoolers' debate, discussion and academic language (Kim et al., 2016; LaRusso, Kim, et al., 2016), and high schoolers' epistemic development (Lee et al., 2016; Shanahan et al., 2016). Accordingly, the RfU research featured teachers' professional development in the ways and means of innovative reading comprehension instruction. The nature of this professional development—ongoing (LARRC, Farquharson, & Murphy, 2016), detailed (Wanzek & Vaughn, 2016), and sometimes on call and responsive to particular needs in particular situations (Connor et al., 2018)—was critical. Teaching strategies to students at an unprecedented young age (e.g., comprehension monitoring to pre-K students; LARRC, Johansson, & Arthur, 2016), and teaching to the twin targets of learning to read while reading to learn (LaRusso, Kim, et al., 2016) demanded that teachers learn new instructional approaches. As well, providing instruction that fostered students' ability to identify and then engage in evidence-based argument (Goldman et al., 2019) was preceded by teachers' learning the requisite strategies and mindsets that they would eventually scaffold for their students. Professional development was accomplished through attention to both theory and practice. For example, at a theory level, READI devised professional development that helped high school teachers understand the nature and workings of disciplinary and epistemic knowledge, which in turn informed their disciplinary teaching practices (Goldman et al., 2019). While grounded in theory, all of the RfU studies eventually focused on the practical aspects of enacting these curricula. Teachers' capacity to learn and adopt new practices related to new

> Usually I am very lecture based rather than text based. It was neat to see students pull things from text that I hadn't necessarily considered or chosen. They turned into little experts because of the text.
>
> *—RfU Participating Teacher*

curricula, strategies, and mindsets was expected to take time and involve challenges. As students in the RfU studies learned new strategies and stances related to reading comprehension, so teachers learned new practices for teaching and supporting this comprehension. This model of teacher professional development in which university and K–12 educators collaborate to translate theory into practice and improve that practice in the context of classroom instruction is reminiscent of the National Writing Project model (Wood & Lieberman, 2000).

As described in the RfU studies, there are necessary areas of growth in teachers' declarative, procedural, conditional, disciplinary, and epistemic knowledge. For instance, fostering students' ability to read like a historian differs from helping students read and learn the facts of history, or from reading like a scientist or literary critic. Modeling strategies for students so that they might question authors' use of claim and evidence, or reconcile competing accounts of scientific cause-and-effect phenomena, differs from teaching summarization strategies. Teachers must also help students learn to be comfortable with the uncertainty of the knowledge they are acquiring. Students may find that they cannot determine the truthfulness of two opposing eyewitness accounts, but they might be able to determine which account is more credible. This stands in contrast to many students' school experience of reading to determine facts and the "right" answer.

Fidelity of implementation is a perennial challenge (Foorman, Dombek, & Smith, 2016), and involves both "logistics and establishing an infrastructure for ensuring adequate implementation" (Gersten, 2016, p. 113). Fidelity of implementation was a goal across the RfU teams, and the sum of the RfU studies provides a tutorial in conceptualizing, working toward, and analyzing fidelity of implementation. The research designs used in the RfU studies required considerable a priori efforts, including resources, to establish guidelines for building and maintaining fidelity. Looking across the continuum of RfU fidelity of implementation efforts one encounters a range of efforts to account for fidelity. Some studies focused on measuring teacher adherence to researcher-created instructional scripts and routines (Connor et al., 2018), while other studies saw professional development as a means of establishing fidelity (Goldman et al., 2019). A related continuum illustrates that professional development might focus on the practical (i.e., did teachers follow the script of a particular lesson?) and the theoretical (i.e., was teachers' understanding of epistemic knowledge evident in a particular lesson?). In the final analysis, few studies used metrics of professional development (e.g., high versus low knowledge or implementation) to predict student performance.

As many RfU fidelity efforts were linked to professional development, they simultaneously measured fidelity and enhanced the probability that teachers in treatment classrooms were adept at teaching with new curricula. FCRR provided its already experienced instructional assistants (mainly former teachers and/or graduate assistants) with professional development that ranged from 6 to 12 hours of initial training, followed by 3 to 6 hours of "booster" professional development (Connor et al., 2018). Teachers also utilized electronic bulletin boards to post teaching-related questions and to submit responses to weekly implementation quizzes. Accompanying these efforts was fidelity monitoring, by which teachers were observed in person and provided immediate feedback. Formative ratings provided a means for amending instruction, while summative ratings told the overall story of fidelity of implementation.

LARRC, Johansson, and Arthur (2016) investigated the effectiveness of teacher professional development in relation to fidelity using classroom teachers' online surveys following completion of the professional development module, three observations of lessons, lesson logs for every lesson, and an end-of-unit teacher survey and guided interview. CCDD (LaRusso, Donovan, & Snow, 2016) used summary coaching reports, implementation challenge checklists, and semistructured interviews, including open-ended questions about implementation, along with teacher case summaries that described implementation progress and barriers and student completion of instructional materials. These measures both reported current implementation fidelity and shaped future fidelity efforts. Furthermore, CCDD scholars (LaRusso, Donovan, & Snow, 2016), in both the STARI and WG interventions, documented structural challenges to fidelity, which included lack of sufficient teaching time and difficulties with new programs.

To summarize, the learning involved in professional development that supports teaching comprehension parallels students' ongoing learning related to reading comprehension. In essence, teachers and students involved in many of the RfU projects were working on parallel learning tracks. While students were learning new strategies, routines, and stances to become better learners, teachers were acquiring knowledge and routines to become better at scaffolding that student learning. This is not to say that the pathway to deeper learning on the part of teachers or students was always clear, well marked, or free of obstacles. For example, as we reported in Chapter 5, teachers in the PACT intervention were able to learn and implement some teaching routines better than others. In particular, routines involving scaffolding more basic skills and knowledge (engaging prior knowledge or teaching key unit vocabulary) exhibited greater uptake from teachers than routines intended to promote higher-level practices such as close reading, interpretation, and critical reading (Wanzek & Vaughn, 2016). When queried about barriers to the uptake of new routines, teachers using the CCDD's WG and STARI (LaRusso, Donovan, & Snow, 2016) interventions cited interference with other, often more high-stakes, initiatives, such as state test preparation or covering the adopted curriculum or the normal variations of student behavior in classroom cultures. Even so, working closely with teachers in supportive teacher learning community settings, researchers were able to promote changes in teacher practices in the direction of key intervention principles (see, in particular, Goldman, Britt, et al., 2016; Goldman, Snow, & Vaughn, 2016) among intervention teachers in comparison to business-as-usual teachers. Moreover, in some instances, changes in teacher practice were associated with gains in student achievement (Lawrence, Rolland, Branum-Martin, & Snow, 2014, Wancek & Vaughn, 2016). Hard work, supportive settings for teachers to try new perspectives, and staying the course on everyone's part seem to be consistent threads in the more successful ventures.

The approaches to professional development and measuring teaching by the RfU teams built upon rich traditions established in decades of work in enacting (e.g., Coburn, 2003) and measuring (Rowan & Correnti, 2009) changes in teachers' knowledge, beliefs, and practices in reform-motivated curriculum projects. Thus, these approaches were not entirely new or ground breaking, but the professional development they embodied was exceptional in its durability (many of the teacher learning communities lasted over several years), focus (enacting particular curricula and measuring uptake of their key

components and principles), and engagement (finding a setting in which teachers could develop ownership of the reform curriculum).

Insights About "Moving the Needle"

Several scholars within the RfU consortium (e.g., Catts, 2018; Lonigan, Burgess, & Schatschneider, 2018; Phillips, Kim, Lonigan, & Connor, 2015; Piasta, LARRC, & Jiang, 2016; Wanzek, Swanson, Vaughn, Roberts, & Fall, 2016) as well as those outside the RfU initiative (e.g., Elleman, Lindo, Morphy, & Compton, 2009; Fuchs et al., 2018; Lesaux, Kieffer, Faller, & Kelley, 2010) have noted how hard it is to move the needle on reading comprehension measures, especially distal measures, with even the most comprehensive, well-designed, and faithfully implemented of interventions. Earlier, in Chapter 5, we raised the same concern, adding to the mix the findings of Hill, Bloom, Black, and Lipsey (2008) and Lortie-Forgues and Inglis (2019).

The vexing question is: "Why?" Informed by the RfU teams, other scholars in the field, and our own experience as researchers in the same community, we offer several plausible explanations, ever cognizant of the perils involved in asserting or even implying causation.

Insensitive Measures

Perhaps the most common argument for the sparse and small reading comprehension effects in the pedagogical literature is the lack of instructional sensitivity of the assessments that have traditionally been used (see Chapter 3 in this volume; Pearson, Valencia, & Wixson, 2014). We just have not had or used measures of reading comprehension that are sufficiently sensitive to the types of interventions implemented in the RfU initiative or a host of other instructional programs that have surfaced over the past several decades, roughly since comprehension rose to prominence in reading pedagogy in the early phases of the cognitive revolution (Pearson & Cervetti, 2017).

There is substantial evidence documenting the difficulty of moving the needle for the very sort of distal measures we tend to demand as evidence of far transfer: intervention-unrelated standardized measures such as the Gates-MacGinitie Reading Test (GMRT) or the Woodcock-Johnson IV Tests of Achievement (WJ-IV). For example, Scammacca, Roberts, Vaughn, and Stuebing (2015), in examining intervention studies from two time periods (1980–2004 and 2005–2011), noted an average effect size for distal measures of comprehension of .24 across both time periods. By contrast, the comparable effect size for ALL measures across both sample time periods was .49. Interestingly, they also found across-the-board decreases in all effect sizes for the 2005–2011 sample; a finding that they attributed "at least in part to increased use of standardized measures, more rigorous and complex research designs, differences in participant characteristics, and improvements in the school's 'business-as-usual' instruction that often serves as the comparison condition in intervention studies" (p. 369). Similar findings (greater effect sizes for proximal investigator-designed assessments over standardized measures) have been reported by Bloom, Hill, Black, and Lipsey (2008) and Moran, Ferdig, Pearson, Wardrop, and Blomeyer (2008), in which the proximal-distal comparison was .56 versus .30.

To the credit of the RfU initiative, a major goal of the Educational Testing Service (ETS)–FCRR portfolio of the RfU work (Chapter 3 in this volume; Sabatini & O'Reilly, 2013) was to develop measures of reading comprehension that reflected more ambitious goals, such as applying the knowledge gained from text comprehension to novel problems and projects and assessing a full array of contributing skills and knowledge. In fact, GISA, the RfU assessment system designed to measure the former sort of application of reading comprehension, did reflect a modest effect size and a modicum of instructional sensitivity for CCDD's WG intervention and the READI science intervention, both of which placed a premium on using text understandings as a source of evidence to warrant arguments. Additionally, CCDD's STARI intervention, with its emphasis on ensuring that struggling readers have an opportunity to bolster their entire repertoire of skills within the context of carefully designed thematic units, found an instructionally sensitive set of specific component measures in the RISE assessment.

In contrast, and no doubt partly due to the belief that neither RISE nor GISA was a good fit for their interventions, the other three teams (LARRC, FCRR, and PACT) for both narrative and expository texts chose either more widely used distal measures of reading comprehension (for PACT, the long-standing GMRT and for FCRR, the WJ-IV) or a specially crafted measure (for LARRC, it was their own listening comprehension measure, roughly modeled on the Qualitative Reading Inventory; Leslie & Caldwell, 2006). For PACT, even though they used the GMRT as their distal measure, they developed the Assessment of Social Studies Knowledge (ASK) measure of knowledge acquisition as their most proximal and the Modified Assessment of Social Studies Knowledge and Reading Comprehension (MASK) measure (comprehension items collected from the released items of many state assessments) as anchoring a spot on the continuum between proximal and distal. As reviewers, we were puzzled that neither of the early grade teams availed themselves of the extensive battery of well-designed, carefully validated measures within the FRA assessment system.

One conclusion that can be drawn from the RfU pedagogical portfolio (see the effect size analysis at the outset of Chapter 5) is that the instructional (in)sensitivity of distal outcomes did play a role in shaping our conclusions about how far an intervention might travel from decidedly proximal to increasingly distal measures. Above and beyond the need for assessments with instructional value and sensitivity, it is important to highlight that, like with any other construct, moving the needle on reading comprehension is also a function of transfer. Transfer (Barnett & Ceci, 2002; Day & Goldstone, 2012) is very difficult to achieve and evaluate in education in general, and in reading in particular (Gick & Holyoak, 1980, 1983; Pearson et al., 2014). Moving forward, we need to consider how transfer interacts with the purpose and grain size of reading assessments. For example, if students make progress on assessments of component skills, does that progress also emerge in a more global assessment of reading comprehension, such as GISA? Transfer—being able to apply new skills and ideas in settings and on measures that differ from the instructional context—is, and always has been, a worthy goal (perhaps the gold standard) of curriculum and instruction enterprises. But it is—and likely always will be—a challenge to achieve.

Design and Implementation Issues

A second commonly offered explanation is faulty design and implementation of the interventions themselves. Historically, pedagogical experiments in reading have often been inadequately designed and implemented. They have been guilty of one or more of the following flaws. The studies:

- were underpowered (especially to detect small effects) due largely to inadequate sampling;
- employed samples of convenience rather than intentional or random samples;
- failed to last long enough for treatments to take effect;
- were implemented in schools that did not really want or value the interventions in order to satisfy the demands of random assignment;
- failed to assess fidelity of implementation; and
- until very recently, ignored modern, sophisticated statistical analyses.

Thus, we have not, as a field, been able to say much about either the validity of any effects obtained or, more germane to the current concern, the trustworthiness of null or weak results. Instead, when we did not find significant and/or robust results, we have typically argued that some set of situational issues (usually design, duration, or measurement issues) conspired to sabotage our attempt.

However, as we suggested in our reporting of the curriculum and instruction portfolio in Chapter 5, poor design and implementation were not issues in the RfU initiative. The intervention portfolio is expansive and complex, with a wide range of independent variables, including many of the potentially malleable factors discussed extensively in Chapter 2. These were well-designed and well-implemented RCTs. All of the interventions emanated from a theoretical base (but not always the same theoretical base) about the nature and development of reading comprehension. Hence, it is reasonable to assume that they possessed a kind of prima facie construct validity. They detailed explicit models (or theories of action) of how particular facets of the reading comprehension puzzle can be shaped in instructional settings to elicit changes in performance. The details of the actual interventions were, in general, as well informed by the wisdom of practice as they were by the theories on which they were built. Teachers were involved as co-designers or critics along the way, often in extensive design research efforts. They employed a range of outcome variables, both proximal measures of whether students learned what was taught and distal measures of how "far" the interventions traveled to more general indices of comprehension or learning. They measured teaching as well as learning, always documenting what actually occurred in the intervention classrooms and, most often, in the business-as-usual control groups. In contrast to many prior efforts in pedagogical research, these were well-powered efforts, with samples sufficiently large and well defined to detect even small effects. About the only implementation standard they may not have met is the length of the treatment; most study directors would have wished for more time, more even than a single school year, for their interventions to take root in classroom ecologies. Even so, in comparison to most intervention studies, these were substantial periods of enactment, ranging from 8 to more than 20 weeks.

In short, given the care with which the interventions were designed and implemented, there was every reason to believe, going into the RCT phase of the RfU

initiative, that if there were effective interventions to be found, they would be found in this initiative. Conversely, if the effects proved to be null, weak, or small, that might be the truth of the matter; such results might be all that should be expected from this sort of intensive effort to move the needle. In short, improvements in reading comprehension itself might be more difficult to achieve than previously thought.

Unrealistic Expectations

At the outset of this section, we suggested that many who conduct and review experimental work on curriculum and instruction bemoan the difficulty of finding large and statistically reliable effects. We also noted the empirical reviews suggesting that the typical effect sizes on distal measures in reading hover in the lower regions of the small (.20 to .49) range. We all want large effect sizes, but they are seldom forthcoming and may, in fact, be flatly unrealistic.

We think the same perspective can be applied to the RfU portfolio. As we note in Chapter 5, it is all too easy to look across the results presented in Chapters 4 and 5 with a glass-half-empty perspective. The effects could have been stronger, significant results more plentiful, and results more consistent across groups and measures. We suggest an alternative "glass a bit more than half full" interpretation, namely, that the aggregate RfU results provide grounds for cautious optimism and guidance for future reading comprehension instruction. We argue that the perspective of disappointment, to mix our metaphors, misses the forest for the trees. If, as we believe, the RCTs and efficacy studies within the RfU possessed reasonably robust designs, psychometrically and conceptually trustworthy measures, adequate power, sufficient dosage/duration, and sensitive statistical analyses, then perhaps we have collectively set our sights too high for achievable effect sizes.

Another interpretation is possible, even plausible: Although many results were uneven and varied across multiple RCTs, some promising patterns emerge when we take a broader view of the collective work accomplished during the RfU initiative. The RfU results suggest that carefully developed and orchestrated multicomponent (and intersectional) instruction, when implemented with fidelity by teachers who are supported by robust professional development, can yield effects that are strong enough to move the needle on reading comprehension and a host of related measures, such as vocabulary, knowledge acquisition, application, and many enabling skills. The needle might not move as radically as we desire, but it most certainly has moved in a positive direction. With continued investment in coordinated, collaborative, and extended efforts like the RfU, the field of education is much more likely to see significant progress in instruction and resultant reading comprehension.

Challenge

The RfU studies in which teachers were observed teaching comprehension (Wanzek & Vaughn, 2016) or queried about implementation barriers (LaRusso, Kim, et al., 2016) offer some additional clues about why it is hard to move the needle. First, there is a lot to get in the way of barriers to serious implementation, as LaRusso, Donovan, and Snow (2016) learned. At the top of the list is test prep, which annually disrupts instruction

in the spring, followed closely by the time demands associated with adhering to the school or district curriculum and dealing with student behavioral disruptions. Second, as Wanzek and Vaughn (2016) discovered, teachers are much more likely to possess the knowledge and material resources to implement instruction that targets the "low-hanging fruit" of the curriculum (invoking or providing relevant prior knowledge and teaching key words) than the harder-to-reach fruit (engaging in close reading, interpretation, application, or critique).

Even so, we know from the RfU work that these higher-order activities can be implemented, as we are reminded from the changes in teachers' practices in the READI work on evidence-based argument (Goldman, Britt, et al., 2016), and that when implemented with fidelity (as in CCDD's WG intervention) they can mediate student learning (Lawrence, Crosson, Paré-Blagoev, & Snow, 2015). These contrastive findings suggest, first and foremost, that this kind of teaching and learning is hard, for teachers as well as students. It further suggests (see R. Anderson, personal communication, September 17, 2019; Sun et al., 2020), for work going forward, that careful monitoring to ensure fidelity not only to the treatment but also to goals of both cognitive and affective engagement is required. Many of the RfU multicomponent interventions put a premium on collaboration and conversation. It seems reasonable to conclude, from the successes that the RfU did achieve, that engagement in higher-order talk within collaborative discussions about interesting and even controversial texts might be the most plausible pathway toward more successful outcomes for students and teachers.

This bundle of requirements (collaboration, talk, and thought-provoking texts) might also explain why we struggle to achieve even modest effects. At the very least, this is an important endeavor for next steps in unpacking the pedagogical puzzle around reading comprehension and learning in the presence of texts. But it is equally as important to provide considerable support for both teachers and students when we ask them to stay the course in these collaborative endeavors. This is as true for learning communities that support teachers' focused, ongoing efforts as it is for students when we ask them to collaborate with their peers in challenging comprehension, critique, and composing tasks.

Moving the Needle Differentially

An explicit goal of educational reform in the United States (and surely in much of the rest of the world) is to close the achievement gap between the educational haves and have nots. This goal often comes couched in a moral imperative such as, "America's problem is not the overall low achievement of its students, it's the unconscionable gap between x and y," where we can fill in the x and y blanks with any of several pairs: majority and minority, rich and poor, native English speakers and English learners (ELs). Students' percentile rank within the overall achievement distribution is remarkably stable from grade to grade. As long ago as 1988, Juel documented this phenomenon in a longitudinal study of students who struggled with reading (Juel, 1988). Within the RfU portfolio, Lonigan (2016) found a similar resistance to differential change in percentile rank in the longitudinal analyses of student growth over time. The story is that, left to the natural ebb and flow in curricular and pedagogical forces, students are likely to maintain their place in the achievement distribution.

But neither the Juel nor the Lonigan analysis involved interventions that intentionally try to disrupt this stable pattern of achievement across the years. Evidence that we could close the achievement gap would mean that we found student characteristic-by-treatment interactions that benefited lower achievers more than high achievers. That is, in a perfect world, all students would make growth over time, but those who started out low would make differentially greater growth than their initially higher-performing peers. Such a pattern was found only occasionally in the RfU work, as with the increased growth of ELs and other language-minority learners in vocabulary and in perspective taking, compared to English-only students, within the CCDD WG curriculum (see Chapter 4). As we suggested there, the news on student characteristic-by-treatment interactions is mixed and complicated. For some interventions, such interactions did not surface; where PACT worked, for example, it worked equally to the benefit of all the identified subgroups. For other interventions (e.g., the FCRR collection), the patterns of interactions were so complex and inconsistent that they defy explanation: in some grades, for some groups, the intervention outpaced business as usual (BAU), but then the pattern flipped at other grades, with BAU (very occasionally) exhibiting greater growth. Likewise, in some studies, these interaction effects could not be examined because students were selected into the study based on low-level skill in reading comprehension or one of its component skills. Nonetheless, those studies often demonstrated main effects that suggested, at the very least, that students selected into the study outperformed students with similar pre-intervention skills.

To study the differential impact of an intervention on students coming into the study with differing characteristics or performance profiles, researchers must make deliberate choices about sampling, design, and analysis. Without an even more substantial investment than was made within the RfU initiative, it is unrealistic to expect that any single efficacy trial could simultaneously account for both student characteristics and pre-intervention performance profiles in determining what works for whom. Researchers need to make these sampling, design, and analytic choices informed by either theory-based or policy-driven priorities. This means that related lines of inquiry, conducted systematically over time, are needed to establish a complete picture of what moves the needle differentially for our most underserved learners.

Where Does This Analysis Leave Us?

One might be tempted, after encountering all of the factors that deter us from our goal, to retreat altogether from the effort to move the needle. To the contrary, we think that the intractability of the problem, and the signs of promise unearthed by the RfU, should motivate us to redouble our efforts to solve it. As we have suggested, the glass-half-full perspective on what we did learn gives us a foothold to resume the quest. More realistic expectations, coupled with building more multicomponent curricula, reframing instruction as a cultural pursuit, accepting the challenge of the diligence it takes to sustain interventions inside classrooms and within teacher learning communities, and expanding our portfolio of innovative assessments of reading comprehension may be the basis—and the best hope—of future efforts to address the goal of moving the needle.

QUESTION 3: BENCHMARKS FOR GAUGING PROGRESS OF THE RFU INITIATIVE

In Chapter 1, we grounded the RfU initiative in a number of contextual factors—theories, practices, policy initiatives, and trends—that had risen to a level of influence in the first decade of the 21st century such that they necessarily influenced what could or should be done in the name of the RfU initiative. In return, these contextual factors were likely to be influenced by the RfU work; in that sense, they provide convenient benchmarks for assessing the influence of the RfU. Here, we address two consequential theories influencing the development of the RfU initiative, the Simple View of Reading (SVR) and the RAND model of reading comprehension. Then we address adolescent literacy and its sibling construct of disciplinary literacy, both influential developments in the 2010s.

The Simple View of Reading

The SVR served as an explicit jumping-off point for the RfU initiative (IES, 2009, p. 5). The SVR was originally intended to provide a broad model for understanding the role of decoding in reading comprehension and potential sources of reading disabilities (Gough & Tunmer, 1986). The SVR describes reading comprehension as the product of decoding and listening comprehension. In doing so, it specifies that, in general, readers who have underdeveloped skill in quickly and accurately recognizing words (decoding) or in constructing meaning from discourse (listening comprehension) will struggle with reading comprehension. Although each of the contributors is actually quite complex, involving an array of skills and knowledge (see, for example, Francis, Kulesz, & Benoit, 2018), framing comprehension as the product of these two broad contributors has long been viewed as a useful heuristic for understanding sources of reading success and difficulty and shaping the purpose and goals of reading pedagogy. However, the relative simplicity of the SVR also invites scrutiny of its explanatory power.

In assessing the overall contribution of the RfU work in advancing our knowledge about the SVR, we conclude that the RfU effort complicated the SVR substantially by adding to our knowledge about (1) the subcomponents that comprise the key components of listening comprehension and, to a lesser degree, decoding; (2) how those components shift in relation to one another and to the ultimate reading comprehension outcome across the development span for pre-K through grade 12; and (3) what other, including exogeneous, factors need to be considered to allow us to explain more of the variance in reading comprehension, for both the youngest and older readers. Now to the evidence that warrants this conclusion.

Several RfU studies examined the validity of the SVR. This research confirmed previous research that had established the validity and credibility of the model: the vast majority of the variance in reading comprehension is accounted for by readers' skill in decoding and language comprehension, at least among elementary-age students (e.g., LARRC, 2015a, 2015b, 2015c; LARRC & Chiu, 2018). Moreover, early language and code-related skills predict the components of the SVR later in school (LARRC & Chiu, 2018). Wang, Sabatini, O'Reilly, and Weeks (2019) provided evidence for a nonlinear relation between decoding and reading comprehension and the identification of a threshold; below this threshold decoding was only weakly related to reading

comprehension and reading comprehension performance was limited. Decoding above this threshold positively predicted performance in reading. RfU findings like these further enhance the credibility of the model and provide plausible hypotheses about malleable factors that can inform pedagogical interventions.

The SVR was also the backbone of the pedagogical portfolio of both of the primary grade RfU teams—FCRR and LARRC. Many of the single-component interventions of FCRR (e.g., Language in Motion, Comprehension Monitoring and Providing Awareness of Story Structure, Morphological Awareness Training, or Enacted Reading Comprehension) or the constituent components of the multicomponent LARRC intervention (language comprehension, comprehension, monitoring, vocabulary, or text structure awareness) can be viewed as an attempt to expand the infrastructure of the listening comprehension factor in the basic SVR formula (reading comprehension [RC] = decoding [DEC] × listening comprehension [LC]).

Similarly, two of the three major assessment efforts, the FRA of FCRR and the RISE of ETS (see Chapter 3 of this volume), could be viewed as attempting to provide at least a partial answer to the question, "What would you need to assess, if you wanted to assess the major internal components of the three variables (reading comprehension, listening comprehension, and decoding) in the SVR?"

The RfU research also sheds light on limitations of the model. For example, the conceptualization and representation of decoding and language comprehension as "necessary, and thus, of equal importance, for reading comprehension" (Hoover & Tunmer, 2018, p. 304) serves the broad conceptual model, but it may obscure the complex dynamic relations among key variables when applied to understandings about reading development and reading instruction. As the RfU studies attest, in practical terms, the role of components and subcomponents shifts across age and comprehension skill level; in particular, the explanatory power of the decoding component attenuates across grades (e.g., LARRC, 2015b; Lonigan et al., 2018).

Several additional issues regarding the clarity and utility of SVR were raised or left unresolved by the RfU research. For example, it is still unclear what subcomponents belong in each of the two broad SVR constructs. For example, does vocabulary adequately index listening comprehension (LARRC, 2015a; Wagner, Herrera, Spencer, & Quinn, 2015)? Should additional components be explicitly acknowledged in the model (e.g., where should one place the powerhouse factor of declarative world knowledge)? Are there underlying factors (such as fundamental cognitive components like memory or attention) that explain the substantial shared variance between decoding and listening comprehension found in many empirical studies of the model (Catts, 2018; LARRC & Chiu, 2018; Lonigan et al., 2018)?

Although the model accounts for most of the variance in reading comprehension in the primary grades, it may not provide sufficient guidance for the development and application of interventions. Indeed, as Gough, Hoover, and Peterson (1996) declared: "Only a fool would deny that reading is complex. Reading clearly involves many subprocesses, and those subprocesses must be skillfully coordinated" (p. 1). In focusing on two broad predictors of comprehension that are underspecified and difficult to distinguish in the earliest grades (Lonigan & Burgess, 2017), the model offers less guidance about the particular underlying factors that will affect some students' reading comprehension later in school.

Similarly, explaining comprehension for older students may involve unpacking the infrastructure of the SVR (e.g., what is entailed in the listening comprehension component?) or augmenting it with additional facets, such as those investigated in other models. For example, the FCRR team subscribed to a longitudinal elaboration of the SVR called the "lattice model" that accounts for the reciprocal relations between decoding and listening comprehension, as well as other cognitive processes, over time (Connor et al., 2014). Ahmed, Francis, York, Fletcher, Barnes, and Kulesz (2016) validated the Direct and Inferential Mediation (DIME) model (Cromley & Azevedo, 2007) in which background knowledge, vocabulary knowledge, reading comprehension, word reading skill, inference making, and reading strategy use all make significant direct contributions to comprehension in adolescence. Using the RfU data, Francis, Kulesz, and Benoit (2018) also examined an alternative model, one they dub the Complete View of Reading (CVR*i*), that accounts for idiosyncratic variation based not only on readers but also on texts by unifying discourse-based cognitive models of reading comprehension with the SVR. They found evidence of variation in rates of reading growth over time that reflect not only variation between readers in reading skills, but also between texts, which shows evidence of differential impact on readers of differing levels of achievement. For example, expository texts and more difficult texts have a negative impact on fluency (i.e., causing students to read more slowly), but especially so for better readers who adjust their reading rate more than poorer readers as they encounter more challenging texts. According to Francis et al. (2018), these findings suggest that models like the SVR that attribute comprehension entirely to component skills may overlook important variation in how individuals approach the task of reading comprehension across different situations and texts (reflecting the task/activity dimension of the RAND model). As a result, they may thus overlook potential pathways for intervention (see Valencia, Wixson, and Pearson [2014] for examples of what these pathways might look like).

In fact, the CCDD and READI work was based on the hypothesis that the SVR declined in relevance to middle grades reading because it obscured or ignored key elements that are crucial to success in reading literature, history, and science in the upper grades. For CCDD, these elements were academic language skills, perspective-taking skills, and reasoning skills (LaRusso, Kim, et al., 2016). For READI, they included the discourse conventions that render oral and written texts discipline specific and the complex set of reasoning skills that define evidence-based argumentation within disciplines (Goldman, 2018). One might argue that the listening comprehension component of the SVR covers academic language, but that interpretation obscures the fact that we are more likely to see than *hear* academic language; the major site for exposure to it is in literate contexts. Similarly, social perspective taking, which starts early with the development of Theory of Mind (Brown-Schmidt, 2009), is not fully accounted for in the broad label of listening comprehension because it involves the ability to infer and project the likely different perspectives of multiple participants in a social scenario using more than just linguistic cues. Finally, given that many texts in literature, science, and social studies require following a multistep, often probabilistically or conditionally stipulated cascade of sequential and/or causally related claims, the ability to follow the logic of complex arguments comes into play as a determinant of successful comprehension (Goldman, 2018; Snow, 2018).

As we suggested at the outset of this analysis, the RfU work advanced our understanding of the SVR by complicating the range of subcomponents that influence listening comprehension and decoding, the shift in influence of within-word and language factors across the developmental span, and the gradual entry of exogenous variables as explanatory factors.

In a recent article (Hoover & Tunmer, 2018), two of the developers of the SVR note that the original intent of the model was to suggest that, "at the broadest level of analysis," reading comprehension is determined by decoding (or word recognition) and language comprehension (p. 304). It is at the broadest level of analysis that the SVR continues to be most useful. It still provides a useful heuristic for conceptualizing and discussing the major "clusters" of factors that account for reading comprehension. The work that remains to be completed is to better understand, and ultimately validate, the key components that constitute the components, particularly the listening comprehension component, across levels of development. However, the collective RfU findings suggest several promising avenues, not only for a better elaborated and more global theory of reading comprehension, but also one that better specifies promising pathways for intervention.

The RAND Reading Study Group

In 2002, the RAND study group posed a set of challenges that could serve as a blueprint for guiding a research agenda specific to reading comprehension. For example, the RAND group heuristic suggested that future research should focus on the independent and joint influence of the reader, task or activity, and text, all of which are nested within sociocultural contexts, on comprehension.

Among the reader factors that received significant attention in the RfU research was reader knowledge, including the quality of that knowledge (see CALI, PACT, and READI) as well as its range. In particular, the RfU (see Chapter 5) portfolio moved us beyond the familiar triad of declarative, procedural, and conditional to include both disciplinary and epistemological knowledge.

The RAND study group proposed that there was a lot to be learned about the influence of text features on comprehension. The RfU teams extended the range of text features to include unfamiliar content using complex language forms (Shanahan et al., 2016), novel syntactic constructions, discourse organization, linguistic markers, multisyllabic words (academic language), and metalinguistic terms. In addition, the RfU researchers studied the role of **s**equencing texts to build vocabulary and knowledge. The texts included ambiguous story characters, unexpected plot developments, and the representation of contrasting positions. Text features were considered in designing all of the intervention and curriculum materials; for example, the SoGen units developed by CCDD intentionally offered relatively small chunks of text (providing information in lists of facts to be sorted into "pro and con" for the debates, for example), rather than longer, denser paragraphs. Furthermore, alternative text types (videos, cartoons, etc.) were used both to present information and as targets of analysis in both WG and READI units. Texts were an especially important influence on the instructional routines and settings within READI, with a special emphasis on scaffolding and supporting students as they grappled with complex, challenging texts;

in particular, disciplinary tasks and social supports were critical in helping students feel comfortable with challenging texts, even leading to cognitive and affective reflection about their encounters with texts.

With respect to readers' tasks and activities related to comprehension, the RfU teams augmented the more traditional classroom practices of recalling and summarizing by having students determine the meanings of unfamiliar words and constructs, analyze text structures, recognize intertextual references, integrate information across texts, and transform text-based information into knowledge that could be used to construct arguments, explanations, and even reports and projects. This work was especially prominent in the work of the adolescent teams (PACT, CCDD, READI), but it was also present in FCRR's CALI. The adolescent teams also expanded study of the purposes for which readers read to include solving problems using text-based information, critiquing arguments, and building arguments (PACT, CCDD, and READI).

The RAND study group raised questions about the role of direct comprehension instruction versus instruction that was embedded in inquiry and authentic reading. Supporting ways to embed comprehension in inquiry and reading for authentic purposes was prominent in the RfU work, especially in the teams that focused on adolescents. For example, PACT attended to cognitive and motivational aspects of the reading process in the design of their interventions; CCDD examined academic language, perspective taking, and reasoning skills; and READI focused on reading and reasoning in different disciplines by attending to oral discourse frames, text genres, and academic language that distinguish disciplines of history, science, and literature. A number of the interventions attended to establishing an explicit purpose for reading that went beyond answering questions or passing a test; for example, the interventions used essential questions (e.g., PACT) or explicit unit goals connected to students' lives and experiences (e.g., PACT and CALI), juxtaposed texts (e.g., READI), conducted highly focused debates (e.g., WG), and used peer participation structures (pair-share, team-based learning; e.g., PACT, READI). In both WG and STARI, students read texts that were chosen to be of interest and relevance to readers (e.g., immigration, nontraditional families), that were organized in thematic units, and that posed discussable questions. Both programs made efforts to align texts and topics to curriculum standards. They emphasized concepts and vocabulary that were specific to the disciplines, and they selected texts to deepen knowledge.

The RAND study group wondered about the relative power of various instructional delivery systems. There were multiple modes of delivery systems investigated across the RfU sites. Most of FCRR's CTT curricula used scripted approaches, but the CALI curriculum used semiscripted lessons, and Word Knowledge e-Book produced technology for independent but guided practice in reading for meaning. PACT produced its own modules for history and also developed technological curriculum assets. CCDD produced supplementary curricula for both WG and STARI. READI did not produce curricula, but instead engaged in close design collaboration with teachers planning around district and school curricula, and often, as suggested earlier, involving nontextbook texts. Most teams used professional development, in-class coaching, and professional learning communities to build extended, not single-shot, teacher learning opportunities. Above all, all teams concerned themselves with both supporting and/or measuring the quality and fidelity of implementation.

The contextual factors identified in the RAND document, such as economic resources, ethnicity, neighborhood, and school culture, did not figure prominently in the RFU portfolio, most likely because of the focus in the request for applications on development, assessment, and pedagogy. In the recommendations for future research that we identify in the next section, we suggest how these contextual factors might figure more prominently in future comprehension research.

Adolescent and Disciplinary Literacy

As we suggested in Chapter 1, adolescent literacy had gained traction in the language arts field in the first decade of the 21st century due in no small part to a systemic effort on the part of the Carnegie Corporation to highlight an emerging groundswell of theoretical and practical work on the reading problems facing adolescents in content-laden secondary classes (Biancarosa & Snow, 2006; Snow & Moje, 2010). With attention focused on adolescents, it seemed a natural step to consider the question of whether the practices of these subject-matter classes were general (applying to all subjects) or subject specific. This distinction led many scholars to the idea of disciplinary literacy as a construct we could use to characterize the goals and challenges of reading, writing, and thinking in what we have traditionally labeled subject-matter or content-area classes (science, history, mathematics, and sometimes even literature). Chief among the perspectives arising from this work was the idea that while there might be general reading, writing, and learning practices, there were also likely to be subject- or discipline-specific practices—or at the very least discipline-specific instantiations of more general practices (Lee & Spratley, 2010; Shanahan & Shanahan, 2008). Also prominent in the disciplinary literacy perspective was the idea that the language of texts and talk about key ideas varied across disciplines—that there were indeed discipline-specific vocabulary and discourse patterns, and even ways of thinking and epistemologies. Phrases like "thinking like a historian" or "reasoning like a scientist" became more common. So how did the RfU initiative influence the ways we think about adolescent and disciplinary literacy?

Combined, the RfU teams that focused on older students demonstrated that the notion of transitioning from learning to read to reading to learn is a false dichotomy. Students in grades 5–12 must learn new strategies, stances, and forms of knowledge to fully comprehend school texts. Recall that three RfU projects (PACT, CCDD, and READI) were funded to attend specifically to older students (grades 5–12). Even though the projects differed in important ways, they were united by their interest in addressing the unique challenges experienced by students as they move from the intermediate grades of elementary school into middle and secondary school. Specifically, these challenges include (1) the amount of unfamiliar content presented in texts, rendering less effectual the typical strategy of encouraging students to use their prior knowledge to make connections and draw inferences; (2) the complexity of academic language encountered in text (including unfamiliar, multisyllabic words and less familiar (and seldom used) syntactic constructions); and (3) the task demands associated with, for example, integrating information from multiple texts, critiquing arguments for claims made in texts, and building one's own arguments from text-based evidence.

Goldman, Snow, and Vaughn (2016) summarized the similar practices that emerged across their three projects, given these challenges. The first of these is

active, purposeful, engaged reading, which entailed identifying explicit goals for reading that were connected to students' lives—for example, by posing a controversy to which students could relate that would be addressed in the text. Another example included the use of essential questions to which the students returned as they read. To support engaged reading, all three projects used nontextbook texts, often replacing textbooks with shorter texts that were sequenced in increasing difficulty and contained information germane to the essential question, or that supported the construction of arguments and or explanations. The second common practice was *social support for reading*. Working in pairs or small groups, students prepared for debates, jointly wrestled with the ideas in the text, and shared common challenges and successes in interpreting and learning from text. Whole-class discussions were used as occasions to model repair strategies and as occasions for teachers to teach disciplinary-specific uses of language and reasoning. The third feature that was characteristic across the projects was *promoting deeper learning* by activating prior knowledge and positioning readers to apply the information they were acquiring to solve a novel problem or articulate an explanation. These three features are critical for understanding the importance of a more "cultural" understanding of comprehension practices as a way of helping students come to terms, as they move through their schooling career, with increasingly challenging and complex literacy activities.

> My practice moved from attention to plot and asking students to make surface connections to characters ("I know men like Rasheed.") to attention to language and the way in which it helps the reader to understand characters, theme, etc. This shift in my planning made a difference in the way students talked about literature.
>
> *—RfU Participating Teacher*

These RfU teams also helped to refine what it means to take a disciplinary stance toward language and learning within secondary classrooms, helping us come to understand "disciplinarity" in several manifestations:

- Representing knowledge, including grappling with the epistemology question of how we know what we know (or do not know);
- Deploying specialized reading comprehension strategies that can help to crack open the puzzles of opaque language, both vocabulary and syntax;
- Engaging in discourse practices that define how we explain phenomena and argue about the validity of competing explanations; and
- Pursuing goals (often taking the form of projects or solutions to problems) representative of the discipline (see Goldman et al., 2019).

This disciplinary knowledge complements the declarative and procedural knowledge that is necessary for literal and inferential interpretation of text; it allows student readers to move beyond literal and inferential comprehension to forms of understanding that include analysis, critique, evaluation, and, above all, integration (Goldman et al., 2019; Shanahan et al., 2016). A clear finding across these adolescent teams is that curriculum and instruction in upper grades must attend to students' ongoing need to learn

to read texts and to participate in tasks of increasing complexity and challenge. That is the essence of our new understanding of disciplinary literacy.

In 2011, Wilkinson and Son proposed that we were on the verge of taking a dialogic turn in comprehension instruction, emphasizing dialogue (talk!) as a medium for paying more attention to discourse and collaboration as a means of improving comprehension and learning in our schools. It seems clear that the RfU teams that focused on adolescents shifted the emphasis of comprehension instruction to just such an agenda; in the work of the adolescent teams, students actively and collaboratively constructed and extracted meaning from texts, used language in the form of discourse to sharpen and deepen their understanding, and applied the knowledge gained from reading, thinking, and talking to solve problems and explain how and why things work the way they do.

QUESTION 4: WHAT MIGHT THE RFA FOR RFU 2.0 LOOK LIKE?

We want to close our stocktaking by looking toward the future and proposing what we think are the absolutely essential initiatives for the literacy research field to undertake in order to "write the next chapter on reading comprehension." To prepare for such a proposal, we begin by summarizing the future research agendas implicit if not explicit in the core chapters on development, assessment, and pedagogy.

Research Priorities for Nature and Development, Assessment, and Curriculum and Instruction

In Table 6-1, we remind readers, in highly truncated and interpreted form, of the future research priorities identified in more elaborated form in Chapters 2, 3, and 5.

TABLE 6-1 Recommendations for Future Research Across the Three Strands

Strand	Issue	Recommendation
Nature and Development	Discrete versus interactive language development	Determine whether rich and broad language experiences develop multiple aspects of language concurrently compared with the independent development of discrete language skills.
	Individual differences in cognitive and attentional skills	Examine whether metacognitive, cognitive, and attentional skills have a critical role in comprehension for particular groups of students, suggesting different pedagogical pathways to improvement.
	Knowledge as a broader mediator	We know a great deal about the mediating role of knowledge for comprehension but not for attentional and retrieval processes on the way toward more facile inferencing or monitoring.
	Linking vocabulary and knowledge	If the semantic and conceptual facets of word knowledge are emphasized over definitional knowledge, there might be synergistic growth in both vocabulary learning and knowledge acquisition.

continued

TABLE 6-1 Continued

Strand	Issue	Recommendation
Assessment	Authenticity	Evaluate the validity and utility of using knowledge gained from comprehension as a deep index of comprehension.
	Theory: process versus componential measures	Use evidence from assessments to evaluate competing theories, such as assembly versus orchestration of key components.
	Instructional sensitivity	Develop measures of global reading literacy for younger readers while also refining those for older readers.
		Determine elements of global literacy appropriate at different age levels, populations, and disciplines.
		Evaluate the everyday utility of global measures.
	Instructional value	Determine if training test-identified specific skills improves more general comprehension performance.
		Evaluate the feasibility of formative measures of global comprehension to complement summative measures.
	Complexity	Explore how to increase the depth of learning required to complete tasks in order to expand the ceiling for comprehension measures.
	Prior knowledge	Examine the increase in explanatory power of assessments by including prior knowledge probes as possible mediators.
Curriculum and Instruction	Emergent bilingual learners (EBs)	Additional research focused on all underserved, but especially EB, populations to help teachers develop more effective practices and close gaps.
	Pedagogical theory (assembly versus orchestration)	Compare the relative merits of assembly versus orchestration models of acquiring skills and knowledge.
	Engagement	Evaluate ways to embed engagement and motivation as inputs (malleable factors), outcomes (measuring the constructs), and mediators (catalysts for enhancing comprehension and learning).
	Text	Examine ways of positioning text in a more central role in our pedagogical research, as a malleable factor rather than simply a medium for discussion.
	Critique	Evaluate the role that a mindset for critique plays in shaping a purpose for close reading and comprehension (see Recommendation 1 below).
	Measuring teaching	Find ways to describe "business-as-usual" conditions with the same care and detail with which interventions are described.
	Transfer	Develop better approaches to scaling and describing the degree of alignment between assessments and interventions.
	Metacognition	Determine optimal approaches to teaching metacognition as the natural counterpart to comprehension and comprehension-related tasks.
	Knowledge	Determine the breadth and depth of prior knowledge that are necessary for comprehension and knowledge application.

Overarching Research Priorities for Reading Comprehension

In addition to these highly specific recommendations (some of which, such as the assembly-versus-orchestration issue, emerge in all three strands), there are several recommendations for a future RfU agenda that are more overarching in character and, as such, are elaborated on below.

Recommendation 1: We need to incorporate the relatively new perspectives of new literacies, digital literacies, and multiliteracies into our comprehension research portfolio. As we suggested in Chapter 1, it was never the expectation that the RfU portfolio would necessarily extend to the recently minted cluster of theories and practices that includes perspectives, practices, and affordances that are relative newcomers to the literacy scene. In fact, terms like "new" and "digital" literacies are not explicitly mentioned in the RfA for the RfU. Whereas the RfU initiative is clearly grounded in a long tradition of cognitive and, to a lesser extent, social perspectives on how we understand and use what we read in the service of learning, these new traditions are grounded more in the epistemological and theoretical traditions of sociocultural or critical perspectives on literacy. As such, they represent opportunities for cross-fertilization across these currently independent research efforts. We think that the work informing the nature, development, assessment, and instruction of reading comprehension, as instantiated by the RfU initiative, has as much to learn from these new traditions (especially in situating comprehension research squarely in the contexts in which its purpose is instantiated) as the new traditions have to learn from the more cognitively grounded work exemplified by the RfU (especially when it comes to research methods that can be used to warrant explanatory and causal accounts of key relationships).

In this brief review of research on technology-related reading comprehension research, as well as research related to multimodal meaning making (both digital and nondigital) and reading comprehension in out-of-school contexts (see Fitzgerald, Higgs, & Palincsar [2020], a white paper available on the National Academy of Education website, for a more extensive treatment of these developments), our goal is to highlight future directions for reading comprehension research that complement those conducted by the RfU teams. As reading increasingly shifts from traditional print to screens, online platforms, and other digital representations in school, work, and community spaces, readers need increasing facility using search engines to locate information, critically evaluate online information to determine the reliability of the text(s) identified, and use online communication tools, such as email, blogs, or infographics, to communicate information. Across the past decade, researchers have investigated a number of questions about reading comprehension with digital text, particularly in the context of the Internet (Leu et al., 2015). Examples include investigations of readers' use of strategies during online reading, pointing to the interplay of new and traditional reading strategies (e.g., Cho & Afflerbach, 2015; Goldman, Braash, Wiley, Graesser, & Brodowinska, 2012); facilitative and detrimental cognitive and social processes during online inquiry (Coiro, Sekeres, Castek, & Guzniczak, 2014); contextual factors that may influence online research and comprehension (e.g., Kennedy, Rhoads, & Leu, 2016; Leu et al., 2015); and how learners evaluate the quality of online information (Coiro, Cosacarelli, Maykel, & Forzani, 2015). We suspect that these will continue to be important lines of inquiry.

In a related area of research, a number of studies have focused on identifying facilitative and detrimental cognitive and social processes during K–12 students' online inquiry. While some cognitive and social processes appear to facilitate students' performance on online research and comprehension tasks, others inhibit performance (e.g., Castek, Coiro, Guzniczak, & Bradshaw, 2012; Cho, Woodward, & Li, 2017; Coiro, Sekeres, et al., 2014; Delgado, Vargas, Ackerman, & Salmerón, 2018). While studies are emerging that differentiate more and less successful online reading, there is limited research that speaks to instructional practices and tools that K–12 teachers might adopt in order to foster student's performance on online research and comprehension tasks, and how instructional practices and tools might differ across grade levels; this is another area ripe for inquiry.

While the Internet has the potential to democratize access to vast quantities of information, it also places unprecedented responsibility on readers to evaluate the quality and reliability of the information they encounter online (McGrew, Breakstone, Ortega, Smith, & Wineburg, 2018). Future research in this area could productively include observational research in classrooms to understand curriculum and practices teachers are already using to support students to interpret and evaluate digital text in online spaces, as well as design-based implementation research using multiple qualitative and quantitative research methods and conducted in collaboration with teachers, schools, and districts to design, test, and iterate upon the design and enactment of curriculum materials to support digital literacy. In addition, while some research in this area has focused on students in the elementary grades, the vast majority of the research on online reading comprehension has been conducted in secondary and postsecondary contexts, suggesting that the field needs to expand research efforts to earlier grades in order to better understand how students develop strategies and skills for online reading over time.

There is increasing interest in multimodal literacy, most germane to this report being multimodal composition as a form of comprehension assessment. For example, Kesler and colleagues (2016) studied the digital stories created by fifth graders to share their interpretations of historical fiction novels. Analyses suggested that students' multimodal designs showed inferential skills, metaphorical thinking, and their understandings of character motivation. The digital stories also made visible to researchers and teachers the limits of students' understandings, such as misconceptions about plot sequence and gaps in background knowledge of historical context. These studies underscore a synergistic relationship between reading and writing.

A number of possible future directions for multimodal comprehension research are warranted. First, the field would benefit from longitudinal research that follows students over time to determine the effects of scaffolded learning in designed digital environments and the implications for students' achievement in and beyond the classroom. Second, although findings suggest that multimodal nondigital texts are useful tools that can support comprehension in K–16 settings and among diverse learners, it would be helpful to know how teachers might support students in learning how to interpret and synthesize communicative modes as they read. Studies suggest that interpretation of even mundane multimodal texts such as picture books or textbooks is a complex process, and one that requires thoughtful guidance and ongoing opportunities to practice. Finally, there is still relatively little comprehension research that focuses on critical literacy and multimodality. While some researchers have explored how

young people understand and evaluate multimodal texts using explicitly sociocritical lenses (e.g., Ajayi, 2015; Begoray, Higgins, Harrison, & Collins-Emery, 2013), studies that consider how integration of modal resources can support learners' inferential and sociocritical understandings of texts are still uncommon.

Finally, studies of comprehension in out-of-school settings have attended to varied contexts, disciplines, and age groups, including a STEM program for nondominant middle school girls (Pinkard, Erete, Martin, & McKinney de Royston, 2017), an afterschool literacy program for recent-arrival immigrant teenagers (Park, 2016), and a summer science and data literacy camp for high school students (Sommer, Hinojosa, Traut, Polman, & Weidler-Lewis, 2017). While there is clear interest in what young people read and how they make meaning of texts outside of school in digital and nondigital environments (e.g., Hutchinson, Woodward, & Colwell, 2016; Jiménez & Meyer, 2016), a sustained line of inquiry related to reading comprehension in out-of-school spaces is not yet clear. Indeed, many of the questions raised by Hull and Schultz (2001) in their review of research related to out-of-school literacy learning remain salient directions for future research almost two decades later. For example, more research is needed to understand reading comprehension in out-of-school spaces and its relationship to in-school learning, including how to bridge students' out-of-school worlds and lived experiences with classroom practice, how to leverage learning in afterschool and other "school-like" spaces in the classroom, and how to support teachers to view and leverage students' out-of-school meaning-making practices as assets for classroom learning.

Recently, Ito et al. (2020), summarized a decade of research, conducted by the Connected Learning consortium to address gaps between in-school and out-of-school learning. Ito et al. (2020) advocate for research that asks, for example, (a) how the field can optimally use "the growing abundance of free and open learning resources to support the learning and interests of diverse young people"; (b) how "new media [can] be mobilized to forge shared rather than divergent interests and literacies between young people, parents, and teachers"; (c) what "new literacies [are] required by the new media ecosystem"; (d) "what forms of measurement, documentation, and evaluation can capture learning across settings"; and (e) how "factors such as social connection, affinity, and belonging influence learning" (Ito et al., 2020, p. 66).

Recommendation 2: We need to develop more precise tools for evaluating the implementation of interventions by incorporating insights from the relatively new field of improvement science. Like most educational researchers, reading researchers are more prone to be guided in their work by developments within rather than outside their own fields of study. But in light of what the RfU teams learned, particularly about just how hard it is to maintain the momentum needed to sustain implementation fidelity—and even more particularly for sustaining collaborative deeper learning practices—perhaps the time has come for scholars who do efficacy studies and RCTs within curricular settings to incorporate even more principles and tools from other fields. In particular, we think that pedagogical researchers have much to learn from the relatively new but rapidly exploding field of improvement science (Bryk, Gomez, Grunow, & LeMahieu, 2015; LeMahieu, Grunow, Baker, Nordstrom, & Gomez, 2017). Important in the field of improvement science is moving toward metrics that assess not only what individual players are learning (e.g., measures of *student* learning or *teacher* fidelity) but also indicators of *system*

learning, where the degree to which entities like schools, districts, and collaboratives are also assessed for the enhancements or barriers they construct in reform efforts. In the process, some constructs change. So, for example, implementation fidelity gets replaced by implementation integrity (LeMahieu, 2011), where the consequential index is not how closely the implementation of the reform matches the "ideal" but how well it is situated in a particular context of implementation. We think research efforts, even tightly controlled RCTs, would benefit from a more ecologically sensitive approach to examining the constraints and affordances of implementation, especially when we have compelling evidence of their consequential influence on research outcomes.

Recommendation 3: We need to add both breadth and depth to our study of the knowledge-comprehension relationship. The aphorism that we learn what is new in terms of what we already know has been with us probably from the onset of human cognition—as a matter of folk wisdom—and from the early days of educational and psychological studies of human cognition (e.g., Thorndike, 1917)—as a matter of empirical documentation. And the RfU scholars often referred to it as a key factor in both research design (e.g., Vaughn et al., 2015) and interpreting results (Goldman, Snow, & Vaughn, 2016). In particular, we see two directions for this expansion, one that looks inward to the RfU work and the other more outward looking.

Testing the power of the RfU-expanded view of knowledge to help us understand and improve comprehension. We acknowledge the importance of the RfU initiative's contribution of highlighting the role that disciplinary and epistemic knowledge—over and above the traditional triad of declarative, procedural, and conditional knowledge—play in describing and improving what students must do to read complex content with deep understanding. At the same time, we assert that we have much to learn about the potential value added of these newer forms of knowledge. At a basic level, we do not know how independent these allegedly distinct forms of knowledge are. Do they develop independently or in concert with one another? Are these two new categories important only for older students, beginning perhaps in middle school, or are they equally important for younger readers? In what ways do students really learn to read like historians or apply knowledge like scientists as they advance through school? If one looks at the Common Core State Standards (and other related state standards documents) or even the Reading Framework for the National Assessment of Educational Progress, our current sources of guidance assume a march toward disciplinarity in thinking about pedagogy and assessment in the service of reading for understanding; indeed, the disciplinary grounding of the RfU work in curriculum and instruction reinforces that perspective. But there are basic and applied research efforts that should be undertaken before we dismiss the idea that there might also be some value in more generic constructs and instructional practices. We think the expansion from the RfU work is important and influential; however, prudence suggests that we continue to examine and refine the power of this expansion.

Expand the scope of the work on the relationship between knowledge and comprehension. In addition to incorporating research on newer categories of knowledge championed in the RfU portfolio, there is still a great deal of unfinished business on the knowledge–comprehension relationship within the realm of more conventional categories of knowledge, such as the familiar triad of declarative, procedural, and conditional knowledge.

Building knowledge within language arts instruction. The RfU research has added to the substantial body of research documenting the significant, positive impact of topic knowledge on reading comprehension, particularly among adolescent readers. Children and adolescents are asked to read texts on a wide range of topics in their lives as students. Starting early in building knowledge of the topics they are likely to encounter is one of the most promising ways to ensure they will successfully comprehend the increasingly complex texts they encounter. As many literacy researchers have pointed out, we have often viewed reading primarily as an opportunity for strategy and skill development, even when the texts are content rich (e.g., Neuman & Celano, 2006; Norris et al., 2008; Palincsar & Duke, 2004), and we have reduced time devoted to content area instruction in the early grades, often without considering the consequences for students' continued literacy development. Even so, the teaching profession is faced with the strong likelihood that the English Language Arts (ELA) block will continue to dominate curricular space (over science and social studies) at the elementary level. We should evaluate opportunities for students to engage with content-rich reading and learning from the earliest years of schools, and ask how ELA instruction can be put to work in building students' knowledge of the natural and social world. A starting point would be to think of literature, as some RfU efforts did, as rich in the content of understanding the human experience—with an emphasis on the big themes of love, friendship, conflict, betrayal, empathy, interacting with the environment, and the like. Literature may prove to be as rich a source of knowledge as science and history.

Leveraging knowledge for other facets of literacy development. The RfU research—and much research on knowledge and comprehension—has focused on the role of topic knowledge in helping students comprehend text by filling gaps and establishing conceptual coherence. A small, but intriguing body of research suggests that knowledge may have a broader role to play in literacy development, supporting students' incidental acquisition of word knowledge as they read (e.g., Barnes, Ginther, & Cochran, 1989; Cervetti, Wright, & Hwang, 2016; Kaefer, Neuman, & Pinkham, 2015; Pulido, 2004), and supporting their acquisition of comprehension strategies (Gaultney, 1995). Future research might well focus on this sort of reciprocity. A common heuristic among practicing teachers goes something like this: if you are presenting a new process for students, situate it in familiar content; and if you are presenting new content, situate it in a familiar process. Investigating the efficacy of such a heuristic could add valuable insights to how we think about the knowledge–literacy relationship.

Leveraging student interests and cultural knowledge. An essential and complex question for future research is how to leverage students' experiential and cultural knowledge in the interest of their literacy development. Studies have demonstrated that cultural knowledge supports students' text comprehension (e.g., Bell & Clark, 1998; Kelley, Siwatu, Tost, & Martinez, 2015; McCullough, 2013; Pritchard, 1990; Pulido, 2004). Although there is promising research demonstrating that cultural knowledge impacts text comprehension, this type of knowledge has yet to be used purposefully in classroom instruction with the goal of supporting students' reading comprehension. There remains substantial work ahead in bringing together two rich research traditions: (1) research on instructional programs that build reading comprehension, and (2) research documenting the efficacy of sociocultural funds of knowledge (Moll,

Amanti, Neff, & Gonzalez, 1992) and both culturally sustaining (Paris, 2012) and culturally relevant (Ladson-Billings, 2014) pedagogy.

Recommendation 4: More of our work on comprehension needs to be directed toward populations currently underserved in U.S. schools. The list of currently marginalized populations is long because it includes cultural and minoritized groups and children of poverty irrespective of race, ethnicity, or home language. But at the top of the list should be emergent bilingual or translingual learners. The particular irony of EBs is that, even though they bring rich language experiences to the classroom, we seem unable to exploit their first language or interlingual/translanguage (first- to second-language connections) resources to craft effective programs for deep reading experiences in English as a second language. Developing curriculum, as well as assessments, that exploit their linguistic resources is a special challenge that scholars of comprehension need to embrace.

Two loosely coupled but separate issues complicate this recommendation. First, we need an explicit effort to include underserved populations in experimental studies for equity purposes. If they are included just incidentally (as a part of a random sample of the entire population, for instance), they will be underrepresented and inappropriate conclusions and recommendations about what works for particular populations will be made. Second, there are important theoretical issues about the relationship of language to knowledge and comprehension that can be addressed only if they are included. This is doubly important for EBs because, for them, knowledge is constant, but proficiency in the two languages in which they operate will vary. To fail to target this population is a missed opportunity to better our understanding of the relationship between language, knowledge, and comprehension.

> Close reading skills became a foundation of the class. We stopped reading just to read, but we began dissecting while we read. My ESL students were able to easily and readily admit when they came across words they struggled with. They began developing tools to deal with words they struggled with. They were able to visualize their reading throughout a text. They were able to set goals before reading a text.
>
> *—RfU Participating Teacher*

Recommendation 5: Writing, especially writing in response to reading and learning from text, is a likely candidate for improving reading comprehension. Writing as the natural complement to and outcome of reading comprehension (Collins, Lee, Fox, & Madigan, 2017; Graham & Hebert, 2011) was implicit in all of the middle and high school interventions—CCDD, PACT, and READI. Sometimes it took the form of group work that required students to collaborate on a joint project (PACT), sometimes the development of arguments about key issues in the text (READI), and sometimes short syntheses and perspective taking on key issues across a set of texts (WG). But in all cases, the writing tasks served the function of promoting integration of key ideas unpacked in one or more texts. Much remains to be examined vis-à-vis the role that writing plays in promoting integration and analysis of key textual ideas. An underexplored area is the role that writing can play in units in which writing is deployed

systematically as students encounter multiple texts along the way to producing some sort of culminating product (an argument, an essay, or even a website design) that documents what students have understood and learned across texts they have read.

Recommendation 6: We need to redouble our efforts to understand, measure, and organize instructional experiences to promote students' language skill and knowledge. The RfU teams embraced language as a major component of and/or contributor to reading comprehension in all three strands of their efforts—nature and development, assessment, and pedagogy. Language is reflected most broadly in the logic of the SVR (RC = LC × DEC). In this regard, one of the RfU teams (LARRC, 2017) has presented the case that the LC component in the SVR model can more profitably be thought of as oral language comprehension than just listening comprehension, to drive home the point that it is language, not just listening, that is necessary to have a better understanding of the nature and development of reading comprehension. Because this recommendation overlaps considerably with Recommendation 7 regarding the relative merits of assembly versus orchestration, we defer our list of language-related possibilities to the next recommendation.

Recommendation 7: Given the prevalent tension within the RfU initiative between the assembly and orchestration models of skill acquisition, the field (perhaps with the leadership of IES) should undertake a major national initiative, including meta-analyses of existing research and new research studies, to evaluate the relative merits of competing theories of the process and pedagogical models of delivery. Albeit with different terminology, the issue of which metaphor—assembly or orchestration—better captures the character of reading (and reading comprehension) development arose in each strand. Chapter 2 referred to the discrete versus connected development of skills. Chapter 3 contrasted process versus componential assessments of reading comprehension. Chapter 5 discussed the tension between assembled versus orchestrated approaches to individual skill instruction, acknowledging that this tension revealed, at its core, a pedagogical grain size issue.

As we suggest in a similar recommendation for pedagogy in Chapter 5, the RfU teams varied considerably in their theoretical position on this tension. Anchoring the atomistic components end of the continuum was FCRR, with its theoretical grounding in the lattice model (and its implicit search for the ideal set of components for a given student), and its quest, along with LARRC, to populate the LC factor in the SVR formula (RC = LC × DEC) with a curated collection of language structures and routines that might ultimately drive reading comprehension. At the orchestrated activity end of the continuum stood READI, with its commitment to situating comprehension practices within the context of discipline-based learning modules that employed collaborative learning, close reading of texts to acquire knowledge to use in constructing evidence-based arguments, and engagement in the discourse practices of the discipline. The work of CCDD and LARRC leaned toward the READI end of the continuum, and PACT seems best positioned squarely in the middle. Much work needs to be completed on this important but enormously complex issue.

Conduct close examinations of the skill infrastructure of older readers. We have very elaborate analyses of the changing interrelationships among subword level, vocabulary, and

comprehension skills from pre-K through grades 4 or 5. Save for DIME and the recent work from the RfU (e.g., Francis et al., 2018; Jones et al., 2019), we do not possess a rich database for older readers. We need to expand our understanding of these interactive developments during adolescence.

Follow through on the logic of the FCRR approach. As we have suggested, one can conceptualize the FCRR approach as answering the question, "What might the infrastructure of the LC term in the SVR look like if we unpacked it with the same care and fervor as has been accomplished for the DEC term over the last 30 years?" Even though the initial attempt to accomplish that goal was only partially achieved (the mixed results and unfinished analyses reported in Chapter 4)—and even though the evidence for orchestration is stronger than for componential assembly—we think there is merit in staying the course to determine whether and what key malleable facets of LC might look like. In particular, moving beyond simple indicators of vocabulary acquisition (i.e., selecting definitions or words to fill a sentence slot) to consider more nuanced aspects of vocabulary and syntactic and pragmatic aspects of language is necessary before closing down such a line of inquiry.

Consider the possibility of middle-ground approaches. Here we suggest that there may be some middle ground between the "assembly" assumption that students only learn what we teach (so we make sure to teach everything separately and explicitly to some level of mastery) and the "orchestration" assumption that some combination of close reading, rich discussion, and collaboration in applying the fruits of comprehension (i.e., the knowledge and insight one acquires from such routines) to authentic real-world tasks will naturally improve skill infrastructure without the cumbersome baggage of heavy-duty skill and strategy instruction. Middle-ground positions might include:

* Emphasizing some "mini-assemblages" or skill clusters (e.g., causal reasoning, predicting, and inferring), and
* On-demand excursions into explicit instruction for components only when formative assessments suggest a mini-intervention.

We have too many convictions and too little empirical evidence to resolve or manage this tension. It is wise, we think, to devote more resources and conceptual energy to understanding and managing, if not resolving, these tensions.

Recommendation 8: In future initiatives in which a separate team is charged with the responsibility of developing relevant assessments for the entire network, employ a different model of assessment development and utilization. We identify two issues regarding the relationship between assessment and its use in evaluating matters of reading development and pedagogy. One is focused on timing, and a second on common measures.

Ensure lead time for assessment development. If the network involves a separate team devoted to the assessment of the core construct under study, along with the enabling skills that feed into it, provide the assessment group a substantial head start (3 years at a minimum) if the core teams examining development and pedagogy are expected to use these measures in their work.

Require common measures across teams. Regardless of the source of the assessments, IES should insist on a core of common measures across projects that focus on similar populations (e.g., kindergarten through grade 3 or grades 6–8). While in the RfU initiative three of the teams (CCDD, READI, and FCRR) used GISA (although FCRR did not use it for their core efficacy studies and CCDD used it only for WG) and both READI and CCDD used RISE (as a pretest control variable for READI and as a progress indicator for STARI within CCDD), there was no common measure across all teams and even when the same measure was used, as noted, it was used differentially. At the very least, common measures across teams would allow for more credible observations (but not direct comparisons) of outcomes across projects. The construct of common measures has been a part of cooperative research since the 1960s, when the First Grade Studies (Bond & Dykstra, 1967) required each of its 22 separately funded and enacted projects to use the same measures for both outcomes and key covariates. It was also a feature of the Follow-Through Studies (Stebbins et al., 1977). A core of common measures, with project-specific options, seems both wise and easy to implement. Even better would be initiative-tailored common measures, the very goal intended within the RfU initiative.

Recommendation 9: Issues of affect and conation should be at the forefront of reading comprehension research. Since the onset of the cognitive revolution in the 1970s, pedagogical research about reading comprehension has been dominated by cognitive strategies and skills. The same could be said of curricular and pedagogical practices devoted to reading comprehension in our schools. The teams in the RfU initiative sought a path to reading achievement marked by innovative curriculum and dedicated teacher professional development in teaching comprehension. While this innovative disposition resulted in many attempts to bring affective or conative factors into the work, as we documented in Chapter 5 (see the Metacognition section on p. 236 and the Engagement section on p. 240), the teams maintained a strong emphasis on cognitive strategy and skill.

While necessary for reading success, cognitive strategies and skills cannot do the job on their own. Readers must be motivated and engaged, and they must possess the self-efficacy that helps power them through challenging texts and tasks. This is especially so for students who struggle in our schools. These are often students whose affective and conative dispositions are ill fitted with the school version of successful reading. We have a critical mass of research that examines students' affect and conation in relation to reading development and reading achievement. Going forward, innovative research should propose the productive marriage of cognition, affect, and conation. All three are essential to development and achievement, yet it is rare to encounter reading comprehension research or instruction based on this acknowledgment.

THE GRAND QUESTION

We close by providing our answer to the grand question: What rewards did we reap from the RfU initiative? There are a few different ways to answer the question.

In the core chapters, we have answered that grand question finding by finding, study by study, issue by issue, theme by theme, and insight by insight across the

portfolio of work produced by these six teams across the past decade—the 5 years of funding plus the extensions and the lingering trail of publications.

We can answer the question with numbers. We could tell you that more than 100 scholars worked with hundreds of teachers and thousands of students in scores of schools in the majority of states across the duration of the RfU initiative. They produced more than 300 research publications, the vast majority of which appeared in top-tier refereed journals. They improved the lives of the students reached by their research by achieving average effect sizes for interventions in the very respectable range of .20 to .80—a sizable and reliable advantage over the control groups. We could even tell you how many downloads and citations their research reports and curricular materials have garnered across the years.

Or we could tell you a completely different story to mark their achievement. We could tell you what a singular qualitative achievement it was to persuade all of those teachers (and all of those students) in all of those schools to work hard to acquire new knowledge, new routines, and new expectations for what students can and should do, and what an achievement it was for the teachers to go on to deliver suitable instruction to all of those students—to get the students to do things that were out of their comfort zone.

We could assert that it is hard, but not impossible, to move that stubborn, sticky reading achievement needle. But it takes a lot of effort and stamina to do so. Teachers have to overcome the temptation to pick the low-hanging pedagogical and curricular fruit and search instead for the higher-hanging and more rewarding fruit—close reading, critique, and a search for evidence to support explanations and arguments. Teachers have to ignore, or work around, the barriers of the required curriculum and misguided accountability schemes with all of the test prep.

Even so, we know from the RfU work that the higher-hanging fruit can be reached, and that when the practices hanging up there are implemented with integrity, they can mediate student learning. This suggests key elements of the success of the effort:

- For most of the projects, there were strong and supportive professional learning communities that maintained high standards and offered sustained support in the form of coaching and careful monitoring.
- Those communities allowed teachers to implement and even sustain engaging but challenging practices.
- Those practices promoted wide and deep student engagement in collaborative discussions about interesting and edgy texts.
- Those conversations were on the pathway not just to comprehension but to applying what students learned to explanations and arguments about important ideas.

What this means is that the job of comprehension is not complete until one uses the resulting understanding to *do* something—tell a story, explain a situation, argue with an author or a classmate, or maybe even plan to change the world. In short, one reading of the RfU is that it has given us a glimpse of what an alternative culture of comprehension pedagogy might look like. The RfU initiative has led us part of the way down that path. And the legacy they left us—both in terms of what we learned and still need to learn—is surely an important road map for taking the next steps in unpacking this important pedagogical puzzle around comprehension and learning in the presence of texts.

REFERENCES

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M. A., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology, 44*, 68–82.

Ajayi, L. (2015). Critical multimodal literacy: How Nigerian female students critique texts and reconstruct unequal social structures. *Journal of Literacy Research, 47*(2), 216–244.

Barnes, J. A., Ginther, D. W., & Cochran, S. (1989). Schema and purpose in reading comprehension and learning vocabulary from context. *Reading Research and Instruction, 28*(2), 16–28.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637.

Begoray, D., Higgins, J. W., Harrison, J., & Collins-Emery, A. (2013). Adolescent reading/viewing of advertisements: Understandings from transactional and positioning theory. *Journal of Adolescent & Adult Literacy, 57*(2), 121–130.

Bell, Y. R., & Clark, T. R. (1998). Culturally relevant reading material as related to comprehension and recall in African American children. *Journal of Black Psychology, 24*, 455–475.

Biancarosa, G., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Bloom, H. S., Hill, C. J., Black, A. B., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289–328.

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 2*(4), 5–142. doi: 10.2307/746948.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review, 16*, 893–900.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press.

Castek, J., Coiro, J., Guzniczak, L., & Bradshaw, C. (2012). Examining peer collaboration in online inquiry. *The Educational Forum, 76*(4), 479–496.

Catts, H. W. (2018). The Simple View of Reading: Advancements and false impressions. *Remedial and Special Education, 39*(5), 317–323.

Cervetti, G. N., Wright, T. S., & Hwang, H. (2016). Conceptual coherence, comprehension, and vocabulary acquisition: A knowledge effect? *Reading and Writing, 29*(4), 761–779.

Cho, B. Y., & Afflerbach, P. (2015). Reading on the Internet: Realizing and constructing potential texts. *Journal of Adolescent & Adult Literacy, 58*(6), 504–517.

Cho, B. Y., Woodward, L. & Li, D. (2017). Epistemic processing when adolescents read online: A verbal protocol analysis of more and less successful online readers. *Reading Research Quarterly*. doi: 10.1002/rrq.190.

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12.

Coiro, J., Coscarelli, C., Maykel, C. & Forzani, E. (2015). Investigating criteria seventh graders use to evaluate the quality of online information. *Journal of Adolescent and Adult Literacy, 58*(7), 546–550.

Coiro, J., Sekeres, D. C., Castek, J., & Guzniczak, L. (2014). Comparing the quality of third, fourth, and fifth graders' social interactions and cognitive strategy use during structured online inquiry. *Journal of Education, 194*(2), 1–15.

Collins, J. L., Lee, J., Fox, J. D., & Madigan, T. P. (2017). Bringing together reading and writing: An experimental study of writing intensive reading comprehension in low-performing urban elementary schools. *Reading Research Quarterly, 52*(3), 311–332.

Connor, C. M., Phillips, B. M., Kaschak, M., Apel, K., Kim, Y. S., Al Otaiba, S., Crowe, E. C., Thomas-Tate, S., Johnson, L. C., & Lonigan, C. J. (2014). Comprehension tools for teachers: Reading for understanding from prekindergarten through fourth grade. *Educational Psychology Review, 26*(3), 379–401. doi: 10.1007/s10648-014-9267-1.

Connor, C. M., Phillips, B. M., Kim, Y. G., Lonigan, C. J., Kaschak, M. P., Crowe, E., Dombek, J., & Al Otaiba, S. (2018). Examining the efficacy of targeted component interventions on language and literacy for third and fourth graders who are at risk of comprehension difficulties. *Scientific Studies of Reading, 22*(6), 462–484.

Cromley, J., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*, 311–325.

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist, 47*(3), 153–176.

Delgado, P., Vargas, C., Ackerman, R., & Salmerón. L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review, 25*, 23–38. doi: 10.1016/j.edurev.2018.09.003.

Elleman, A., Lindo, E., Morphy, P., & Compton, D. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness, 2*, 1–44.

Fitzgerald, M., Higgs, J., & Palincsar, A. S. (2020). *New media, new literacies: Implications for reading for understanding*. Washington, DC: National Academy of Education.

Foorman, B., Dombek, J., & Smith, K. (2016). Seven elements important to successful implementation of early literacy intervention. *New Directions for Child and Adolescent Development, 154*, 49–65.

Francis, D. J., Kulesz, P. A., & Benoit, J. S. (2018). Extending the Simple View of Reading to account for variance within and across readers: The Complete View of Reading (CVRi). *Remedial and Special Education, 39*, 274–288.

Fuchs, D., Hendricks, E., Walsh, M. E., Fuchs, L. S., Gilbert, J. K., Tracy, W. Z., Patton III, S., Davis-Perkins, N., Kim, W., Elleman, A. M., & Peng, P. (2018). Evaluating the efficacy of a multidimensional reading comprehension program for at-risk students and reconsidering the lowly reputation of tests of near transfer. *Learning Disabilities Research and Practice, 33*(1), 11–23. doi: 10.1111/ldrp.12162.

Gaultney, J. F. (1995). The effect of prior knowledge and metacognition on the acquisition of a reading comprehension strategy. *Journal of Experimental Child Psychology, 59*(1), 142–163.

Gersten, R. (2016). Commentary: The tyranny of time and the reality principle. *New Directions for Child and Adolescent Development, 154*, 49–65.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*(3), 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38.

Goldman, S. R. (2018). Discourse of learning and the learning of discourse. *Discourse Processes, 55*(5–6), 434–453.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, & Brodowinska, K. (2012). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly, 47*(4), 356–381.

Goldman, S. R., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C. D., Shanahan, C., & Project READI. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist, 51*(2), 219–246.

Goldman, S. R., Greenleaf, C., Yukhymenko-Lescroart, M., Brown, W., Ko, M. L. M., Emig, J. M., George, M., Wallace, P., Blaum, D., & Britt, M. A. (2019). Explanatory modeling in science through text-based investigation: Testing the efficacy of the Project READI intervention approach. *American Educational Research Journal, 56*(4), 1148–1216. doi: 10.3102/0002831219831041.

Goldman, S. R., Snow, C., & Vaughn, S. (2016). Common themes in teaching reading for understanding: Lessons from three projects. *Journal of Adolescent & Adult Literacy, 60*(3), 255–264. doi: 10.1002/jaal.586.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a Simple View of Reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1–13). Mahwah, NJ: Lawrence Erlbaum Associates.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10.

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*, 172–177.

Hoover, W. A., & Tunmer, W. E. (2018). The Simple View of Reading: Three assessments of its adequacy. *Remedial and Special Education, 39*(5), 304–312.

Hull, G., & Schultz, K. (2001). Literacy and learning out of school: A review of theory and research. *Review of Educational Research, 71*(4), 575–611.

Hutchinson, A. C., Woodward, L., & Colwell, J. (2016). What are preadolescent readers doing online? An examination of upper elementary students' reading, writing, and communication in digital spaces? *Reading Research Quarterly, 51*(4), 435–454.

IES (Institute of Education Sciences). (2009). *Request for applications: Reading for understanding research initiative.* Retrieved from https://ies.ed.gov/funding/pdf/2010_84305F.pdf.

Ito, M., Arum, R., Conley, D., Gutiérrez, K., Kirshner, B., Livingstone, S., Michalchik, V., Penuel, W., Peppler, K., Pinkard, N., Rhodes, J., Tekinbaş K. S., Schor, J., Sefton-Green, J., & Watkins, S. C. (2020). *The Connected Learning Research Network: Reflections on a decade of engaged scholarship*. Irvine, CA: Connected Learning Alliance.

Jimenez, L. M., & Meyer, C. K. (2016). First impressions matter: Navigating graphic novels utilizing linguistic, visual, and spatial resources. *Journal of Literacy Research, 48*(4), 423–447.

Jones, S. M., LaRusso, M., Kim, J., Kim, H. Y., Selman, R., Uccelli, P., Barnes, S. P., Donovan, S., & Snow, C. (2019). Experimental effects of Word Generation on vocabulary, academic language, perspective taking, and reading comprehension in high poverty schools. *Journal of Research on Educational Effectiveness, 12*, 448–483.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 437–447.

Kaefer, T., Neuman, S. B., & Pinkham, A. M. (2015). Pre-existing background knowledge influences socioeconomic differences in preschoolers' word learning and comprehension. *Reading Psychology, 36*(3), 203–231. https://doi.org/10.1080/02702711.2013.843064.

Kelley, H. M., Siwatu, K. O., Tost, J. R., & Martinez, J. (2015). Culturally familiar tasks on reading performance and self-efficacy of culturally and linguistically diverse students. *Educational Psychology in Practice, 31*(3), 293–313. https://doi.org/10.1080/02667363.2015.1033616.

Kennedy, C., Rhoads, C., & Leu, D. J. (2016). Online research and learning in science: A one-to-one laptop comparison in two states using performance based assessments. *Computers and Education, 100*, 1–21.

Kesler, T., Gibson, Jr., L., & Turansky, C. (2016). Bringing the book to life: Responding to historical fiction using digital storytelling. *Journal of Literacy Research, 48*(1), 39–79.

Kim, J. S., Hemphill, L., Troyer, M., Thomson, J. M., Jones, S. M., LaRusso, M. D., & Donovan, S. (2016). Engaging struggling adolescent readers to improve reading skills. *Reading Research Quarterly, 52*(3), 357–382. doi: 10.1002/rrq.171.

Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harvard Educational Review, 84*(1), 74–84.

LARRC (Language and Reading Research Consortium). (2015a). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50*, 151–169.

LARRC. (2015b). The dimensionality of language ability in young children. *Child Development, 86*, 1948–1965.

LARRC. (2015c). The dimensionality of Spanish in young Spanish-English dual-language learners. *Journal of Speech, Language, and Hearing Research, 58*, 754–766.

LARRC. (2017). Oral language and listening comprehension: Same or different constructs? *Journal of Speech, Language, and Hearing Research, 60*, 1273–1284.

LARRC & Chiu, Y. D. (2018). The Simple View of Reading across development: Prediction of grade 3 reading comprehension from prekindergarten skills. *Remedial and Special Education, 39*(5), 289–303.

LARRC, Farquharson, K., & Murphy, K. (2016). Ten steps to a large, multi-site, longitudinal investigation of language and reading in young children. *Frontiers in Developmental Psychology, 7*, 1–16.

LARRC, Johanson, M., & Arthur, A. M. (2016). Improving the language skills of pre-kindergarten students: Preliminary impacts of the Let's Know! experimental curriculum. *Child & Youth Care Forum, 45*(3), 367–392.

LaRusso, M. D., Donovan, S., & Snow, C. (2016). Implementation challenges for tier one and tier two school-based programs for early adolescents. *New Directions for Child and Adolescent Development, 154*, 11–30.

LaRusso, M., Kim, H. Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., Donovan, S. & Snow, C. (2016). Contributions of academic language, perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research in Educational Effectiveness, 9*, 201–222.

Lawrence, J. F., Crosson, A. C., Paré-Blagoev, E. J., & Snow, C. E. (2015). Word Generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal, 52*(4), 750–786. doi: 10.3102/0002831215579485.

Lawrence, J. F., Rolland, R. G., Branum-Martin, L., & Snow, C. E. (2014). Generating vocabulary knowledge for at-risk middle school readers: Contrasting program effects and growth trajectories. *Journal of Education for Students Placed at Risk, 19*(2), 76–97.

Lee, C. D., Goldman, S. R., Levine, S., & Magliano, J. P. (2016). Epistemic cognition in literary reasoning. In J. Green, W. Sandoval, & J. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 165–183). New York: Routledge.

Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. Final Report from Carnegie Corporation of New York's Council on Advancing Adolescent Literacy. New York: Carnegie Corporation.

LeMahieu, P. G. (2011, October 11). What we need in education is more integrity (and less fidelity) of implementation [Blog post]. Retrieved from http://commons.carnegiefoundation.org/what-we-are-learning/2011/what-we-need-in-education-is-more-integrity-and-less-fidelity-of-implementation.

LeMahieu, P. G., Grunow, A., Baker, L., Nordstrum, L. E., & Gomez, L. M. (2017). Networked improvement communities: The discipline of improvement science meets the power of networks. *Quality Assurance in Education, 25*(1), 5–25.

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly, 45*(2), 196–228.

Leslie, L., & Caldwell, J. (2006). *Qualitative Reading Inventory, 4th edition*. New York: Pearson.

Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly, 50*(1), 37–59.

Lonigan, C. J. (2016, May). The FCRR Reading for Understanding Network: Challenges and Unanswered Questions in Pre-K to Grade 3. Paper presented at the National Symposium on Reading for Understanding. Alexandria, VA.

Lonigan, C. J., & Burgess, S. R. (2017). Dimensionality of reading skills with elementary-school-age children. *Scientific Studies of Reading, 21*, 239–253.

Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the Simple View of Reading with elementary school children: Still simple after all these years. *Remedial and Special Education, 39*, 260–273.

Lortie-Forgues, H., & Inglis, M. (2019). Rigorous large-scale educational RCTs are often uninformative: Should we be concerned? *Educational Research, 48*(3), 158–166. doi: 10.3102/0013189X19832850.

McCullough, R. G. (2013). The relationship between reader response and prior knowledge on African American students' reading comprehension performance using multicultural literature. *Reading Psychology, 34*(5), 397–435.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education, 46*(2), 165–193.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice, 31*(2), 132–141.

Moran, J., Ferdig, R. E., Pearson, P. D., Wardrop, J., & Blomeyer, R. L. (2008). Technology and reading performance in the middle-school grades: A meta-analysis with recommendations for policy and practice. *Journal of Literacy Research, 40*(1), 6–58.

Neuman, S. B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for lowş income and middle-income children. *Reading Research Quarterly, 41*(2), 176–201.

Norris, S. P., Phillips, L. M., Smith, M. L., Guilbert, S. M., Stange, D. M., Baker, J. J., & Weber, A. C. (2008). Learning to read scientific text: Do elementary school commercial reading programs help. *Science Education, 92*(5), 765–798.

NRC (National Research Council). (2002). *Scientific research in education*. Washington, DC: National Academy Press.

NRC. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st Century*. Committee on Defining Deeper Learning and 21st Century Skills. Washington, DC: The National Academies Press.

Palincsar, A. S., & Duke, N. K. (2004). The role of text and text-reader interactions in young children's reading development and achievement. *The Elementary School Journal, 105*(2), 183–197.

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher, 41*(3), 93–97.

Park, J. Y. (2016). "He didn't add more evidence": Using historical graphic novels to develop language learners' disciplinary literacy. *Journal of Adolescent & Adult Literacy, 60*(1), 35–43.

Pearson, P. D., & Cervetti, G. N. (2017). The roots of reading comprehension instruction. In S. E. Israel (Ed.), *Handbook of research on reading comprehension, 2nd edition* (pp. 12–56). New York: Guilford Press.

Pearson, P. D., Valencia, S. W., & Wixson, K. (2014). Complicating the world of reading assessment: Toward better assessments for better teaching. *Theory into Practice, 53*(3), 236–246. doi: 10.1080/00405841.2014.916958.

Phillips, B. M., Kim, Y.-S., Lonigan, C. J., & Connor, C. M. (2015, July). *Language for understanding: Two large-scale studies of small-group language interventions in prekindergarten and kindergarten*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Kona, HI.

Piasta, S., Language and Reading Research Consortium, & Jiang, H. (2016, July). *Targeting lower- and higher-level language skills to support comprehension: Initial results for Let's Know*. Paper presented at the annual conference of the Society for the Scientific Study of Reading, Porto, Portugal.

Pinkard, N., Erete, S., Martin, C., & McKinney de Royston, M. (2017). Digital Youth Divas: Exploring narrative-driven curriculum to trigger middle school girls' interest in computational activities. *Journal of the Learning Sciences, 26*(3), 477–516.

Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly, 25*(4), 273–295.

Pulido, D. (2004). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning, 54*(3), 469–523.

Rowan, B., & Correnti, R. (2009). Measuring reading instruction with teacher logs. *Educational Researcher, 38*, 449–551.

RRSG (RAND Reading Study Group). (2002). *Reading for understanding: Toward a research and development program in reading comprehension.* Santa Monica, CA: RAND Corporation.

Sabatini, J. P., & O'Reilly, T. (2013). Rationale for a new generation of reading comprehension assessments. In B. Miller, L. E. Cutting, & P. McCardle (Eds.), *Unraveling reading comprehension: Behavioral, neurobiological and genetic components* (pp. 100–111). Baltimore: Paul H Brookes.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities, 48*(4), 369–390. doi: 10.1177/0022219413504995.

Shanahan, C., Bolz, M., Cribb, G., Goldman, S., Heppeler, J. & Manderino, M. (2016). Deepening what it means to read (and write) like a historian: Progressions of instruction across a school year in an eleventh grade U.S. history class. *The History Teacher, 49*(2), 1–30.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*(1), 40–59.

Snow, C. E. (2018). Simple and not-so-simple views of reading. *Remedial and Special Education, 39*, 313–316. doi: 10.1177/0741932518770288.

Snow, C. E., & Moje, E. B. (2010). Why is everyone talking about adolescent literacy? *Phi Delta Kappan, 91*(6), 66–69.

Sommer, S., Hinojosa, L., Traut, H., Polman, J., & Weidler-Lewis, J. (2017). *Integrating eye-tracking activities into a learning environment to promote collaborative meta-semiotic reflection and discourse.* Philadelphia: International Society of the Learning Sciences.

Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). *Education as experimentation: A planned variation model* (Vol. IV-A). Cambridge, MA: Abt Associates.

Sun, J., Anderson, R. C., Lin, T., Morris, J. A., Biller, B. W., Ma, S., Nguyen, K. T., & Scott, T. (2020). Children's engagement during collaborative learning and direct instruction. Manuscript under review.

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology, 8*(6), 323–332. doi: 10.1037/h0075325.

Valencia, S., Wixson, K. K., & Pearson, P. D. (2014). Putting text complexity in context: Refocusing on comprehension of complex text. *Elementary School Journal, 115*(2), 270–289.

Vaughn, S., Roberts, G., Swanson, E. A.,Wanzek, J., Fall, A. M., & Stillman-Spisak, S. J. (2015). Improving middle-school students' knowledge and comprehension in social studies: A replication. *Educational Psychology Review, 27*, 31–50.

Wagner, R. K., Herrera, S. K., Spencer, M., & Quinn, J. M. (2015). Reconsidering the Simple View of Reading in an intriguing case of equivalent models. *Journal of Learning Disabilities, 48*, 115–119.

Wang, Z., Sabatini, J., O'Reilly, T., & Weeks, J. (2019). Decoding and reading comprehension: A test of the decoding threshold hypothesis. *Journal of Educational Psychology, 11*(3), 387–401.

Wanzek, J., Swanson, E., Vaughn, S., Roberts, G., & Fall, A.-M. (2016). English learner and non-English learner students with disabilities: Content acquisition and comprehension. *Exceptional Children, 82*, 428–442.

Wanzek, J., & Vaughn, S. (2016). Implementation of a text-based content intervention in secondary social studies classes. *New Directions for Child and Adolescent Development, 154*, 31–48.

Wilkinson, I. & Son, E. (2011). A dialogic turn in research on learning and teaching to comprehend. In M. L. Kamil, P. D. Pearson, E. Moje, P. Afflerbach (Eds.), *Handbook of reading research (volume IV)* (pp. 359–387). New York: Routledge.

Wood, D. R., & Lieberman, A. (2000). Teachers as authors: The National Writing Project's approach to professional development. *International Journal of Leadership in Education, 3*(3), 255–273.