The Role of Text Difficulty in Small-Group Reading for Bilingual Students

by

Laura Beth Kelly

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved March 2018 by the
Graduate Supervisory Committee:

Margarita Jimenez-Silva, Co-Chair
Lindsey Moses, Co-Chair
Laida Restrepo

ARIZONA STATE UNIVERSITY

May 2018

ABSTRACT

How hard should the books be in elementary small-group reading? This study explored text difficulty for bilingual students reading below grade level in third grade. Using a convergent parallel mixed methods design, I used qualitative methods to analyze students' engagement and discussion during small groups and single case design to evaluate students' fluency and reading comprehension after reading and discussing texts in small groups.

Six Spanish-English bilingual students, split into two groups of three, participated in twelve, 30-minute, small-group reading sessions. Students in Group 1 read approximately one year below grade level, and students in Group 2 read approximately a year and a half below grade level. In six of the twelve sessions, students read and discussed texts matched to their reading levels, and in the other six they read and discussed texts one year ahead of their reading levels. I assigned matched and difficult texts across the twelve days by blocked randomization.

I analyzed video transcripts of each session to understand students' engagement (focus of engagement, strategies, and interaction) and discussion (inferential vs. literal responses, instances of verbal participation). At the end of each session, students reread and retold the book the group had read and discussed that day to produce a fluency (words correct per minute) and comprehension (ideas correctly retold) score.

Findings were complex and revealed that different levels of texts have both advantages and drawbacks. Key findings included: For fluency, half of the students benefited from matched texts. The other half read difficult texts with similar fluency to matched texts. For comprehension, text difficulty did not matter for anyone except one

student, and for him it only had an effect on 3 of 12 days. Group 2 engaged much more with texts and ideas in difficult books and with pictures in matched books. Group 1 had more inferential/interpretive responses with matched texts, and Group 2 had more inferential and interpretive responses with difficult texts. Most students participated evenly regardless of the difficulty of the text under discussion. However, two students talked more when discussing matched texts.

DEDICATION

To Joseph and Lisette

ACKNOWLEDGMENTS

I thank Joseph, who never doubted that I should get a Ph.D.

I thank the committee members for providing support and feedback throughout the dissertation process. Your availability and promptness has made this process smooth.

I especially thank Mr. Hayes, who made it easy to conduct research in his school; Ms. Gaspar, who welcomed me to work with her students; and the six students who participated—Alyssa, Jack, Rosa, Sarah, Gabriella, and Elise. They made the research a joyful process!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

My third-grade classroom had the standard horseshoe table for small-group reading. Blue lockers lined the entire back wall behind that table. On top of those lockers, literally wall to wall, sat a year's worth of leveled books. For each week of the curriculum, I had 24 books: 6 advanced, 6 on the high end of our grade level, 6 on the low end of our grade level, and 6 below grade level. They were real books by real authors and sold in real bookstores, and the curriculum publisher had determined a readability level for them all. Inside the metal blue lockers, I had more leveled books from years' worth of discarded curriculum. So, every day when I called kids back to read, I wondered, which books would promote the most literacy progress? I had not found that following the conventional wisdom about matching books to readers accelerated learning for my students. I used low books for students with low reading levels, and when we finished, nothing had changed. Something felt off.

Elementary teachers in the United States often provide individualized reading instruction by grouping students of similar reading levels together for small-group reading instruction with texts close to the students' reading levels. Historically, teachers have understood texts to use in these groups at three levels: frustration (texts that students read with less than 90% accuracy), instructional (texts that students read with 95% accuracy), and independent (texts that students read with 99% accuracy). These basic understandings of text difficulty (Betts, 1946), while not uncontested (Halladay, 2012) or originally based on empirical research (Allington, McCuiston, & Billen, 2015), have

played a significant role in guiding many teachers' decisions about text selection for small-group reading, with most teachers selecting texts at the instructional level.

However, the widespread adoption of the Common Core State Standards (CCSS) has changed the landscape concerning text difficulty in elementary classrooms. College and Career Readiness Anchor Standard 10 from the language arts standards requires students to "read and comprehend complex literary and informational texts independently and proficiently" (Council of Chief State School Officers, 2010). Beginning in second grade, the standards raise the Lexile level bands (a measure of text difficulty) for each grade. Because of these new expectations, elementary students, including those learning English, now grapple with increasingly complex texts at school, texts often historically considered "frustration level." Yet, many research questions remain about the ideal text difficulty level for students at different grades, English proficiency levels, reading proficiency levels, and instructional contexts. Further, the role of text difficulty specifically for students learning English (emerging bilinguals) and the affective factors associated with small-group reading (such as students' engagement and discussion) remain underexplored. This dissertation study aims to fill that gap.

Purpose Statement

I conducted this study to compare students' discussion, engagement, comprehension, and fluency when using books matched to their levels and books one year ahead of their current levels. (Because I worked with students reading below grade level, "one year ahead" meant texts from their grade level.) I wanted to understand how text difficulty affects the group dynamic and the students' achievement in small-group reading.

Significance of Problem

Most elementary teachers spend some time working with small groups for reading instruction (Ford & Opitz, 2008), and many students need this support to make grade-level progress in reading (Gersten et al., 2008). For some students, small-group reading with their classroom teacher is the first (or even only) extra support they receive when school officials first notice that they need literacy help (D. Fuchs, Fuchs, & Vaughn, 2014). Studying what levels of text most facilitate discussion, engagement, and literacy achievement matters because so many students participate in small-group reading and schools rely on these groups as a form of intervention for students not making progress in the regular curriculum.

Research Questions

In small-group reading sessions for bilingual students reading below grade level...

1. How does text difficulty impact student engagement?
    a. How do students engage with the books?
    b. What comprehension strategies do students use in discussion?
    c. In what ways do students interact to participate in discussions?
2. How does text difficulty impact small-group discussion?
    a. What types of responses (literal, inferential/interpretive, and incorrect) do students share?
    b. How often do students participate?
3. What is the effect of text difficulty on reading comprehension?
4. What is the effect of text difficulty on reading fluency?

Definition of Terms

When I refer to small-group reading, I mean the practice of teachers pulling aside a group of students for supplemental reading instruction beyond what the class received as a whole group. I refer to the students who participated in this study as bilingual: they all had a home language of Spanish and varying degrees of proficiency in English. Their school district considered them all (at one time) English Language Learners (ELLs) although two no longer received support as ELLs. All of them continued to develop literacy in English and Spanish through their school's dual language program. When I talk about "striving readers," I mean that they read below grade level, and I made this determination based on data their teacher shared about whether they met district-determined benchmarks on literacy assessments.

I describe text difficulty in great detail in the literature review. Throughout this dissertation, I continually refer to "matched texts" and "difficult texts." Matched texts mean books that, according to the text leveling and assessment system I used, fell within the instructional level of the students. Difficult texts came from a difficulty band one year ahead of the students' current reading levels.

I measured outcomes in terms of students' quantitative performance on typical literacy assessment tasks: comprehension and fluency. I describe those measures in the methods section. I also compared students' discussion (how they participated and responded) and their engagement (what strategies they used and how they interacted) through qualitative analysis. I define these terms and explain the measures in more detail in the methods chapter.

Brief Review of Related Studies

Researchers have conducted little empirical research about the effect of text difficulty on student outcomes. Much theoretical writing advocates one difficulty level or another (Allington et al., 2015; Hiebert & Mesmer, 2013). Writings for teachers suggest instructional protocols using either matched or difficult text (D. Fisher & Frey, 2012; Fountas & Pinnell, 1996). Assessment studies that provide no instruction produce unsurprising findings about how students read on-level versus struggling with difficult texts (Amendum, Conradi, & Liebfreund, 2016; Benjamin & Schwanenflugel, 2010). Some studies that connect student outcomes to text difficulty report these findings incidentally because they did not set out to examine text difficulty explicitly (Ehri, Dreyer, Flugman, & Gross, 2007; Stahl & Heubach, 2005). In the few cases where researchers did attempt to manipulate text difficulty, findings were mixed (Morgan, Wilcox, & Eldredge, 2000; O'Connor et al., 2002; O'Connor, Swanson, & Geraghty, 2010). In other words, while many have written on the subject, the relationship between text difficulty and student reading outcomes remains ambiguous.

Theoretical Framework

This mixed methods study rests on a pragmatic theoretical framework that draws from sociocultural views of reading and the work around reading comprehension put out by the RAND Reading Study Group (RAND Reading Study Group, 2002). The interest in students' engagement and discussion derives from a sociocultural understanding of reading as a fundamentally social process. The consideration of students' comprehension and fluency stems from a view of reading that considers individual reader abilities in processing text, something emphasized by the RAND group.

Sociocultural View of Reading

Wertsch (1998) explained that sociocultural analysis delves into the relationships between actions and their cultural, institutional, and historical contexts. Sociocultural theories emphasize that learning occurs as people participate in communities and become enculturated into the practices of those communities (Lave & Wenger, 1991). Less experienced people learn from the more experienced, and with their support begin to approximate the practices and skills of experts (Vygotsky, 1978). As Brickhouse (2001) explained, learning extends beyond acquiring knowledge: "it is a matter of deciding what kind of person you are and want to be and engaging in those activities that make one a part of the relevant communities" (p. 286). Numerous researchers have emphasized the importance of community and cultural connection in education: for learning to occur, people need to want to adopt the practices of a particular group in order to be considered a member of that group (Gutiérrez & Lee, 2009).

When students demonstrate engagement in reading, they position themselves as competent members of a literate community. Sociocultural researchers have defined reading as the "motivated, strategic, conceptual, social interaction with text and written language" (Guthrie & Anderson, 1999, p. 20) and emphasized that "acts of reading take place not in a void, but in a stream of cultural practices" (p. 24). Sociocultural approaches to engagement research allow researchers to "better understand what engagement-related … practices work, for whom, where, under what circumstances, when, why, and for how long" (Lawson & Lawson, 2013, p. 462). Bundick et al. (2014) agreed and suggested the importance of a sociocultural approach to describe "the complexities of classroom contexts" (p. 6) as well as "the interaction of students and their environments" (p. 7).

McCarthey et al. (1999) explained the centrality of the social context for reading researchers:

> Researchers interested in studying the nature of engaged reading and the instruction that supports it, must consider the social context as an integral resource system rather than as a confounding or extraneous variable in a traditional research design. (p. 47)

Sociocultural research differs sharply from traditions that seek to control or minimize the effects of context. Goncu and Gauvain (2012) even questioned the validity of experimental research that, they suggested, "strips children of their natural contexts" (p. 128). Indeed, it would be difficult to study engaged reading without adopting a theoretical lens that emphasizes the social context in which reading occurs.

In this study, the sociocultural lens drove the instruction with the students. Our reading groups revolved around student-driven discussion. Furthermore, the sociocultural orientation guided my interpretation of the data (in depth qualitative analysis of transcripts), and the way I foregrounded and prioritized data about student engagement and discussion when ultimately making recommendations for instructional practice.

The RAND Reading Study Group's View of Reading

Alexander (2012) suggested that reading competence is multidimensional and one theoretical framework alone cannot explain it. While she admitted the importance of context, she suggested that researchers cannot escape that reading happens in the mind and body, and thus does have a strong cognitive component. Perry (2012) noted several critiques of sociocultural frameworks in literacy. She suggested that they "are limited in their ability to explain what actually happens when an individual learns to read and

write—that is, when someone learns how to decode, encode, and otherwise make sense of written text" (p. 65). The RAND Reading Study Group put forth a framework that addresses some of these concerns.

The RAND model emphasizes the sociocultural context at all points, but it centralizes comprehension as a product of interaction between a reader, a text, and an activity (RAND Reading Study Group, 2002). Visual representations of the model often show the reader, text, and activity as overlapping circles (like a three-part Venn diagram) surrounded by a larger circle representing the sociocultural context. This group defined reading comprehension as "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (p. xiii). They intentionally referred to "extracting" and "constructing" in order to emphasize "both the importance and insufficiency of the text" (p. 11) in driving reading comprehension.

The first component of the RAND model refers to readers, and the report described in detail the ways individual readers differ. It elaborated sociocultural differences like the way schools position students, the type of instruction they make available to them, and home language and literacy practices. Differences between readers also occur when students have different vocabulary, oral language, knowledge of language structure, non-linguistic abilities (like working memory or attention), motivation, goals, discourse knowledge, background knowledge, metacognition, and strategies. Indeed, much reading research has shown the importance of these individual capacities for successful comprehension.

Various researchers have emphasized the role of students' integrating lexical and semantic knowledge to make sense of print (Perfetti & Stafura, 2014). In synthesizing

research on second language literacy, Lesaux and Geva (2006) highlighted the importance of vocabulary knowledge and syntactic skills for learning to read in English as a second language. In discussing teaching academic English to emerging bilingual students, Scarcella (2003) accentuated metalinguistic awareness and reading strategies.

In fact, focusing on the traditional cognitive domains of reading (phonemic awareness, phonics, fluency, vocabulary, and comprehension; see National Institute of Child Health and Human Development, 2000) does benefit young children's reading achievement (Sparks, Patton, & Murdoch, 2014), including those children learning to read in a second language (Ehri et al., 2007). And, comprehension skills like questioning and activating background knowledge (Taboada, Tonks, Wigfield, & Guthrie, 2009) and inferring (Oakhill & Cain, 2012) contribute uniquely to reading comprehension.

Yet, in keeping with its broad sociocultural orientation, the RAND group highlighted how these factors—which do make a real difference in reading outcomes—do not always reflect stable individual differences. Rather, many children who have not developed these areas strongly attend schools where they are "victims of inadequate instruction" (RAND Reading Study Group, 2002, p. 82) and may also come from homes that have not provided the kinds of early literacy experiences that U.S. schools tend to value and reward. The group also acknowledged that all of these abilities can change due to the act of reading itself. For example, in the process of reading, students expand their vocabularies and background knowledge, and that growth contributes to future reading success. Thus, the RAND model did not conceptualize differences in cognitive subskills like vocabulary as static differences that fatalistically determine reading achievement.

The RAND model also includes texts and activities as central elements with which readers interact on their way to building comprehension. Issues like vocabulary, linguistic structure, discourse style, and genre all contribute to the potential difficulty of a text. Historically, many studies of comprehension have paid scant attention to the texts students read, but text characteristics play a central role in the skills that students have to marshal to successfully comprehend (McNamara & Kendeou, 2011). And, how difficult a text is for a particular reader can hinge in large part on what the reader has to do with it—the activity.

The authors of the RAND report described students as "high-need" or "low-need" based on the amount of instructional support they require to comprehend successfully. However, importantly, they acknowledge that any readers can become high-need or low-need depending on the text and activity before them. They explained that a "high-need" reader can become "very successful in an instructional setting in which the teacher attends to this student's needs while selecting texts, designing tasks for him or her, and deciding how to structure the context to best support the student's participation and learning" (p. 30). In other words, texts and activities play just as critical a role as readers do in instructional contexts that lead to comprehension.

The RAND model allowed me in this study to centralize texts and consider how their difficulty affected students' reading. Because of the emphasis on individual reader skills in this model, I collected data about students' fluency and comprehension as measured through a retell. I agree with the RAND Study Group that these variables matter for developing higher levels of comprehension. Yet, I also share their critical

caveats that the field currently has limited ways to measure these variables, and both fluency and comprehension develop over time and in response to instruction.

Thus, in this study I embraced Alexander's (2012) claim that one theoretical framework cannot capture the multidimensional character of reading. I grounded my exploration and analysis of students' engagement and participation in a sociocultural view of literacy that emphasizes the social context of reading. Yet I also benefited from an expanded model of reading comprehension that foregrounded reader and text variables to help me understand what happens when students tackle difficult text.

CHAPTER 2

LITERATURE REVIEW

In this literature review, I situate this study in the context of developing second language reading comprehension for elementary students and instructional strategies that support this goal. I broadly discuss factors that contribute to reading comprehension, including skills that support and lead to comprehension as well as instructional practices that assist students in comprehension growth. This discussion of literacy development includes factors unique to emerging bilinguals who must comprehend written text in a second language. I then briefly discuss reading levels, both of students and texts. I also describe small-group reading instruction: its history, common programs, programs developed especially for bilingual students, and two instructional protocols that emphasize text level. For these two protocols, guided reading and close reading, I offer an overview, describe how they work for bilingual students, and describe cautions that researchers have suggested about using each approach. I also briefly survey research related to informational texts, the type of texts students read in my study. In a review theoretical issues related to text difficulty, I describe Krashen's comprehensible input hypothesis, Vygotsky's zone of proximal development, and the work of many reading researchers from a variety of theoretical frameworks who support either matched or difficult texts. Finally, I report the existing (but sparse) research base relating text difficulty to student outcomes—particularly discussion, engagement, comprehension, and fluency.

Factors Impacting Reading Comprehension for Bilingual Students

Many factors contribute to elementary students' reading comprehension. Literacy development is similar for students learning English and fluent speakers of English, but some important considerations apply uniquely or especially to emerging bilinguals. For both groups, research has identified predictors (both academic skills and instructional practices) that contribute to reading comprehension.

Similar Early Literacy Development between Emerging Bilinguals and Fluent English Speakers

The current research base suggests that similar factors predict reading achievement for students learning English and fluent English speakers and that these students benefit from similar, robust instructional practices.

Predictors. The National Reading Panel identified five key areas of literacy instruction: phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Institute of Child Health and Human Development, 2000). Phonemic awareness, phonics, fluency, and vocabulary all contribute to comprehension; researchers and educators generally consider comprehension the goal of reading and value the other areas for the ways in which they support comprehension, rather than for their own sake. Strong word level skills (phonemic awareness and phonics for decoding, as well as vocabulary) predict later reading achievement (Foorman et al., 2006; Perfetti & Stafura, 2014). Mastering these basic literacy skills quickly in the early grades sets students up for later success with reading comprehension (A. E. Cunningham & Stanovich, 1997; Sparks et al., 2014). Research has substantiated the predictive value of early phonics mastery and vocabulary breadth for later reading achievement both with English monolingual students

and emerging bilinguals (Yesil-Dagli, 2011). Thus, word level skills play an obvious role in predicting comprehension. When students have mastery of word level skills, they theoretically read with automaticity without expending mental energy on the effort required to decode words. Thus, they have more cognitive resources available to engage the text's meaning.

In addition to word level skills, oral language and vocabulary breadth powerfully predict reading comprehension. Numerous studies have validated the importance of oral language (Kendeou, van den Broek, White, & Lynch, 2009; Roth, Speece, & Cooper, 2002; Spira, Bracken, & Fischel, 2005). Students need the ability to comprehend and communicate spoken ideas to comprehend written ones. Strong oral language skills give students a sense of syntax and semantics that they need to make meaning of a text. Related to the need for strong oral language, breadth of vocabulary also plays a critical role in predicting comprehension (Elleman, Lindo, Morphy, & Compton, 2009; Stahl & Fairbanks, 1986). Even if students can decode successfully, they cannot comprehend the text if they do not know the meaning of the words!

Comprehension rests on several underlying cognitive skills. For example, students who comprehend well activate their background knowledge about the topic in the text (Elbro & Buch-Iversen, 2013; Mayer, 2012). Using background knowledge helps them make connections between the text and their lives, the world, and other texts, which supports comprehension. Strong readers also integrate information across the text to make inferences (Dunst, Williams, Trivette, Simkus, & Hamby, 2010; Florit, Roch, & Levorato, 2014; Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012; McNamara & Kendeou, 2011). Inferencing ability contributes uniquely to comprehension. Strong

readers apply knowledge about text structures (Oakhill & Cain, 2012), monitor their understanding as they read, take steps to correct breakdowns in comprehension (such as rereading a passage), identify important information in a text, and ask themselves questions as they read (Pearson, 1991). These cognitive habits of strong comprehenders support the process of making meaning from a text.

Sociocultural variables also affect comprehension. For example, students need to feel motivated to read a text and have an authentic purpose for reading it (Nolen, 2007). They need to view themselves as readers with strong self-efficacy to tackle texts and an identity as a literate person who enjoys reading (McCarthey, 2001). Researchers have criticized the National Reading Panel for omitting these critical factors and thus suggesting that areas like motivation make little difference for comprehension (Williams, Hedrick, & Tuschlnski, 2008).

Comprehension rests on the foundational predictors described in this section: word-level skills, oral language and broad vocabulary, cognitive processes that support sense-making, and motivating, positive-identity reinforcing contexts. The development of second language comprehension appears to rest on these foundational predictors as well (Amendum & Fitzgerald, 2010), suggesting some similarities in the process of comprehending for monolingual and emerging bilingual students.

Instructional practices. In addition to identifying underlying skills and cognitive practices that predict reading comprehension, research has also identified instructional practices that support comprehension. Again, strong literacy instruction for fluent English speakers also benefits emerging bilingual students (Amendum & Fitzgerald, 2010; Gersten & Baker, 2000). Key suggestions include providing direct instruction in

comprehension strategies and reading processes, incorporating classroom discussions, and providing high quality texts in a positive classroom climate.

Both emerging bilingual and fluent English-speaking students benefit from comprehensive programs that include the five areas of reading identified by the National Reading Panel (Amendum & Fitzgerald, 2010). Specific instruction in comprehension strategies following the gradual release of responsibility and teaching text structure helps students comprehend (Reutzel & Smith, 2004; Shanahan et al., 2010). Teachers can implement read alouds and think alouds to model the underlying cognitive processes described above, such as inferring or determining important information.

Numerous studies support the benefits of discussion for comprehension (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009). A quality discussion leads to deep thinking and lets students negotiate and co-construct meaning. Students often find discussions motivating, and discussions also support the development of oral language, which reinforces comprehension.

Lastly, teachers who provide numerous high-quality, high-interest texts in positive classroom climates support the comprehension growth of their students. Thematically-related texts build knowledge, enhance comprehension, motivate students, and support language development (Gelzheiser, Hallgren-Flynn, Connors, & Scanlon, 2014; Guthrie, Wigfield, & VonSecker, 2000). Children need access to expansive classroom libraries (Guthrie, Schafer, & Chun-Wei, 2001) and time to engage the books in authentic literacy activities (Duke, Purcell-Gates, Hall, & Tower, 2006), or activities that people actually do with texts beyond school contexts, to grow their comprehension.

Special Considerations for Bilingual Students

For all the similarities in underlying comprehension skills and beneficial instructional practices, teaching bilinguals and monolinguals reading is not the same. It is important to guard against the idea that teaching emerging bilinguals is "just good teaching" and that a teacher needs no specialist knowledge for the task. In fact, reading teachers of students learning English need knowledge of the roles of culture and language in learning and reading, in-depth understanding of the nature of language, appreciation for the value of the first language, and a clear vision of the importance of oral language to effectively work with emerging bilinguals (de Jong & Harper, 2005). They should also understand the role of background knowledge and the limitations that students learning English may face if their English oral language skills restrict the extent to which they can communicate what they already know about a topic (Bernhardt, 2009).

Emerging bilinguals benefit from thoughtful native language support in literacy instruction (Gersten & Baker, 2000). When they have a strong literacy foundation, including reading strategies instruction in their native language, they do better in English reading comprehension (Klingner, Artiles, & Barletta, 2006). Similarly, Mendéz, Crais, Castro, and Kainz (2015) found that bilingual vocabulary instruction produced growth in both languages. Comprehension skills do transfer across languages, and second language reading seems to be a function of second language proficiency (oral language) and first language reading (comprehension) (Lesaux & Geva, 2006).

While oral language predicts reading comprehension for all students (discussed previously), this relationship is especially critical for emerging bilingual students. Oral language encompasses issues like students' vocabulary knowledge, listening

comprehension, syntactic skills, and metalinguistic knowledge, all of which impact reading comprehension (Lesaux & Geva, 2006), and its importance as a predictor of comprehension cannot be overstated (Kieffer, 2012). Providing support in oral language development at school is critical for students learning English who may spend much of their out-of-school time in monolingual (not English-speaking) communities. Several researchers have identified the importance of robust vocabulary instruction for emerging bilinguals (Gersten & Baker, 2000; Klingner et al., 2006) since vocabulary predicts comprehension for them as it does for fluent English speakers (Silverman & Hines, 2009). Emerging bilinguals from some language groups can benefit from vocabulary instruction in cognates, which provide a valuable resource for comprehension (Proctor & Mo, 2009) especially of academic texts (Lubliner & Hiebert, 2011). Research has shown that on word-level tasks, students learning English tend to perform comparably to fluent English-speaking peers when they receive equivalent instruction. However, they often do not achieve equally on measures of comprehension, likely due to differences in English vocabulary and oral language (Lesaux & Geva, 2006). Thus, providing extra support in vocabulary and oral language is critical to leveling the playing field for emerging bilinguals and establishing opportunities to learn that will support comprehension.

Classrooms with many books plus additional home reading support may be especially beneficial for students learning English in terms of comprehension and motivation (Amendum & Fitzgerald, 2010). While all students benefit from home literacy connections (Morrow & Young, 1997), the additional support offers a needed advantage to emerging bilingual students who may not have English language literacy materials at home.

Finally, when students learning English qualify for additional reading support, it helps them to participate in intense and comprehensive programs. Programs that include multiple components (Snyder, Witmer, & Schmitt, 2016) and have an established scope and sequence (Amendum & Fitzgerald, 2010) provide a greater benefit than incidental teaching and support focusing on just one aspect of reading. For example, Snyder et al. (2016) recommended that even if a teacher is only concerned with improving one aspect of reading (such as comprehension) that the existing research base suggests that extra support should focus on multiple components of reading (such as comprehension and fluency) to create the greatest impact. Snyder et al. (2016) also found that intense interventions (such as those occurring three times per week or more) had greater effects for students than long, but less intense, interventions. Thus, they recommended intensity over length and suggested teachers consider increasing the intensity of extra support for a student if the current frequency does not produce greater achievement. Notwithstanding, Snyder et al. (2016) did note that reading comprehension interventions take longer to produce practically significant results than interventions that focus on phonemic awareness, phonics, fluency, or vocabulary.

Reading Levels

A variety of tools help teachers determine the reading levels of their students and of texts. Knowing the levels of students and texts helps teachers create matches between students and books so that students can avoid spending time with texts too easy or too difficult to help them grow as readers.

Assessing Students' Reading Levels

Informal reading inventories. Various published informal reading inventories are commonly used in schools (Mellard, McKnight, & Woods, 2009) to assist teachers in determining students' approximate reading levels, placing students in reading groups, and matching students to books (Nilsson, 2008). These reading inventories serve as formative assessments that help teachers identify a reasonable starting point for literacy instruction for each student, and they help teachers update their understanding of student accomplishments and needs throughout the school year (Afflerbach & Cho, 2008).

Informal reading inventories measure accuracy, fluency, and comprehension. Teachers administer them to students individually several times per year (often at the beginning, middle, and end of a school year). The assessment begins with a student reading a word list to help the teacher determine where to begin the inventory. Based on the student's results from the word list, the teacher selects a passage for the student to read. After the reading, the student answers comprehension questions. This process continues until the student can no longer produce a fluent reading and successfully answer the comprehension questions. The highest level at which the student can read the text with accuracy (usually 95-98%) and answer the comprehension questions (usually at 70-89% accuracy (Betts, 1946)) is the student's instructional level.

Popular informal reading inventories used widely throughout the United States include the Fountas and Pinnell Benchmark Assessment System (Pinnell & Fountas, 2010), the Developmental Reading Assessment (Beaver, 2012), and the Qualitative Reading Inventory (Leslie & Caldwell, 2016). The commercial market contains several

others, and while some variation exists (Nilsson, 2008), they largely follow the format described above.

Computerized testing. Some schools determine student reading levels with computerized assessment. For example, over 50,000 schools in the United States use the Accelerated Reader program, linked to the STAR Reading Assessment (Renaissance Learning, 2015). Students take the STAR Reading Assessment in a ten- to fifteen-minute session. During the session, the computer adapts the difficulty of the questions according to student performance. Students answer cloze questions by selecting the appropriate vocabulary word to complete a sentence. STAR Reading computes students' scores by comparing their performance with established lists of vocabulary recommended by grade. The software assigns students an instructional reading level at which they could comprehend 80% of the text. For example, a student assigned the reading level 2.6 should comprehend 80% of the text appropriate for second graders in the sixth month of second grade. These reading levels tie to the Accelerated Reader program, which levels books using the ATOS formula. The ATOS formula most commonly reports readability ratings in terms of grade levels, so the same scale measures both students and books. The ATOS formula considers word length, word grade level, sentence length, and book length in its calculations (Nelson, Perfetti, Liben, & Liben, 2012).

Students can also take computerized tests to determine their Lexile level. The Lexile Framework for Reading relies on a scale for both books and readers that rangers from below 200L for new readers to beyond 1700L for advanced readers (Nelson et al., 2012). It uses sentence length to measure syntactic complexity, and it calculates semantic complexity based on average word frequencies. Any standardized reading level

assessment can yield a Lexile score (Lennon & Burdick, 2004). Many schools use the Scholastic Reading Inventory in which students read passages of varying difficulties (adjusted by the computer according to their performance) and provide the missing word.

Other tools. Schools use other tools to yield information about student reading levels as well. Some districts create their own benchmark reading assessments or used published ones such as DIBELS (R. H. Good & Kaminski, 2002) or Discovery Education. These assessments do not yield an exact reading level, but give some information about students' reading performance in relation to local standards. Others use curriculum-based measures such as running records (Clay, 2001) or commercial measures that come with their reading curricula, such as the Rigby Reads assessment that links to leveled Rigby readers. Teachers in Reading Recovery use an observational survey (Clay, 1993) that includes multiple components (such as letter and word identification) related to early literacy, but identify students' reading levels through a running record as students read a continuous text.

Text Leveling

Assessing students' reading levels only provides useful information to teachers if they also have corresponding information about book levels. Knowing the level of a text allows a teacher to match students and texts of similar levels.

Mesmer et al. (2012) constructed and proposed a tentative theoretical framework for early grades text complexity. They built on the RAND model of reading that centers the reader, text, and activity within its sociocultural setting. They highlighted word, syntax, and discourse structures of individual text as well as drawing attention to the sequence, pace, content, and repetition of the text program. This model recognizes the

importance of the reader, activity, and context while centralizing the text and its quantitative features. The inability to consider qualitative features does limit readability formulas. Readability estimates provide a starting point for leveling text, and human judgment helps make a final decision about the match between a book and a reader.

Traditional readability formulas. The field of text leveling has "no clear 'gold standard' measure of text difficulty against which to compare" (Nelson et al., 2012, p. 17). For nearly a century, researchers have attempted to quantify text difficulty with various readability formulas (Lively & Pressey, 1923), and researchers developed over eighty formulas in the twentieth century (Pearson, 2000). These formulas gauge text difficulty through measures of semantic and syntactic difficulty. Semantic measures consider difficulty at the word level, such as length of word (as in Degrees of Reading Power (Koslin, Zeno, & Koslin, 1987)), number of syllables per word (Fry's readability formula (Fry, 1968)), or appropriateness of word for grade level (Spache's readability formula (Spache, 1953)). Syntactic measures use the sentence as the unit of analysis and commonly evaluate based on number of words per sentence. Some formulas, like SourceRater (Sheehan, Kostin, Futagi, & Flor, 2010), also consider discourse level features beyond the word and sentence level. Research has indicated that on the whole, these formulas reliably predict grade level equivalents in line with the recommendations and experience of expert teachers and assessment data (Nelson et al., 2012).

Several systems of text leveling use the same scale for measuring readers and texts. For example, the Lexile framework, Degrees of Reading Power, and the ATOS formula (used with the STAR assessment and Accelerated Reader) all measure texts and readers on the same scale, so the match is straightforward. The other formulas (above)

provide estimates of text difficulty that teachers can convert to grade level recommendations; these formulas do not measure students and books on the same scale.

The developers of Coh-Metrix attempted to overcome some of the limitations of previous readability formulas by incorporating extensive analysis of discourse level features (Graesser, McNamara, & Kulikowich, 2011). Coh-Metrix relies on five variables (out of over 200 variables considered in the development process) to predict text difficulty. The variables in the final model include: non-narrativity (not following story structure), referential cohesion (the extent to which ideas overlap across sentences), situation model cohesion (the extent to which language builds cohesion, for example with causal words), syntax (sentence simplicity), and word abstractness (the extent to which words do not refer to concrete objects). In contrast to other readability formulas, Coh-Metrix does not produce one estimate of text difficulty, but rather yields five scores, each reflecting a point along a continuum for each variable described above.

Classroom tools. Teachers rely on many tools to determine the level of books in their classrooms. These tools may include publisher provided levels, lists recommending books by grade (often available from public libraries), books recommended by standards documents or state education agencies, and their own experience and judgment. More formal classroom leveling tools include the Fountas and Pinnell text leveling system and Reading Recovery (which only applies to the earliest grades).

The Fountas and Pinnnell text leveling system includes 26 guided reading levels. Books receive their level based on an evaluation on the Fountas and Pinnell Text Level Gradient (Fountas & Pinnell, 1996). The gradient considers quantitative features (sentence complexity, vocabulary, number of words, number of high frequency words,

repeated words, and sentence length). However, it also includes qualitative dimensions requiring human judgment (genre, text structure, content, themes, literary features, illustrations, book and print features). Many trade book and curriculum publishers provide a Fountas and Pinnell reading level for their materials because so many schools use the system. However, teachers can also easily look up the level of a book online or with apps. An official website associated with Fountas and Pinnell lists over 55,500 books, and other websites (such as Scholastic Book Wizard) provide similar information.

Reading Recovery targets students in first grade who have not made expected progress after their year in kindergarten. Thus, Reading Recovery levels only cover early literacy. A published list of books (National Reading Recovery Project, 1995) places books at twenty levels for early readers based largely on the sight words they contain. Teachers level books not on the list according to their judgment (Peterson, 1991). Research has found that sentence length, word length, number of pages, number of words, and semantic properties of words reliably predict the assigned Reading Recovery level (Hatcher, 2000).

A need for individualized considerations in leveling. Clearly, text complexity varies for each student, text, and task. Fitzgerald et al. (2015) discussed the concept of "individualized text complexity" (p. 37), recognizing that complexity varies according to factors associated with the student (for example, interest and background knowledge), text (for example, genre and structure), and task (for example, pleasure reading or assessment). Many factors, both quantitative and qualitative, affect text complexity, and ultimately the judgement of complexity for classroom purposes often rests with the

teacher. After criticizing informal reading inventories for their inconsistencies in identifying student reading levels, Burns et al. (2015) admitted,

> It would be difficult to design an assessment system that does generalize well to reading levels of authentic books because successful reading is dependent on background knowledge, vocabulary from the given text, and text structures, which would be almost impossible to assess or take into account for an individual student. (p. 443)

This section has detailed issues pertaining to text difficulty, or quantitative measures of the readability of a text and parallel quantitative measures of students' abilities to read different levels of text. It has not explored the full range of issues that may make a text *complex*, as opposed to *difficult*.

Other factors affecting complexity. Multiple dimensions affect the complexity of a text for a particular student. Appendix A of the CCSS addresses text complexity and specifically highlights three elements of text complexity: qualitative factors, quantitative factors, and reader- and task-related factors. Qualitative factors include levels of meaning, text structure, language use, and knowledge demands. These categories include questions of whether the text has implicit or explicit meanings, the ways in which graphics interact with the text, whether the material is chronological, the amount of figurative language used, and the cultural or content background knowledge needed to understand. Quantitative factors include issues such as word length and frequency, text cohesion, sentence length, and overall passage length; computer software usually measures these factors and generates a readability score such as a Lexile. Reader and task

considerations include professional judgments that teachers must make about students' motivation and interest as well as what they will ask students to do with the text.

Appendix A of the CCSS suggests that "multiple quantitative measures be used whenever possible and that their results be confirmed or overruled by a qualitative analysis of the text" (p. 8). This suggestion makes sense given the problematic nature of readability formulas that often fail to capture linguistic (such as text structure, level of abstraction, or vocabulary difficulty), cognitive, and discourse-level features (Foorman, 2009). However, the CCSS does provide "text complexity grade bands and associated Lexile ranges" (Appendix A, p. 8), and the qualitative evaluations of text complexity require substantial linguistic knowledge and professional teaching experience. Thus, many teachers will rely on quantitative measures for initial judgments of text complexity, and most research has conceptualized text complexity by quantitative measures. This dissertation will do the same, but I will use the term "text difficulty" to acknowledge that by using readability scales for text selection I do not capture the full range of what makes a text complex.

Small-Group Reading Protocols

Over the past century, students who struggle with reading have received increasingly more support designed to enhance their reading achievement in school. In the first half of the 1900s, this extra support focused primarily on word level skills, but beginning in the 1970s, instructional support increasingly targeted reading comprehension (Scammacca et al., 2016). This trend continued through the 80s and 90s when researchers and teachers targeted comprehension through teaching metacognitive strategies like self-monitoring and summarizing. By the 1990s, most additional support

provided to striving readers targeted comprehension, and researchers expect this focus to remain prominent in the foreseeable future (Scammacca et al., 2016). These shifts in instructional support paralleled larger changes in general reading instruction and reading research. Over the past century, teachers shifted from viewing reading as pronouncing words to making meaning (Pearson, 2000).

Small-group instruction as a general practice and a way of supporting striving readers began to increase substantially in the 1980s to now for two reasons. First, the 1980s saw the rise of the whole language movement, which grounded children's literacy instruction in authentic literature, discussion, and writing (Pearson, 2000). The research and educational climate at this time gave more attention to qualitative research and the sociocultural context in which children learn to read (Alexander & Fox, 2013). Then, small-group support for students not reading at grade level increased dramatically with the passage of No Child Left Behind in 2000 and the implementation of the Response to Intervention framework in schools (D. Fuchs et al., 2014). With this change, schools and teachers faced increased accountability for the reading achievement of all students. Response to Intervention frameworks often involved grouping students who did not respond to the standard curriculum into small groups for more intensive reading support (Gersten et al., 2008).

Research has indicated that providing students extra comprehension support generally works (Edmonds et al., 2009). Murphy et al. (2009) reported a meta-analysis of literature discussion strategies designed to support comprehension. They found that these strategies generally resulted in substantial comprehension gains especially for students performing below average and students learning English. They concluded, "Evidence

suggests that discussions about and around text have the potential to increase student comprehension, metacognition, critical thinking, and reasoning, as well as students' ability to state and support arguments" (p. 743). In other words, small-group reading that provides students opportunities to discuss texts has great promise.

Survey of Small-Group Reading Instruction

A variety of instructional approaches to small-group text discussion have proven comprehension benefits for students (Murphy et al., 2009). Discussion formats that engage students with texts and invite open-ended discussion have multiple benefits. Santori (2011) noted the importance of classroom discussion (situations in which students have "textual agency" or engage their own views with texts) for both comprehension and developing a healthy view of what it means to read and comprehend: "…students need a school-sanctioned space where they have textual agency if we want them to view reading comprehension as an ongoing, fluid, social, cognitive, and linguistic process, rather than a discrete set of skills and strategies…" (p. 199). She explained that talking about text involves students in actively thinking about meaning, considering multiple ways to interpret, and taking into account their own experiences, all practices that improve their comprehension. While protocols differ, the opportunity to discuss texts in an authentic conversation as opposed to engaging in a teacher-directed question and answer routine (Cazden, 1988) helps students comprehend (VanDeWeghe, 2007).

Some small-group reading routines emphasize strategies instruction, and others call for open-ended conversation and higher-level thinking. In reciprocal teaching, students take on the role of teacher and generate their own questions. Reciprocal teaching emphasizes the comprehension strategies of predicting, asking questions, summarizing,

and clarifying (Palinscar & Brown, 1984). Transactional strategies instruction highlights interaction between a teacher and students in a reading group and the interaction between the group and the text they discuss (Pressley et al., 1992). In transactional strategies instruction, the teacher facilitates a discussion of the text by encouraging all the students to participate in the meaning-making process and to use strategies to generate alternative interpretations of the text.

Literature circles and book clubs are broad terms used interchangeably to refer to students selecting books or literature of interest and then coming together to discuss what they read (Raphael & McMahon, 1994). Questioning the author invites students to reflect on the author's meaning in a text. Students consider questions such as "What is the author trying to say?" or "What does the author mean by that?" (Beck, McKeown, Sandora, Kucan, & Worthy, 1996, p. 387). The developers of this discussion strategy aimed to promote deep understanding and interpretation. These instructional routines have all supported comprehension development (R. Brown, Pressley, Van Meter, & Schuder, 1996; Lysynchuk, Pressley, & Vye, 1990), including for emerging bilinguals (Carrison & Ernst-Slavit, 2005; Klingner & Vaughn, 1996; Martínez-Roldán & López-Robertson, 1999).

Text-Based Small-Group Reading Specifically for Students Learning English

Emerging bilingual students have long participated in small-group reading instruction with fluent English speakers. In some cases, teachers have adapted their practices for students learning English (Avalos, Plasencia, Chavez, & Rascón, 2007; DaSilva Iddings, Risko, & Rampulla, 2009) or offered the same small-group reading instruction as they do to fluent English speakers without making changes (Neufeld &

Fitzgerald, 2001). Thus, to some extent, the distinction between small-group reading and small-group reading for emerging bilinguals is artificial. Researchers have documented the effectiveness of several of the small-group practices described above for students learning English. And, some of the approaches described in this section have worked with fluent English speakers.

In collaborative reasoning, students read a text that presents an unresolved issue about which students may take multiple points of view (Zhang & Dougherty Stahl, 2011). Collaborative reasoning discussions, in contrast with traditional approaches, tend to elicit more student talk, fewer teacher questions, and more explicit references to the text (Anderson, Chinn, Waggoner, & Nguyen, 1998). Another approach, instructional conversations, are essentially "good classroom discussions" (Goldenberg, 1992, p. 318); they promote analysis, critical thinking, and reflection about text. Instructional conversations have proven effective for supporting bilingual students' reading comprehension and writing abilities (Saunders & Goldenberg, 1999). In another series of studies, students with native Hawaiian backgrounds and speaking local dialects participated in reading lessons as part of the Kamehameha Early Education Project (KEEP). These small-group lessons focused on comprehension of stories in the basal reader and incorporated quick interactions between the teacher and students. The interactions mirrored the native Hawaiian discourse pattern of "talk story" in which everyone jointly participates in co-narration (Au, 1980) with positive comprehension results. In proactive reading (Vaughn, Mathes, Linan-Thompson, & Francis, 2005), small groups of first-grade students read controlled texts containing familiar vocabulary and decodable words to develop fluency and comprehension. In the comprehension

strand of the program, students make predictions, relate their background knowledge to the text, retell what they have read, and answer questions about story elements and new learning.

These programs or instructional protocols do not represent an exhaustive list of text-based reading support for emerging bilingual students. However, these examples show a range of formats that benefits the comprehension of students learning English. These approaches expose students to texts of varying degrees of difficulty and authenticity. Crucially, in each approach, children respond to texts and discuss them in a small group. Collaborative reasoning, instructional conversations, and the KEEP project emphasize continual high levels of participation as students relate texts to their lives, challenge each other, and work out their understanding as a group. Proactive reading adds additional reading components beyond comprehension, a recommended approach (Snyder et al., 2016). It has less emphasis on discussion, but still does engage students in predicting, connecting background knowledge, retelling, and sharing new learning. Across the approaches, discussing texts emerges as a key commonality.

Most of the approaches described for emerging bilingual students and the general population do not place a strong emphasis on the readability of the text. To the extent that the approaches thrive on discussion, teachers commonly select texts complex enough to talk about, but few of these instructional programs give great attention to the readability of the text in relation to the students' reading levels. (Though, proactive reading does use vocabulary-controlled texts.) Guided reading, which relies on texts closely matched to students' levels, and close reading, which uses texts considerably above students' levels prove two exceptions to this general lack of concern for text level in small-group reading.

Small-Group Reading that Emphasizes Text Level

The Common Core State Standards recognize that "...students who struggle greatly to read texts within (or even below) their text complexity grade band must be given the support needed to enable them to read at grade-appropriate level of complexity" (Appendix A, p. 9). The question remains whether complex texts themselves or more traditional on-level texts provide the best means of accomplishing this goal. Below I describe two instructional protocols that aim to support students in comprehending text and making grade level progress in reading. However, they differ in the level of texts recommended to match with students.

Instructional level text: guided reading. In guided reading, teachers meet with small groups of students to provide reading instruction just at or above students' reading levels, or at the "instructional level" by Betts's (1946) categories. (See Table 1.) The widespread practice of guided reading in the United States typically involves around six students of similar reading levels meeting with the teacher about three times per week for twenty minutes each time (Ford & Opitz, 2008).

Table 1
Betts's Reading Level Descriptions

Reading level	Word recognition	Comprehension
Independent	99%+	90%+
Instructional	95-98%	75-89%
Frustration	90% and below	50% and below

Overview. Fountas and Pinnell (1996) described guided reading as "good first teaching for all children," rather than as instructional support exclusively targeting striving readers. Teachers implement guided reading with children grouped together based on similar developmental and reading ability. The flexible groups can change according to teacher observation and assessment of student needs. Text selection plays a paramount role in guided reading. The teacher selects a text within the instructional range of the students. Fountas and Pinnell (2012) explained that "teachers have learned to avoid the daily struggle with very difficult material that will not permit smooth, proficient processing—no matter how expert the teaching" (p. 270).

A guided reading lesson typically follows the format of: text introduction, reading, discussion, teaching points, word work, and extensions (Fountas & Pinnell, 1996). The teacher introduces the text, but does not provide such a thorough summary that the students no longer need to read it to understand. Students then read independently while the teacher listens in and offers support or prompting as needed. In the discussion and teaching points, the teacher focuses on comprehension strategies. Optional extensions at the end of the lesson involve writing or drawing in response to the text.

Ford and Opitz (2011) outlined several key principles of and assumptions undergirding the guided reading. They suggested that guided reading rests on the assumptions that all children can read; a skilled teacher is crucial; students need lots of reading practice; and instruction should emphasize metacognition, comprehension strategies, and higher-level thinking. They explained that guided reading is one tool in an effective literacy program (not the only tool), and that the goal of guided reading is independent reading for meaning that brings "joy and delight" (p. 235). Commercial

reading curricula frequently provide leveled books intended to support teachers in matching texts and readers (Ford & Opitz, 2008).

For emerging bilinguals. Some teachers have found students learning English hesitant to participate in guided reading and noted that these students watch others and follow their lead during small-group literacy instruction (Neufeld & Fitzgerald, 2001). To counteract this lack of participation, Purdy (2008) suggested supporting emerging bilinguals in guided reading through questioning, vocabulary instruction, encouraging collaborative talk, and cultural sensitivity. Indeed, guided reading may require modification to best meet the needs of students learning English. Some teachers have grouped emerging bilinguals together into stagnant low reading groups (DaSilva Iddings, 2005) that defy the spirit of flexible grouping described by the architects of guided reading (Fountas & Pinnell, 1996). Others have found that modifying their practices offers the support emerging bilinguals need.

DaSilva Iddings et al. (2009) provided a microethnographic account of a guided reading session in which a monolingual English teacher facilitated the reading comprehension of a group of Spanish-speaking emerging bilinguals. These authors described the role of the teacher as facilitating "purposeful and meaningful interactions within the social context of the text discussion, which helped the ELLs hypothesize about the story and enhanced their oral language development and use of text-specific vocabulary" (p. 60). In their analysis, they highlighted how the teacher reiterated student contributions, shared the instructional space with the students, activated and built on prior knowledge, and focused on central themes in the text. This approach to guided reading allowed students to serve as "co-participants in the interactions and, as such, [they were]

much more apt to experience a sense of agency as they co-constructed meaning from the text" (p. 59-60).

Avalos et al. (2007) suggested a modified guided reading format for emerging bilinguals. They advocated for spending three or more days with the same text and giving special attention to the cultural relevance of the texts. While they generally suggested the same guided reading format as Fountas and Pinnell (1996), they added a teacher read aloud for the purpose of modeling fluency and starting a discussion that fosters comprehension and vocabulary development. Thus, teachers can modify guided reading to support the literacy development of students learning English.

Cautions. Some have objected to the intense focus on leveling books, determining reading levels of children, and making a good match between the two, all areas central to guided reading (Fountas & Pinnell, 2005). Glasswell and Ford (2010) called for moderation:

> …in maintaining a focus on assigning numbers or letters to texts as labels that represent their 'difficulty,' we can lose sight of what matters in reader-text interactions. ...hindering teacher judgment and masking the transactions that occur between a reader, a text, and the social context in which they read. (p. 57)

These researchers feared an undermining of teacher judgment and a failure to consider qualitative and contextual factors when matching readers to books.

Fountas and Pinnell (2012) have responded to similar concerns by acknowledging the need to look beyond merely the quantitative level of a book and consider the demands the text makes of readers. They highlighted multiple features affecting text complexity: genre, structure, content, themes, literary features and language, sentence complexity,

vocabulary, illustrations, and book and print features (such as length, font, and supports like glossaries). They positioned the ability to analyze texts as "important teacher knowledge that takes time to develop" (p. 278) and emphasized that "the text gradient and leveled books are a teacher's tool, not a child's label, and should be de-emphasized in the classroom. Levels are for books, not children" (p. 281).

Difficult text: close reading. Close reading grew out of literary pedagogy (instead of literacy pedagogy) and has a long history in secondary and collegiate literature classes. However, its past (as a literary analysis tool) and its present (as a literacy instructional tool) breed disagreement about how to define and implement close reading in new contexts, like elementary classrooms and the small-group reading that happens there. This disconnect stems from "different theoretical and epistemological beliefs about reading, language, text, literacy, and schooling" (Fang, 2016, p. 108) in literary and literacy instruction. Most descriptions of close reading assume that the teacher selects a text that the students will find complex.

Overview. Brown and Kappes (2012) defined close reading as "investigation of a short piece of text, with multiple readings done over multiple instructional lessons. Through text-based questions and discussion, students are guided to deeply analyze and appreciate various aspects of the text" (p. 2). They suggested that this instructional practice offers students the opportunity to practice "logical arguments and critique the reasoning of others" (p. 2). More simply, Fisher and Frey (2014a) simply described close reading as "a purposeful rereading of a complex text." It may involve critical examination of text organization, vocabulary, key details, arguments, inferential meanings, author's

purpose, connections to other texts, and consolidating information from the text to form one's own opinion (D. Fisher & Frey, 2012).

Most models of close reading include an individual reading of the text, a group read aloud, text-based questions and discussion, and writing about the text (S. Brown & Kappes, 2012). After summarizing several models of close reading, Fang (2016) concluded that multiple readings, discussions, "careful attention to lexicogrammar as a creative meaning-making resource" (p. 109), and writing all played crucial roles in close reading. He summarized the goal of close reading as "developing engaged readers who are able to comprehend, compose, converse about, and evaluate complex texts in thoughtful, critical ways" (p. 109-110).

Fisher and Frey (2012) conducted a study to determine how to implement close reading most effectively in elementary schools. Elementary teachers observed high school teachers with skill and experience in close reading. In focus groups and through consensus, they distilled the relevant practices for elementary students and then applied their new procedure in their own classrooms. They identified several key elements of close reading:

- The teacher selects a short, complex passage worth spending time on.
- The teacher provides limited background information prior to the students independently reading the text.
- The students read and annotate the text.
- The teacher creates tasks and questions that require repeated readings and textual evidence.

For elementary school, the teachers suggested and piloted several modifications. At lower levels, teachers may begin by reading the text aloud rather than having students confront it independently first. Teachers may also rely on their expertise to determine how much "frontloading" to do. Background knowledge plays a crucial role in reading comprehension (Mayer, 2012), and failing to provide or activate it may not serve young students well (Snow & O'Connor, 2016).

Fisher and Frey (2013) described a combination of close reading and guided reading meant for use with elementary students. They suggested an initial independent read for students to become familiar with the text. During this read, students would annotate the text for patterns, connections, and confusions. The teacher would offer a think-aloud to support comprehension and then ask text-based questions to drive discussion and opportunities to reread.

Text-dependent questions to propel the discussion might include questions on the general topic (main idea, overall themes, point of view), key details, vocabulary and text structure, author's purpose, inferences, opinion, and intertextual connections (D. Fisher & Frey, 2012). Teachers can group their questions into literal (what does the text say?), inferential (what does the text mean?), structural (how does the text work?), and practical (what does the text inspire you to do?) (D. Fisher, Frey, Anderson, & Thayre, 2014).

Proponents suggest many benefits from close reading. It avoids extensive "before reading" practices like picture walks and K-W-L charts that critics suggest take too much instructional time and do not require students to actually read the text (Pearson, 2013). Teachers across the curriculum can use close reading to help students develop disciplinary literacies. For example, close reading of science texts can help students

explore the discipline-specific language and discourse of science (Lapp, Grant, Moss, & Johnson, 2013; Rosebery, Warren, & Conant, 1992).

Previous research has indicated that instructional support can help students read otherwise difficult texts at instructional levels of fluency (Burns, 2007). While close reading does not incorporate the same pre-teaching of specific words that Burns employed, it is possible that teachers who implement close reading with robust instructional support may use it to make otherwise difficult texts into instructional level texts for students.

For emerging bilinguals. In a large interview study (D. Fisher & Frey, 2014c), teachers expressed initial concerns about implementing close reading with students learning English and students with special educational needs, but when they tried it, they generally found that the instructional protocol provided enough support and that these students did benefit from the instruction.

Martin and Rose (2012) described a style of close reading that they termed "Reading to Learn" or R2L. In the preparing to read stage, the teacher read aloud and provided an understandable summary. In the detailed reading stage, students read, paraphrased, and discussed challenging but brief passages sentence by sentence. They followed up with writing. Schools that implemented the model experienced twice the expected rate of literacy development with the strongest benefit going to students learning English and indigenous students.

Functional language analysis is a type of close reading in which students analyze the text at a linguistic level (Fang & Schleppegrell, 2010). They explore the main topics by analyzing each clause and identifying relationships between words. They evaluate text

organization as they consider clause beginnings, combinations, and cohesion. And, they explore judgments and points of view through analyzing word choices. Students of teachers receiving professional development on functional language analysis, particularly emerging bilinguals, made greater gains on a standardized measure of disciplinary literacy and wrote more effectively than students whose teachers did not have the training (Fang & Schleppegrell, 2010). Wong Fillmore and Fillmore (2012) described a similar instructional routine in which students learning English spend ten to fifteen minutes per day unpacking one complex sentence. In their experience, this miniature close reading has produced gains for emerging bilinguals on state English Language Arts tests.

Cautions. Snow and O'Connor (2016) raised several concerns about close reading with complex texts. They anticipated student frustration, declining motivation, and exacerbated educational inequities associated with reduced opportunity to learn as students encounter texts they struggle to read. These authors objected to "widespread reliance" (p. 2) on close reading because they believed it would not be possible (or beneficial) to remove the important role of background knowledge from the reading process and that close reading procedures privilege textual information too much. They suggested that prior knowledge, moral judgment, logic, common sense, and social norms also provide valid sources from which to construct arguments, and so, in their view, limiting students to textual evidence artificially constrains the authenticity of arguments and claims they may make regarding texts.

Fisher and Frey (2014a) anticipated some of these concerns and suggested that thoughtful reading teachers would spend some time developing their students' reading stamina with texts they could read independently and other time developing their

students' reading strength with complex texts and close reads. They positioned close reading as one part of literacy instruction, but not the only strategy teachers would use. They suggested that interactive read alouds, shared readings, think alouds, guided reading with leveled texts, discussion, independent reading, and writing all have a place in the literacy curriculum. They concluded that "to abandon these practices in favor of close reading exclusively would be akin to having a toolbox with only one tool" (D. Fisher & Frey, 2012, p. 180). Even the Common Core State Standards, which have driven the push to more complex texts, contain the acknowledgement that students need to "experience the satisfaction and pleasure of easy, fluent reading" at times (Council of Chief State School Officers, 2010, Appendix A, p. 9).

Reading Informational Texts

Informational texts include works in which the "primary purpose is to convey information about the natural or social world" (Duke & Billman, 2009, p. 110). This designation excludes narrative non-fiction, such as biographies. Attention to informational texts with primary grade readers has grown in recent years both in response to research pointing out students' historic lack of access to these texts (Duke, 2000a) and the expectation that students demonstrate comprehension on both literary and informational text under the CCSS (Council of Chief State School Officers, 2010). Children typically enjoy reading informational texts (Mantzicopoulos & Patrick, 2010) and grow in their reading abilities when they read informational texts (Hiebert, 2005; Kamil, 2008).

Aspects of informational text can both increase and decrease the difficulty of this genre for young readers. For example, beginning readers may find unfamiliar text

structures and linguistic features (such as timeless verbs, generic nouns, singular forms standing in for a category, rare words, and abstract words) difficult (Duke & Billman, 2009; Nagy & Townsend, 2012). However, in contrast, the frequent repetition of key vocabulary, extra-textual support features (like glossaries and diagrams), navigational features (like headings and tables of content), and inclusion of definitions and explanations in text may make the content more accessible to young readers (Duke & Billman, 2009). How difficult students find informational text in large part rests on their previous background knowledge of the topic and their educational experiences orienting them to this type of reading.

Some research has suggested a need for genre-specific theories of reading comprehension processes. Paris and Paris (2007) reached this conclusion after teaching comprehension strategies to first graders with narrative texts and finding that the benefits did not transfer to informational text. These researchers suggested "...comprehending the different genres requires different genre knowledge as well as different types of cognitive processes, strategies, and prior experiences" (p. 32). Indeed, scholars of informational reading for primary students commonly suggest teaching informational text-specific strategies such as text structure, features of informational text, and main idea and supporting details (Dreher, 2000; Purcell-Gates, Duke, & Martineau, 2007; Shanahan et al., 2010).

Informational texts also offer the important benefit of knowledge. Students reading informational texts do not merely practice decoding and comprehension; they learn about the world! Researchers and practitioners have suggested wide reading of informational texts as crucial for building content knowledge (Gelzheiser et al., 2014),

and several have advocated for considering content knowledge a sixth pillar of reading (in addition to phonemic awareness, phonics, fluency, vocabulary, and comprehension) (Cervetti & Hiebert, 2015; Hirsch, 2006). This knowledge base (acquired through wide reading of informational texts) has particular benefits for students learning English who may have different content and cultural stores of knowledge than those typically valued and rewarded by the school (Bernhardt, 2009), particularly if they have experienced interrupted schooling or missed access to content instruction due to their participation in English language development programs.

Theoretical Issues

Several theoretical issues relate to text difficulty in small-group reading for students learning English. Krashen's (1981, 1985) theory of comprehensible input speaks to the difficulty of the text (the "input"), and Vygotsky's zone of proximal development (1978) speaks to the social supports provided in the small-group reading context. Because both theories address ways of supporting learners by providing either input (Krashen) or a learning challenge (Vygotksy) just a little beyond the learner's present developmental level, some educators have conflated the theories. However, the two theories derive from quite different assumptions (Dunn & Lantolf, 1998; Kinginger, 2001) and thus apply to different aspects of a small-group reading context. This section also includes discussion of the work of scholars who employ various theoretical frameworks to advocate either matched or difficult texts.

Comprehensible Input

Krashen (1985) suggested that "humans acquire language in only one way—by understanding messages, or by receiving 'comprehensible input'…that contains

structures at our next 'stage'—structures that are a bit beyond our current level of competence" (p. 2). Krashen (1985) explained comprehensible input as *i + 1*, where *i* is "the acquirer's current competence, the last rule acquired along the natural order," and *+ 1* is "the next rule the acquirer is 'due to' acquire or is eligible to acquire along the natural order" (p. 101). The theory of comprehensible input phrased this way assumes the natural order hypothesis, that people acquire language structures in a predictable grammatical sequence (Krashen, 1981). The comprehensible input hypothesis fundamentally addresses second language acquisition broadly, not specifically reading. While Krashen and subsequent theorists have failed to quantify *+ 1*, an application of the comprehensible input hypothesis to reading contexts for second language learners would suggest that the teacher provide reading materials slightly, but not greatly, above the learners' current levels. In a sense, studies of text difficulty attempt to quantify the ideal *+ 1* for students. Under this theory, the teacher takes the role of determining the students' *i* (current level) and providing reasonably challenging input (*+ 1*). The theory relies on the metaphor that the human mind, like a computer, processes input and acquires knowledge (Kinginger, 2001).

Krashen defined comprehensible input in terms of grammar. Later researchers working in English as a Second Language would define comprehensible input more broadly (Echevarria, Vogt, & Short, 2016), and Krashen would go on to propose wide reading as a way to accelerate students' English acquisition (Krashen & Williams, 2012). However, applying the original theory to small-group reading for elementary students does raise some suspicion because of the focus on acquiring grammar rules in order and the way that texts for young children do not really facilitate that goal.

Within the original comprehensible input framework, social interaction does not play a key role except to the extent that people use social interaction to obtain comprehensible input (for example, by asking a conversation partner to repeat or rephrase something). Research within this vein does not evaluate social interaction as a unit of analysis; rather it only considers social interaction in addressing "how social context functions as a support mechanism for processes that take place in the autonomous mind" (Kinginger, 2001, p. 422; see also Dunn & Lantolf, 1998). This information processing view of language and communication contrasts sharply with the theoretical milieu in which Vygotsky conceptualized the zone of proximal development.

Zone of Proximal Development (ZPD)

The zone of proximal development is a metaphorical social space in which children succeed at a task slightly beyond their developmental level because they receive support from a teacher or peer. Vygotsky (1978) described it as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). He suggested the instructional importance of the ZPD, stating "what is in the ZPD today will be the actual developmental level tomorrow—that is, what a child can do with assistance today she will be able to do by herself tomorrow" (p. 87). Vygotsky believed that what children could do with support provided more important information about their mental development than what they could do alone. Cazden (1981) neatly summarized the ZPD as "performance before competence" (p. 5), which highlights how in the ZPD a learner

receives support to successfully perform a task and then applies that support to later perform the task independently.

The ZPD does not independently exist across contexts. It varies for each learner, and in any given context the participants jointly construct ZPDs through interaction (Wells, 1999). Social interactions form the crux of the ZPD. The help children receive becomes part of their own thinking that guides their future development. So, the ZPD does not merely address the concept of development through assistance. Rather, a key idea behind the theory is that the assistance provided to the learner must drive maturing mental functions that push the learner into higher levels of development as they internalize assistance from others (Chaiklin, 2003). Thus, a focus on teaching in the ZPD defies instructional organization around discrete skills and subskills or assessment of students' independent practice after guided practice. Moll (1990) explained, "The focus…is not on transferring skills…but on the collaborative use of mediational means to create, obtain, and communicate meaning" (p. 13). He suggested the important outcome of teaching in the ZPD is "the ability of children to participate in qualitatively new collaborative activities" (p. 13).

The ZPD is not a theory of language learning or reading, but grows out of sociocultural theory which addresses human development in its social, cultural, and historical context. In sociocultural theory, social interaction is the primary focus of the research process. Educators and researchers who emphasize reading in social contexts have applied the ZPD to reading instruction. For example, Van den Branden (2000) offered fifth grade students reading in a second language several levels of text difficulty, but also provided students the opportunity to discuss the difficult texts. She found that

discussion allowed students to "look for the meaning of unfamiliar input collectively" (p. 437), leading to positive impacts on comprehension. She concluded:

> ...comprehension problems, especially those that arise as a result of a gap between the learner's current level of language proficiency and the proficiency needed to comprehend input with which the learner is confronted, inherently contain rich potential for further language development. It is exactly through bridging these gaps that learning may come about. (p. 438)

In this study, students bridged the gaps through discussion with each other and the researcher, and the researcher found that this social interaction allowed children to support each other to higher levels of comprehension.

Clay and Cazden (1990) provided a Vygotskian analysis of Reading Recovery. They referred to teachers selecting texts "on an increasing gradient of difficulty" (p. 219), (reminiscent of the discussion of comprehensible input), but they also discussed the social interaction between the teacher and student. In describing this interaction, they explained, "…the scaffold of teacher support continues, always at the cutting edge of the child's competencies, in his or her continually changing zone of proximal development" (p. 219). Of course, this language indicates the ambiguity of defining the zone of proximal development, a problem noted by other researchers (Wertsch & Rogoff, 1984) and suggests the need for highly trained and professional teachers (a cornerstone of Reading Recovery) able to identify and provide instruction within a shifting ZPD.

Finally, Miller (2003) analyzed open-forum literature discussions among secondary students to study how they facilitated learning in the ZPD. She found that in classrooms with teachers who supported the discussion,

> …a problem of understanding is jointly pursued, the context becomes a supportive social space in which mutual assistance creates new ways of talking and thinking about text—that is, such discussion creates a zone of proximal development…Over time the dialogic strategies moved inward to become part of students' repertoires for meaning-making. In varying ways each teacher mediated specific habits of mind by lending her "structuring consciousness" to enable students to think in increasingly complex ways about texts, knowledge, and the world. (p. 312)

Studies like this one emphasize the important role of discussion in comprehending texts, and they show how teachers and students can collaborate in discussion to create a ZPD for students to tackle texts more complex than what they might comprehend independently.

Thus, the ZPD provides a worthwhile lens for understanding how learners discuss texts of varying difficulty levels in small-group reading. Considering the interactions between the students and the teacher allows researchers to understand the type of supports or scaffolds (Wood, Bruner, & Ross, 1976) that students take up and offer each other as they collaborate to comprehend texts of different difficulty levels.

Voices in Support of Difficult Text

Some have suggested that offering students complex texts provides motivation. Explaining that "grappling with rich, complex texts is an exciting, thought-expanding experience" (S. Brown & Kappes, 2012, p. 5), these authors suggested that complex texts offer readers advanced concepts that lead to engaging discussions. They also described that close reading across the curriculum will prove "an effective strategy for deepening

content knowledge and learning to read like an expert in all academic disciplines" (p. 4). These authors suggested that in working with complex texts through close reading, students develop skills and new knowledge, both of which will motivate.

Some researchers have suggested that complex texts may emphasize challenge and avoid stagnation, and that perhaps the call for complex texts is a "message long overdue" to correct the previous "intense concern about avoiding frustration" (Mesmer, 2015, p. 84). Some have positioned access to complex texts as an equity issue, suggesting that historically not all students had the opportunity "to productively struggle with complex texts" (S. Brown & Kappes, 2012, p. 2), but that providing this opportunity "can be an important strategy to accelerate and deepen …learning [for students with lower reading skills]" (p. 2). Brown and Kappes (2012) adamantly insisted that "close reading cannot be reserved for students who are already strong readers; it should be a vehicle through which all students grapple with advanced concepts and participate in engaging discussions *regardless of their independent reading level*" (p. 2, emphasis added). Complex texts offer students greater access to models of complex language and content knowledge.

The revised publishers' guide to developing CCSS-aligned curricular materials (Coleman & Pimentel, 2012) argued for the importance of all students accessing complex texts. The authors explained,

> The CCSS hinge on students encountering appropriately complex texts at each grade level to develop the mature language skills and the conceptual knowledge they need for success in school and life. Instructional materials should also offer advanced texts to provide students at every grade with the opportunity to read

> texts beyond their current grade level to prepare them for the challenges of more complex text.
>
> All students (including those who are behind) have extensive opportunities to encounter grade-level complex text. Far too often, students who have fallen behind are only given less complex texts rather than the support they need to read texts at the appropriate level of complexity. Complex text is a rich repository of ideas, information, and experience which all readers should learn how to access, although some students will need more scaffolding to do so…
>
> …students whose reading is developing at a slower rate also will need supplementary opportunities to read text they can comprehend successfully without extensive supports. … Students who need additional assistance, however, must not miss out on essential practice and instruction …to help them read closely, think deeply about texts, participate in thoughtful discussions, and gain knowledge of both words and the world. (p. 3)

This lengthy explanation positioned access to complex text as an equity issue, suggesting that it does not benefit learners who struggle to not have access to the same grade-level text (and therefore content) as their peers. The guide both acknowledged that some students need extensive instructional support to comprehend complex texts and made clear that striving readers need ample time with accessible texts. Thus, it did not position complex texts or close reading as the only acceptable way to meet the standards or support literacy development. However, it did insist that close reading of complex texts does have a role in instruction for all students.

Some have cautioned that with existing classroom practices, teachers tend to do more of the cognitive heavy lifting as texts become difficult for students (Burkins & Croft, 2010). These authors suggested that teachers should orient students to the text in ways that teach students how to find support for comprehension from the text itself rather than relying on excessive teacher scaffolding. Close reading helps students understand what they miss with a superficial read, and it reinforces the message that students do have resources for constructing meaning independently (Snow & O'Connor, 2016). Glasswell and Ford (2010) explained that striving readers "can access grade-level appropriate material if we facilitate their interactions with it" (p. 57). They suggested instructional support that allows students to work with grade-level content "is both necessary if we want to accelerate growth, and desirable if we want our below-level readers to see themselves as competent and confident readers" (p. 57). Close reading may in fact provide just the instructional context in which teachers can support students in finding and using resources within texts to support their comprehension rather than relying on teacher scaffolding.

Wong Fillmore (2014) suggested that students learning English particularly stand to benefit from more complex texts. She suggested that

> not only can ELLs handle higher standards and expectations, but ...more complex materials are in fact precisely what they have needed, and lack of access to such materials is what has prevented them from attaining full proficiency in English to date. (p. 624)

She explained that the texts offered to emerging bilinguals are often simplified and offer little example of the academic language and discourse that students need to master to

fully participate in the curriculum. She described the ineffective practices that plague many programs for emerging bilinguals wherein "language required for advanced literacy and learning in school is treated as a prerequisite for working with complex and demanding curricula" (p. 627) rather than seen as a result of the working with with complex curricula. Wong Fillmore admitted that increasing complexity without adding appropriate instructional support would be "disastrous" (p. 626) and went on to describe an instructional routine like close reading at the sentence level that she believed may provide reasonable support for students learning English beginning to work with complex texts.

Researcher Voices Wary of Difficult Text

Yet, researchers have raised concerns that increasingly complex texts may not be appropriate for elementary students learning to read. Appendix A of the CCSS suggests that texts in K-12 classrooms have become easier in recent decades and that this decline in text difficulty level has contributed to falling reading achievement. The standards cite research that demonstrated a decline in high school text difficulty and established a gap between what students typically read at the end of high school and what they need to read two months later at the beginning of college. However, subsequent research that specifically evaluated elementary-level reading texts has found that the texts have either stabilized or become more difficult over the past century for beginning readers (Gamson, Lu, & Eckert, 2013; Stevens et al., 2015). These findings raise concerns about the evidence base used to advocate difficult texts for young readers. Pearson (2013) also raised concerns about the assumptions undergirding the new text complexity expectations. He reflected that the "assumption that we can get students back on the

college and career readiness track by gradually increasing the linguistic complexity of texts required of students in grades 2-12, is, of course, the unknown; it awaits empirical evaluation" (p. 250).

Other researchers have suggested that the new standards are aspirational and not based on developmental research. Noting that two-thirds of fourth graders could not read texts in the 580 to 620 Lexile band on the National Assessment of Education Progress, Hiebert and Mesmer (2013) suggested it was unlikely that third graders would be able to read 790 to 820, the new band under the CCSS. These researchers voiced concerns about falling fluency and engagement for students facing too difficult texts. They did not challenge the need for secondary students to read more complex texts, but pointed out that research has yet to establish any link between reading complex texts in the primary grades and college and career readiness at the end of high school. Further, they pointed out that the end of grade 3 represents less than one third of a student's educational career, but with the new Lexile bands, the standards demand that students reach 58% of their reading competency in this time. Hiebert and Mesmer questioned the wisdom and feasibility of placing the burden of increasing text complexity on the early grades.

Williamson et al. (2013) graphed alternative trajectories showing that intentional school districts could stay within the Lexile bands recommended by the CCSS, but decide at what point they want to shift their students to reading more complex texts. They illustrated the possibilities of moving the shift later to remove the burden of complex texts from young readers just beginning to decode. Their work identified multiple developmental trajectories and advocated for local decision making about the point at which students should begin to work with complex texts with any frequency.

Other researchers have echoed the concern about whether complex text provides an appropriate and inviting introduction to reading for primary students (Shanahan, 2011). Shanahan suggested that even if complex texts hold promise, they will not likely produce much benefit until teachers learn different ways of teaching with them. Historically when faced with texts students could not read, teachers read the texts to the students, supplemented with easier materials, or told and otherwise instructed students in the material they would have read had they been able to read the texts. Valencia et al. (2014) also voiced concern that the standards may not realize their promise if educators responded with overemphasis on text difficulty, but failed to consider the important role of the tasks they ask students to complete with those texts and failed to design tasks specifically intended to support higher levels of comprehension. Other educators have shared this concern about the potential of decentering the reader and the task when foregrounding quantitative measures of text difficulty (Wilhelm, 2015).

Use of complex texts with young readers undermines conventional wisdom about providing students instruction just above their current level. Allington (2009) has explained that for young readers to develop fluency, comprehension, and a large sight vocabulary, "a steady diet of high-accuracy reading is essential" (p. 49). After a survey of research on text complexity, Allington et al. (2015) recommended that teachers continue to use texts that students can read with 95% accuracy or higher until more research about the impact of varying levels of text becomes available.

Other researchers have voiced concern about the effect of raising text complexity for special populations of readers, such as those learning English or those with learning disabilities. Spear-Swerling (2015) expressed that particularly for students with specific

word-level reading difficulties, raising text difficulty may not offer students enough opportunity to practice "in texts that provide a reasonable match to…the skills they have learned" (p. 28). She worried that this mismatch would produce fluency difficulties and encourage readers to adopt what she termed "maladaptive reading habits" (p. 28) like not monitoring comprehension or just guessing at words, both practices that would not support long-term reading progress. She concluded that for students with specific learning disabilities that affect reading, grade-level learning is important, but should not supplant addressing the need to develop foundational reading skills for students. Bunch et al. (2014) explained that text complexity raises special issues for students learning to read in a second language, and they highlighted their concerns about potentially reduced engagement if students learning English receive texts they find too challenging.

Ambiguity about Ideal Levels

In addition to the paucity of research addressing the ideal text-reader match, researchers contribute to the ambiguity surrounding ideal text levels by their language. Pressley et al. (2001) described that good first grade teachers provide their students with texts they described as "just a little bit challenging" (p. 47). In a practice guide for kindergarten through third grade teachers, Shanahan et al. (2010) wrote that teachers should select texts "with word recognition and comprehension difficulty appropriate for the students' reading ability" (p. 1) without defining what might make a text "appropriate." Rasinski and Young (2014) suggested texts "towards the outer limits of a students' instructional level" (p. 3) for using assisted reading to promote fluency. Each of these recommendations leaves classroom teachers unclear on the actual implementation of reader-text matching due to their failure to specify exactly what these terms mean.

An Institute of Education Sciences panel convened to identify best practices for supporting reading comprehension in kindergarten through third grade wrote that they could "not recommend choosing texts that are too difficult for students to read or understand. Students should have opportunities to read somewhat challenging texts. Challenging texts may be most appropriate during activities where there is support available from the teacher" (Shanahan et al., 2010, p. 32). The panel described that defining challenging text fell to the teacher who should consider quantitative factors pertaining to text difficulty as well as qualitative factors like student interest and background knowledge that might make a difficult text more accessible. The guidance about incorporating challenging text when the teacher provides support suggests that the panel may support close reading although their report did not specifically address this practice. Notwithstanding, ambiguity remains around nebulous terms such as "too difficult" and "somewhat challenging."

Existing Research Relating Text Difficulty to Student Outcomes

Existing studies have found mixed results concerning text difficulty and reading achievement. Some have shown that students reading difficult text improved on some measure of reading achievement (D. Fisher & Frey, 2014b; Morgan et al., 2000). Others have shown minimal difference across difficulty levels, with a slight advantage for on-level texts in fluency outcomes (Hiebert, 2005; O'Connor et al., 2002, 2010) or a clear advantage for on-level texts with English learners in first grade (Ehri et al., 2007). In general, many of the existing studies are assessment studies with little to no instructional support provided for students reading texts of different difficulty levels; unsurprisingly, in these studies, researchers find increased difficulty without instructional support

typically reduces both fluency and comprehension (Amendum, Conradi, & Hiebert, 2017). No studies considered affective outcomes such as discussion and engagement, and few studies isolate effects for English learners.

Discussion

I could not locate any studies that related discussion outcomes for students working with various text difficulties. However, one study suggested that different reading instructional contexts bred different types of participation among students.

Santori (2011) did not specifically examine text difficulty, but she did conduct a qualitative analysis of the participation patterns that emerged in three different reading contexts (shared reading, guided reading, and shared evaluation pedagogy—a small-group reading discussion). She found that different instructional structures "fostered distinct comprehension practices by encouraging particular types of student participation" (p. 204). For example, in shared and guided reading the class conceptualized comprehension as getting the right answer or employing strategies, and student participation most commonly took the form of recalling, hypothesizing, discussing vocabulary, and summarizing. However, in the small group emphasizing discussion, students viewed comprehension as joint meaning-making and participated by hypothesizing, recalling, clarifying, and connecting.

Engagement

Guthrie and Anderson (1999) suggested that engagement should play a key role in conceptualizing reading. Researchers have described engaged readers as "motivated to read, strategic in their approaches to reading, knowledgeable in their construction of meaning from text, and socially interactive while reading" (Guthrie, Wigfield, & You,

2012, p. 601). I could not locate any studies in which the researchers set out to systematically examine student engagement as a function of text difficulty. However, several researchers included anecdotal findings and small-scale reports relating to student motivation, on-task behavior, task completion, and focus on the task (constructs distinct from, but related to engagement). Several of the results reported here occurred with secondary or college level participants and may or may not apply to young readers.

Gickling and Armstrong (1978) studied the relationship between task difficulty and on-task behavior, task comprehension, and task completion for first and second graders who normally struggled at school. While they used simple reading and writing skills activities from the basal reader (rather than connected text), they found that students consistently performed poorly in all three areas that they measured when given reading work at their frustration level. Students completed work at the independent level too quickly, causing a rise in task completion, but a decrease in on-task behavior. These researchers ultimately recommended that students work at their instructional level, which for the purposes of their study, they defined as tasks containing between 70 and 85 percent "known elements" (p. 34). Treptow et al. (2007) replicated this study with reading passages and comprehension assessment and confirmed the earlier findings. Their work suggested that the highest amount of on-task behavior occurred when students worked at their instructional level.

Morgan et al. (2000) studied second-grade "delayed readers" (p. 115). They randomly assigned fifty-one second graders to one of three partner reading conditions: on-level materials, texts two years above grade level, and texts four years above grade level. While they did not collect data to formally address engagement and behavior, they

noted in their report that students reading the most difficult texts (those four years above their grade level) seemed to struggle to stay on task and motivated. The authors speculated that this result related in part to denser text with fewer pictures.

Stillman and Anderson (2017) spent two years in a bilingual school observing teachers' implementation of the Common Core State Standards. In this school, teachers and administrators prioritized close reading with complex texts as early as first grade. The researchers described how they observed this focus lead to what they considered compromised practice. Students often gave the appearance of being busy, but closer examination showed that their work with annotations, graphic organizers, and peer discussions associated with close reading did not reflect much comprehension or engagement.

Research with older learners also supports the idea that matched texts support engagement. Several middle grades teachers reported in an action research project that implementing guided reading (with texts matched to the levels of the students) increased their students' motivation to read (Lyons & Thompson, 2012). In a small pilot study, six high school emerging bilinguals participated in a close reading procedure (using complex texts) for two units of study in their history class. The reports of these students indicated falling motivation to read after participating in close reading (Thomason, Brown, & Ward, 2016). Finally, research with college students has found that students experience more mind wandering when reading difficult texts as compared with easy texts, and that this distraction affects their comprehension (Feng, D'Mello, & Graesser, 2013). However, when readers expressed interest in the text, this interest moderated (by

reducing) the impact of the difficulty on both readers' comprehension and focus (Fulmer, D'Mello, Strain, & Graesser, 2015).

Others have suggested that students can find difficult text engaging. For example, Brown and Kappes (2012) claimed,

> Teachers who have implemented Close Reading in their classrooms are finding that being challenged by complex texts is not, as they feared, tripping students up; on the contrary, it is actually motivating students to work harder and think more deeply (p. 5).

Contextually however, this assertion appears to more reflect the anecdotal experience of the authors than widespread data collection.

Fisher and Frey (2014c) conducted interviews with fourth through twelfth grade teachers and focus groups with their students to explore their views about close reading after their school districts began widely implementing it. They found that teachers and students held generally positive views about the practice; they recognized that it provided a high degree of cognitive challenge, and students found it engaging, but tiring. Teachers expressed that they looked forward to working with students in the future who had several years of close reading experience.

Thus, by and large, research results are mixed and the existing research has failed to address directly the impact of text difficulty on student engagement, particularly for young readers.

Comprehension

Several studies have investigated comprehension outcomes in relation to text difficulty either directly or incidentally. Taken together, the studies have mixed findings and do not present a clear picture of an ideal reader-text matching plan.

In a study that compared young readers partner reading with texts at various levels of difficulty (Morgan et al., 2000), the researchers found students in all conditions (on-level material, two years above grade level, and four years above grade level) improved their reading level after a year of partner reading, but the greatest gains (and statistically significantly different from the other two groups) accrued to students in the two years above grade level group. The group reading materials four years above their current grade had greater mean gains than the group reading on-level materials, but the analysis did not indicate statistical significance for this difference. However, the authors speculated that reading more difficult materials benefited students more than on-level materials in dyad reading. They suggested more research to determine the point (grade level above current level) at which difficult material no longer benefits students.

In an evaluation of a comprehensive literacy program (rather than a specific experiment on text difficulty), Stahl and Heubach (2005) found that second grade students demonstrated remarkable reading growth (comprehension and fluency) under Fluency-Oriented Reading Instruction. In this instructional program, students made gains while reading texts much beyond their instructional level, often texts they read with approximately 85% accuracy. The authors suggested that this success with difficult text resulted from the extraordinary amount of scaffolding that students received. In the program, they experienced multiple exposures to the same text; had prior exposure to the

vocabulary; participated in home reading, echo reading, and partner reading; and listened to the texts read aloud. These authors concluded that greater teacher support warrants the use of more difficult texts. In the program, students also participated in independent and partner reading, and the texts they read during this time approximated their instructional and independent levels. Thus, the program had many facets and students read texts of varying degrees of difficulty in different segments of the program. While the results suggest some promise for difficult texts with instructional support, they do not provide the results of a systematic investigation that isolated text difficulty.

In another study, middle school students scoring at the fortieth percentile or below on their state reading achievement test participated in an after-school close reading intervention. In relation to a comparison after-school reading group, the close reading group showed significant improvement on the state reading achievement test and also developed positively in their self-perceptions as readers (D. Fisher & Frey, 2014b). It is unclear whether the benefit derived from the close reading or the additional free reading that the after-school program also involved. Furthermore, it is not clear if elementary students would respond in the same way.

One study compared striving readers randomly assigned to a tutoring condition with texts at their reading level or texts at their grade level (O'Connor et al., 2002). The study found no differences in comprehension between the groups although they noted that the tutors had more opportunity to teach vocabulary in the grade-level condition because these more difficult texts had more advanced words. A later study (O'Connor et al., 2010) compared students reading aloud from instructional and independent level texts (texts they could read with 92% or higher accuracy) with students reading from

frustration level texts (texts they read with 90% or lower accuracy). Again, the research team found no difference in comprehension or fluency after twenty weeks of treatment.

Mathes and Fuchs (1993) compared comprehension and fluency outcomes for fourth- to sixth-grade students in special education resource rooms. They randomly assigned some students to read texts at their instructional level and others at their independent level as measured according to the reading program used in the intervention. These researchers found no difference in outcomes according to the difficulty of texts students read. Importantly, they did not assign any students to difficult texts, so their results only speak to distinguishing between independent and instructional levels. Based on their findings, they suggested this distinction may be less critical, and they also raised the likely possibility that the published texts in their independent and instructional conditions were not so different.

In a small pilot study with six high school emerging bilinguals, researchers reported no comprehension growth or improved content knowledge for students participating in two history units built around close reading (Thomason et al., 2016). Another study found that bilingual high school students who had previously struggled with reading did not necessarily find a difficult text (as measured by conventional readability formulas) so different from an easy text if the difficult text was highly coherent and the easy text was not (Reed & Kershaw-Herrera, 2016). The authors suggested that instructional strategies to help students understand relationships (cohesion) among ideas in text could support effective comprehension for bilingual students faced with conventionally difficult texts. The study also highlighted the reality that multiple factors contribute to a text's difficulty and that students learning English may or may not

find a text difficult depending on which factors contribute to its difficulty on readability scales.

In an assessment study, researchers found that early primary readers' comprehension fell when they read texts that exceeded their grade levels (as defined by an informal reading inventory) even when they could successfully read the texts with 90 percent accuracy or higher (Amendum et al., 2016). Similarly, Treptow et al. (2007) found that students exhibited the highest comprehension when assigned passages at their independent level. Importantly, both studies provided no instructional support, so the researchers could not say if these results would hold had the students received robust scaffolding from their teachers or other peers to help them comprehend difficult passages.

Topping et al. (2008) analyzed a large dataset containing students Accelerated Reader and STAR Reading assessments. The dataset contained information on students' reading achievement, reading behavior, and book selection. From their correlational analysis, the authors found a small but significant relationship between selecting books of moderate challenge and reading achievement. They concluded that students could read books across a span of levels and still boost their reading achievement. They noted, "This range was quite broad...suggesting exact targeting of narrow challenge ranges is not critical for achievement gain" (p. 519).

Thus, the results reported here reflect mixed findings. It seems from studies that targeted young readers and provided robust instructional support that difficult texts may hold some promise although some studies reported minimal differences between groups working with different levels of texts.

Fluency

Fluency refers to students' smooth and automatic reading of text. The traditional components of fluency include accurate decoding, automatic word recognition, and appropriate prosody (stress, intonation, and phrasing) (Kuhn & Stahl, 2003). In a review of instructional practices that promote fluency, Kuhn and Stahl (2003) determined that the best text level relative to the students to use for fluency growth remained an open question. After detailing the results of several previous studies with mixed findings, they stated, "Our best guess is that more difficult materials would lead to greater gains in achievement, but more research is needed on this question" (p. 9). Menon and Hiebert (2011) reconsidered these same studies and located several additional ones, and they concluded that generally researchers found effects for fluency regardless of the text difficulty. They noted that generally in fluency research, researchers "pay scant attention to the features of the texts" (unpaged), but they still argued for fluency passages with which students "are fairly accurate" (unpaged) most especially for striving readers.

A study that compared students tutored in texts matched to their reading levels with students tutored in texts matched to their grade levels found that students who began with lower levels of fluency benefited from the reading level match (O'Connor et al., 2002). Students who already had higher levels of fluency benefited from reading either difficulty level of text.

Second grade students who participated in repeated reading of informational texts designed to have fewer multisyllabic, single-appearing, and rare words made greater fluency gains than a comparison group that read literary texts (with more multisyllabic, single-appearing, and rare words) (Hiebert, 2005). This study did not attempt to match

readers with texts by level, or even provide a quantitative measure of the text difficulty, but did provide substantial evidence that due to the words used, most second graders would likely find the literary texts more difficult. However, the study confounded text difficulty and genre. Notwithstanding, the author concluded that reading simpler texts produced greater benefit for students in terms of fluency.

A small-scale study with third grade students found that when students participated in repeated readings of texts that exceeded their instructional levels (defined for the study as 93-97% accuracy), they made slow fluency gains (Parker & Burns, 2014). After receiving instructional support to help students recognize the words in the texts (ie, make the texts instructional level for the students), the students' fluency rates began to improve. The authors concluded that "high practice without accuracy led to minimal improvements" (p. 88).

Another study also supported matching students with texts they can read accurately. In a study of students participating in Reading Recovery, an intervention program that serves first-grade students not making grade-level progress, researchers found that students made greater literacy gains when they could read the texts used in the intervention with 95% accuracy or higher (Rodgers, D'Agostino, Kelly, & Mikita, 2018). When Reading Recovery teachers used texts that students could only read at 90% accuracy or lower, these texts had a clear negative effect on students' progress.

Taken together, these few studies suggest that students benefit from fluency practice in texts they can read with a high degree of accuracy, but that often they benefit regardless because of the instruction offered to support fluency growth.

Conclusion

Most classroom teachers probably appreciate that they can and should use different levels of texts for different purposes and contexts. Those who advocate guided reading and those who advocate close reading do not suggest that these approaches should constitute the whole of any child's literacy instruction. Burkins and Yaris (2014; Saul & Dieckman, 2005) suggested a framework for reading contexts and text difficulty that involved advanced texts for read alouds, texts on or slightly above grade level for shared reading, instructional level texts for guided reading, and texts at independent level for independent reading. Under this framework, students tackle harder texts with more teacher support, and when students bear the brunt of the decoding efforts, the teacher supplies more accessible texts.

Allington et al. (2015) surveyed research on text complexity and considered outcomes such as time on task, vocabulary acquisition, and self-correction. They concluded that "the best research evidence currently available supports the use of texts that can be read with at least 95% accuracy" (p. 496). They did note that two studies (Morgan et al., 2000; Stahl & Heubach, 2005) in their survey did support the use of complex texts, but they discounted these findings because they did not compare complex texts with on-level texts and because they provided considerable instructional support through re-readings and assisted reading. The authors suggested that this level of support is not typical of classroom reading instruction, and so therefore teachers should not take these findings as encouragement to use difficult texts. However, considerable instructional support with re-readings and assisted readings are in fact key components of close reading and guided reading. Allington et al. (2015) concluded, "Perhaps if

classroom lessons were altered such that these levels of support were available every day for every reader, then it might be beneficial to use texts that can be read at accuracy levels below 95%" (p. 497). Since most teachers do provide such support in small-group reading, it bears exploring how students respond in small groups with different levels of text (Pearson, 2013). Such an investigation would test the wisdom of Stahl and Heubach's (2005) conclusion that more support provided by the teacher (as commonly done in small-group settings) lowers the need for student accuracy with the text. This thinking also aligns with the study's overall theoretical framework which positions reading as the outcome of interactions between the text, reader, and activity. When teachers consider the needs of readers and intentionally select texts and design activities, readers who in some contexts need a lot of support to comprehend become successful comprehenders (RAND Reading Study Group, 2002).

Proponents of accessible texts have written about the widespread use of complex texts in somewhat apocalyptic terms (Allington et al., 2015; Snow & O'Connor, 2016), suggesting that complex texts will doubtless spell the end of literacy progress and motivation for readers who struggle. However, proponents of complex texts do not write about accessible texts in the same way (D. Fisher, Frey, & Lapp, 2012), with even the architects of the standards themselves suggesting that texts students can independently access have a place in the curriculum (Coleman & Pimentel, 2012). Despite all the discussion of this issue, the research base lacks a comparison of small reading groups conducted with texts matched to students' levels and complex texts above students' levels. Certainly, both types of texts have a place in the elementary school day, but research has not explored the impact of text difficulty in the small-group setting on

students' reading engagement, participation, and achievement during these groups. This study aims to fill that gap.

CHAPTER 3

METHODS

Six third-grade bilingual students participated in a mixed methods study on the role of text difficulty in small-group reading instruction. Using qualitative case study methods, I evaluated the students' engagement (interaction, strategy use) and discussion (responses and participation) in each condition, and with quantitative single case design methods, I compared the students' reading comprehension and fluency with each type of text.

Context

Participants

A volunteer classroom teacher recommended students to participate in the study. We collaborated to select students with former or current designation as ELLs and who she considered reading below grade level according to her district benchmark assessments. I made the final two groups of three students based on which students and families agreed to participate in the study. I also considered students' reading levels so I could make homogenous groups.

The students spoke Spanish as a home language and participated in a dual language class at school. For half of their academic day, they received instruction in English, and for half they received instruction in Spanish. This program enrolls native English speakers and native Spanish speakers in the same class. (The small reading groups in this study focused on English literacy.)

The first group of students included Alyssa, Jack, and Rosa. The teacher described all three as "struggling." Jack received services as an ELL because he had

scored intermediate on the state's English assessment. While she did not have data on his Spanish proficiency, his teacher thought he was stronger in Spanish. Jack enjoyed drawing and reading and talking about animals. Rosa was quiet and compliant. Alyssa liked to talk and laugh. Her teacher described her as wanting to do her best and proactive in learning.

The second group read a bit below the level of the first group. These students included Sarah, Elise, and Gabriela. All three received services as ELLs. Elise and Gabriela scored basic, and Sarah scored intermediate on the state English test. In the small-group reading, all three girls enjoyed talking and reading and almost incessantly asking questions, but their teacher described them as "all my quiet kids" and "low." Near the end of the study, Elise received an IEP to address comprehension and fluency in reading. The school planned to pull her from the dual language program in the next grade to focus on academics in English. Her teacher described her as "even lower" in Spanish than in English. The classroom teacher thought Sarah had stronger Spanish than English and that Gabriela felt evenly comfortable in either language.

Small-Group Reading Format

Each group participated in small-group reading for thirty minutes for twelve sessions. The groups met daily except when the school's testing schedule or my travel interfered. Group 1 completed their twelve sessions in three weeks, and Group 2 took four weeks. The thirty-minute time allocation and three-member group size follows current recommendations for extra reading support (Gersten et al., 2007) and resembles the available support resources in many schools. We met in a supposedly quiet office

space in the school, but found students and staff frequently walked through the room on their way to the bathroom. The small-group reading followed this format:

1. Brief text introduction (~2 minutes)
2. Teacher read aloud (~5 minutes)
3. Independent reading (~5 minutes)
4. Text-based discussion (~12 minutes)
5. Comprehension and fluency data collection (~6 minutes)

This format combines the tradition of guided reading lessons (Fountas & Pinnell, 1996) and suggested practices for close reading of complex texts in elementary schools (D. Fisher & Frey, 2012) with the added modification of a teacher read-aloud first because the students were reading in a second language (Avalos et al., 2007). In this study, I compared student performance in different difficulty levels of text (reading level matched or one year ahead of students' current reading levels), not in different instructional conditions (guided reading or close reading). Thus, the instructional format combined elements from both approaches and remained constant regardless of the text. Table 2 provides more detail on each of the elements in the small-group reading sessions. I used informational trade books, reflecting the increased need for students to encounter informational texts in elementary school (Duke, 2000b) and the importance of providing emerging bilinguals content knowledge from wide reading (Bernhardt, 2009; Cervetti & Hiebert, 2015). The consistent use of informational trade books also controlled for any effect of text type on text complexity. Appendix A lists the books used in the study.

Table 2
Small-Group Reading Session Format

Element	Rationale	What it looked like
Brief text introduction	-activate background knowledge (Lesaux & Geva, 2006) -connect to previous learning (Chamot, 2005) -connect to students' lives (Dunst et al., 2010)	-Teacher quickly summarized text topic and structure. -Teacher previewed 2-3 essential vocabulary words.
Teacher read aloud	-model fluency (Braunger & Lewis, 2006)	-Teacher read text aloud. -Teacher briefly clarified vocabulary or concepts as needed. -Students followed along with text.
Independent reading	-provide independent practice (Allington, 2009; Burkins & Croft, 2010) -encourage use of comprehension strategies (Kelley & Clausen-Grace, 2006) -encourage use of decoding fix-up strategies (Morris, 2015)	-Students reread text independently. -Students used sticky notes to mark sections of text they want to discuss.
Text-based discussion	-allow group to co-construct meaning (Van den Branden, 2000; Wells, 1999) -encourage inferential thinking (Collins, 2016)	-Students shared sticky notes. -Teacher asked literal and inferential questions. -Teacher encouraged students to refer to text in discussion.
Comprehension and fluency data collection	-assessment	-One-minute fluency probe -One-minute retell probe

Research Team

I prepared and facilitated the small-group reading sessions, which followed detailed, but not scripted, lesson plans. A trained honors college student served as a research assistant who observed; assisted with data collection (monitoring recording

equipment, delivering fluency and retell probes); and assisted with interrater reliability on quantitative measures.

Text Matching Procedures

I administered the Fountas and Pinnell Benchmark Assessment (Pinnell & Fountas, 2010) to each student to obtain an estimate of their reading levels. I grouped students based on their scores. Because all students in a group did not score exactly at the same level, I created ranges (only spanning three levels) for the groups. When we read books matched to their instructional levels, Group 1 read books ranging from J-L on the Fountas and Pinnell text gradient system, and they read books from the M-O band as their difficult books. For Group 2, the matched books came from the G-I range, and the difficult books came from levels K-M. For both groups, the difficult books were one year ahead of the students' instructional reading levels.

Mixed Methods Design

This study followed a convergent parallel design (Creswell & Clark, 2011), with the simultaneous collection of qualitative and quantitative data. I analyzed each set of data separately and brought the findings from each strand of inquiry together for interpretation at the end of the study. Using qualitative case study methods (Merriam, 1998), I explored each student's engagement and discussion in the small-group reading sessions. I used quantitative single case design methods to collect data on student performance (comprehension and fluency) on the texts read that day in the small-group reading session.

Researchers use mixed methods for many different reasons. This study lent itself to mixed methods because it addresses multiple questions (Bryman, 2006) best answered

by different methods. The research questions about engagement and discussion related to students' experience of the small-group reading session: Were they interested in the text? Did they find ways to become involved in the discussion? How did they show this engagement and participation? Addressing these questions required a narrative description and qualitative analysis of the behaviors and discussion that occured in the group. The research questions about students' reading performance, here measured in comprehension and fluency, lent themselves to quantitative measures on traditional reading assessment probes. Mixing methods in this study permitted a more complete picture (Greene, 2007) of small-group reading experiences for striving readers: it allowed a window both into the students' lived experiences in the group and their performance on the academic task of reading.

Mixed methods research can produce controversy because of the apparent contradiction in the paradigms undergirding the work (Denzin & Lincoln, 2005). Qualitative researchers often approach their work with a more subjective view of reality and an awareness of their own biases as they construct meaning from data (Maxwell, 2012). This approach contrasts with the post-positivist scientific paradigm often driven by hypothesis testing and the collection of quantitative data. However, Hitchcock (2010) has suggested that adopting a pragmatic perspective helps avoid potential philosophical barriers often associated with combining qualitative and quantitative research (see also Creswell & Clark, 2011). A pragmatic approach to research assesses instructional approaches based on their success in the classroom. Thus, researchers working within a pragmatist lens combine their qualitative data with student scores to evaluate their instructional approach. Hitchcock suggested that quantitative work benefits from the

increased knowledge of the context and the perceptions of the stakeholders that qualitative research provides.

Researchers who work in schools know how to adapt their methods in pragmatic ways to meet the needs of these teachers and students (for example, Reinking & Bradley, 2008). In the current era of standards and accountability, teachers track quantitative data from their students and design instruction in response to it. However, most teachers also care about their students' lived experience of school: whether they are engaged and how they participate in the classroom community. Thus, a study that combines an analysis of students' experience (through engagement and discussion) with analysis of student reading performance (through comprehension and fluency) fits well with the current educational climate and the needs of teachers who want to engage their students while simultaneously challenging them to high levels of academic performance.

Hitchcock (2010) suggested several ways to blend qualitative research with single case design. He explained that qualitative work can provide a narrative description of context, address the social validity of interventions, identify unintended impacts, and explain factors hindering or facilitating implementation. He noted that in-depth description of the sample can help address the limited external validity of single case designs, which inherently have few participants. Detailed descriptions help readers determine "if the characteristics of a sample and context of a study make findings relevant to their needs" (p. 54). Additionally, single case designs attempt to show causal relationships, but they do not explain them. Qualitative work, according to Hitchcock, provides "a powerful approach for understanding localized causality" and can answer "culture- or context-specific questions regarding causality" (p. 49). The qualitative

component of a mixed methods case study moves the findings from a causal description (as provided by the single case design) to a causal explanation (as provided by the qualitative analysis). In the context of this study, it makes sense that students' experience of a small reading group (addressed with qualitative data) might shed light on their academic performance in that group (addressed with quantitative data).

Nastasi and Schensul (2005) explained how qualitative data collected in research about instructional protocols helps address the frequent gap between research and practice. They described the tendency of some quantitative researchers to only report numeric findings without documenting challenges, addressing culture and context, describing factors that facilitate or inhibit instruction, attending to social validity, describing what intended outcomes look like (beyond simply reporting standardized measures), and detailing the needs and resources of the target population. They suggested that by attending to these factors with qualitative data, researchers make it more likely that classroom teachers will adopt the instructional protocols that researchers recommend.

Several recent studies have combined qualitative work with single case designs. For example, McKeown et al. (2016) combined single case design, qualitative observations, and quantitative achievement data across groups to measure the impact of a professional development experience for teachers on their second-grade students' writing ability. Dennis, Sorrells, and Falcomata (2016) compared two math interventions on second graders' computation skills using a single case design. However, they also collected qualitative data about the social validity of their intervention. From this data, they learned what teachers attributed student success to, the extent to which teachers

believed students enjoyed the intervention, teachers' perceptions of students' confidence, and whether teachers would recommend the intervention to other teachers. Teachers also reported to the researchers that students applied the strategies from the intervention in the regular classroom and shared them with their peers, an important detail about transfer and social validity that the researchers would not have known without collecting qualitative data.

These two examples notwithstanding, the literature provides few examples of combining qualitative methods with single case design. I only located one example of a researcher who combined a qualitative case study with single case design. Nes Ferrara (2005) conducted a reading intervention with a striving reader in third grade. She collected fluency and comprehension data following the conventions of single case design, but she also wanted to explore "the nature of the lived experience of reading for a less-skilled reader" (p. 215). To this end, she collected qualitative data in daily field notes and informal interviews with the student. She later analyzed this data according to traditional qualitative methods by coding and theming. My study pursued similar goals of exploring the experience of the small group while simultaneously collecting reading performance data; however, in this study the qualitative data collection was more extensive and played a greater role in the final analysis than in Nes Ferrara's study.

Qualitative Data Collection and Analysis

Qualitative data collection included video recordings of the small-group reading sessions, my research journal, and recorded student interviews. I recorded and transcribed each small-group reading session and interview. I transcribed everything students audibly said, gestures such as pointing to the text or acting out a word, and behaviors such as

closing the book early or making a surprised face. These recordings, in conjunction with the research journals, made it possible to analyze engagement practices and discussion contributions that students used in each difficulty condition. I analyzed five of the six transcribed groups for each group and each difficulty condition (so 20 groups total). I discarded the first matched discussion and the first difficult discussion from each group because of problems with the recording equipment on two days, and I wanted to analyze an equal number of groups for each condition. Discarding the first of each type of discussion for each group had the advantage of allowing me to conduct qualitative analysis on the groups after the students had become comfortable with me and our procedures. (Not analyzing these four videos did not affect the quantitative data. I still included fluency and comprehension scores from students on days that I did not analyze videos.) Appendix B provides a summary of the data collected and analysis. Appendix C shows the final codebook.

Recorded student interviews provided some insight into students' perspectives. Each student participated in an exit interview after the last small-group reading session. Table 3 shows questions for the exit interviews.

Table 3
Exit Interviews with Students

1. What did you like about our reading groups? Why?
2. If your friend wanted to know what you do during reading groups, what would you tell them?
3. Was there anything you didn't like about our reading groups? Why?
4. Which books did you like the best? Why?
5. Were there any books you did not like? Why?
6. Which books did you think were easy? Why?
7. Which books did you think were hard? Why?
8. (Book ranking procedure: Student will be shown two books at a time (one reading-level matched and one above level). The interviewer will ask: which of these books did you like better? Why? This will be done six times so that the student answers about all 12 books used in the study.)

Qualitative analysis of the video recordings and research journals addressed the research question about engagement. For the purpose of this study, I defined engagement as *students' affective responses to text and social interaction and strategy use around texts in an effort to construct meaning* (see Guthrie & Anderson, 1999; Unrau & Quirk, 2014). Therefore, I began with a priori codes *strategy use* and *interaction for meaning*. I then developed inductive codes and subcodes based on the data through the constant comparative method (Corbin & Strauss, 2007; Saldaña, 2013). Before the study, I thought subcodes under strategy use might include *questioning, connecting, monitoring, summarizing, inferring,* and *using fix-up strategies.* Upon reviewing the data, I removed *questioning* and *inferring* because they overlapped completely with other codes. (*Questioning* coincided with the *ask* code under *interaction for meaning*, and *inferential/interpretive response* took the place of *inferring*.) I also removed *monitoring* because I found no instances of it in the data. I broke *connecting* into *connection to text, connection to personal experience,* and *background knowledge* based on the variety of

ways that students made connections. Finally, students adopted a few strategies I had not anticipated, so I added the codes *gesture, research, cognates,* and *genre.*

Before working with the students, I brainstormed potential inductive subcodes under *interaction for meaning*. I thought students might *ask, answer, clarify, agree, disagree, peer coach,* and *build on previous comments.* After reviewing the data, I kept each of these codes, but I divided *answer* into *answer peer/self* and *answer teacher*, and I added *introduce new* for when students moved the discussion along to a new topic.

Finally, after the study, I added another category related to engagement. I noticed that when students participated in discussions, they often did not talk about the texts (as in, the words on the page) as I had imagined when planning. Instead, many responses focused on the pictures, photos, and diagrams, and lots of other responses related to ideas introduced in the book, but not specifically to the texts. So, I added the broad code *focus of engagement* and sub-coded for *engagement with pictures through talk, engagement with text through talk,* and *engagement with ideas through talk.* I limited each category with the phrase "through talk" because I did not collect data to analyze other forms of engagement (such as gaze or internal thought processes).

This data analysis strategy mirrored the work of other researchers who have investigated engagement during small-group reading discussions (Almasi, McKeown, & Beck, 1996). These researchers mined video transcripts and field notes for examples of engagement. They analyzed the contexts and participation associated with each coded instance of apparent engagement using the constant comparative method to form tentative key assertions (Erickson, 1986). After reviewing the data for disconfirming evidence,

they coded the instances of apparent engagement to develop categories that explained in greater detail each of their key assertions.

I employed qualitative analysis of the videos and journals for research question 2 about discussion participation. For this study, discussion participation included students' verbal contributions to discussion. I originally planned to use two subcodes under *response: literal* and *interpretive*. I kept these codes, but renamed the last one as *interpretive/inferential*, and I added the code *incorrect* since some responses while falling into the category of *literal* or *interpretive/inferential* also made claims either contradicted or not supported by the texts. I also recorded descriptive statistics on the number of times each student participated with difficult books and matched books.

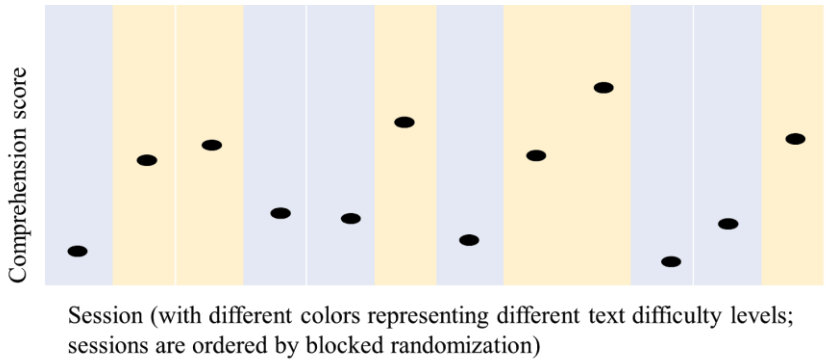
After I completed coding the transcripts, a senior researcher reviewed 25% of the dataset. This researcher confirmed that I consistently applied the codes given the parameters set in the codebook.

Quantitative Data Collection and Analysis: Single Case Design

In single case design (SCD), researchers compare at least two conditions, commonly referred to as A and B. The A phase is often a "baseline" condition when data is collected from the participant without changing their conditions, and in the B phase researchers introduce some treatment or intervention. They then compare the participant's behavior in both conditions, often through graphing. If the treatment made a difference for the participant, researchers expect to find a visually noticeable difference on a graph contrasting the A and B conditions. Single case designs can also compare two different treatments (A and B) as opposed to comparing one treatment with a baseline; researchers refer to this use of single case design as an alternating treatment design.

Figure 1 illustrates hypothetical findings from a SCD comparing reading comprehension under two different instructional protocols.

Figure 1. Hypothetical Single Case Findings



In this hypothetical illustration, the student understands better when the teacher uses the instructional protocol represented by the orange bars. The orange bars all show higher scores, and they contrast clearly with the blue bars that have lower scores. The goal of single case design studies is to demonstrate a relationship between an intervention and a response from the participant. In the case of comparing two treatments, single case design lets researchers compare the relative effectiveness of each.

This design allowed each student to serve as their own control and offered robust conclusions even from the small sample size. The What Works Clearinghouse recognizes alternating treatment designs as providing "rigorous experimental evaluation of treatment effects" (Kratochwill et al., 2010, p. 2). Leaders in the field have recommended them to study classroom literacy interventions (Neuman & McCormick, 1995). Alternating treatment designs allow researchers to compare the relative effectiveness of two different interventions in a short time frame (Horner & Odom, 2014) by exposing all participants to both treatments in a systematic or randomized pattern and comparing the data across

phases. They also offer the opportunity for researchers to provide in-depth description of responders and non-responders to the intervention, a possibility not afforded in group designs that present averages across groups. Alternating treatment designs work when the two treatments do not interfere with each other, and the behavior changes rapidly enough to observe differences between each treatment condition (Plavnick & Ferreri, 2013).

In single case design, each participant constitutes a case, and researchers analyze data at an individual level. One study can have multiple participants, all constituting their own case. In the present study, each of the six students represents a separate case. Having multiple cases within the overall SCD provides the benefit of built-in replication (Kratochwill et al., 2013). If all participants respond similarly across conditions, this finding lends credibility to any claims the researcher makes about causal relationships.

SCD history and role in literacy research. Single case design has a long history in medical research, and it also enjoys widespread use in school and clinical psychology and special education. The recent growth of the field of applied behavior analysis for people with autism has fueled the growth of single case design (Matson, Turygin, Beighley, & Matson, 2012). However, leaders in the field of literacy have recommended SCD in literacy research (Neuman & McCormick, 1995), specifically for identifying effective practices in the context of Response to Intervention (Vannest, Davis, & Parker, 2013). Yet, it seems that few literacy researchers have taken up this call, perhaps because SCD primarily measures performance at a specific time rather than achievement over the long-run (Kazdin, 2011). Indeed, SCD does not find effects that take time to accrue (Neuman & McCormick, 1995) such as long-term reading development. However, a single case design worked for this study because the outcome measures relate to the texts

that students read on a particular day in the small-group reading sessions. That is, the research questions addressed how emerging bilinguals perform on particular texts when provided robust instructional support. Rather than investigating long-term reading *achievement*, this study explored immediate reading *performance*. Students' reading performance in a given context is of interest because, for any proposed instructional protocol, it is important to know if students can experience success under those conditions.

Theoretically, students should spend large portions of their instructional time engaging in tasks at which they can succeed. Experiencing academic success has implications both for the amount of learning that students actually accomplish (C. Fisher et al., 2015) and for students' motivation and self-concept as learners (Ruddell & Unrau, 1994). Thus, if additional instructional support provided in small-group reading with either matched or difficult text does not result in immediate success in terms of student performance, teachers and interventionists should know (Snyder et al., 2016). Lack of successful reading at the time of the reading group most likely portends lack of reading achievement growth over the long-term as a result of the reading group. Conversely, successful performance in the reading group should forecast reading growth over the course of many similar reading groups. Wells (1999) suggested that teaching students within their zones of proximal development allows them to achieve immediate goals and also increases potential for future participation. Teachers should monitor students' immediate success on instructional tasks because "no one ever learned how to be good at anything (especially reading) by doing it poorly every day" (J. W. Cunningham, 2013, p.

138). I adopted single case design to accomplish the goal of exploring the immediate experience of two contrasting instructional conditions for emerging bilinguals.

While SCD is not an extremely prominent method in literacy research, some researchers have used it to contrast student reading performance in two conditions. Researchers have employed alternating treatment designs to compare ways of introducing letters to early readers (Johnston, Buchanan, & Davenport, 2009), summarizing strategies (Schisler, Joseph, Konrad, & Alber-Morgan, 2010; Taylor, Alber, & Walker, 2002), phonological interventions (Zens, Gillon, & Moran, 2009), different styles of feedback when readers miscue (Worsdell et al., 2005), fluency instructional protocols (Begeny, Daly, & Valleley, 2006; Hawkins, Hale, Sheeley, & Ling, 2011; Klubnik & Ardoin, 2010; Noltemeyer, Joseph, & Watson, 2014), methods of comprehension instruction (Güler & Özmen, 2010), and ways of practicing sight words (Didden, Prinsen, & Sigafoos, 2000).

Other single case design researchers have specifically explored literacy outcomes for emerging bilinguals. Bliss et al. (2006) measured the impact of an instructional protocol for supporting a fifth-grade student learning English in recognizing sight words. Gyovai (2009) studied the effect of a commercial reading program on twelve kindergarten and first-grade emergent bilinguals' phonological awareness, fluency, and decoding. Gilbertson and Bluck (2006) used an alternating treatment design to compare the effects of two different instructional paces on four kindergarten emerging bilinguals' letter naming fluency. Thus, SCDs, including alternating treatment designs, can provide useful information about in-the-moment literacy performance in relation to instructional protocols.

The present study. Using an alternating treatment design, all groups participated in reading groups with texts matched to their reading levels and with texts a year ahead of their current instructional levels on the Fountas and Pinnell text gradient system. Since the dependent variables in this study relate specifically to the texts students read that day in the small group (rather than overall reading achievement), multi-treatment interference is not an issue. This study examined student performance, rather than achievement, under two different conditions. How students participate under one condition should not affect how they participate under a different condition. For example, a student's fluency with an on-level text should not affect their fluency with a more difficult text read on a different day. A blocked randomization procedure mitigated any potential order effects and ensured that all groups participated in each condition a total of 6 times. Group 1 read their texts in the randomly assigned order ABABABABABAB, with A representing a day with matched texts and B representing a day with difficult texts. Group 2's random assignment was ABBAABBAABBA. Randomization reduces Type I error (Ferron, Foster-Johnson, & Kromrey, 2003) and improves the scientific robustness of SCDs (Kratochwill & Levin, 2014).

The study met the What Works Clearinghouse guidelines (Kratochwill et al., 2010) for single case design research by:

- Systematically manipulating an independent variable (text difficulty)
- Measuring each dependent variable consistently over time and establishing interrater reliability
- Including five repetitions of the alternating treatments

- Collecting one data point per phase (acceptable for randomized alternating treatment designs) with a total of six data points across all A phases (matched text) and six data points across all B phases (difficult texts)

Data collection and measures. At the end of each reading group, students read for one minute from the book used in that day's group. This recording yielded a fluency score (words correct per minute) (Hasbrouck & Tindal, 2006). Students also provided a retell of that day's text as a comprehension measure (Francis, Fletcher, Catts, & Tomblin, 2005). The research assistant and I established interrater reliability on fluency and comprehension scores. Data collection took two minutes per student (one minute for fluency and one minute for the retell), and so it required six minutes of the reading group each time. While one student completed the quick assessments, the other students wrote and drew in blank books. Students completed the quick assessments across the room from the other group members, and I assessed them in the same order each day. The data collection measured students' reading performance on the books they read in the group; I did not collect global measures of reading achievement. This method allowed conclusions about students' level of success with the specific academic tasks in each difficulty condition.

Students demonstrated comprehension through a retell probe. At the conclusion of each small-group discussion, a member of the research team asked each student, "Please tell me what this book was about like you are explaining it to a friend who has not read it." If a student paused for 5 seconds, the researcher asked, "What else do you remember?" The researcher employed this prompt up to two times. If the student paused

for more than 5 seconds a third time, the researcher ended the retell. The researcher also ended the retell once the student had spoken for one minute.

The research assistant and I independently scored 25% of the recorded one-minute reads and found that we agreed on words read correct per minute within one word. I scored the remaining fluency probes. To establish interrater reliability on correct ideas retold, we independently scored all of the transcribed retells. We considered a statement a correct idea retold if it had a subject and a verb, retold an idea or reasonable inference from the book, and did not repeat an idea the student had already stated in the retell. We originally scored all but one (71 of 72) retell within one idea unit of each other. We met to consider our scores together. We reviewed our criteria for determining an idea unit, independently rescored each retell on which we differed, and discussed any remaining differences. After this process, we came to complete agreement on scores for all the retells.

Retells are a common, but not unproblematic, measure to assess reading comprehension in research (Reed & Vaughn, 2012). They do correlate with overall comprehension (Marcotte & Hintze, 2009) and draw on the important skill of summarizing information. However, they also confound retelling, recalling, identifying main idea, and summarizing and can produce difficulty for second language learners who may comprehend much more than they productively retell in one minute. I hoped to somewhat alleviate these limitations for the bilingual students in this study because by the time they retold the text, they had heard it, read it independently, and engaged in a small-group discussion of it. A timed retell did not afford them a full range of resources for demonstrating their comprehension (like additional wait time and use of their home

language), but it did reflect a common expectation they faced at school to demonstrate comprehension in English reading. Comparing their retells across two conditions provided important data about how they comprehended and retold matched text compared to difficult text despite the limitations of the retell measure.

A retell met the needs of comprehension assessment in this study because students could complete them quickly, and they applied consistently across the twelve different texts students read. In addition, retells rely less on decoding skill than other commonly used standardized measures of comprehension (Keenan, Betjemann, & Olson, 2008).

Given the sociocultural view of literacy that in part undergirds this study, it is important to remember that the qualitative analysis of students' engagement (which includes their use of strategies to make meaning from text) and participation (which includes their literal and interpretive responses in discussion) shed additional light on how students comprehended in each condition. The retell provided important quantitative data that facilitated comparisons between the two conditions; such a measure is common in research that takes a cognitive view of reading. However, in this study, other lenses and data points also addressed reading comprehension.

Many have criticized the common practice of timed readings to measure fluency. Critics have suggested that this assessment encourages students to value speed over comprehension and that it undermines important elements of fluency such as expression that the words read correct per minute score does not capture. However, the timed reading remains a fairly stable and even recommended (Gersten et al., 2008) assessment practice in schools, and researchers have found that it correlates well with students' comprehension (National Institute of Child Health and Human Development, 2000; Snow

& Burns, 1998). Indeed, some research even suggests that young readers may use prosody, an element of reading fluency, as a tool for understanding difficult texts (Benjamin & Schwanenflugel, 2010).

Data analysis and interpretation. I used quantitative methods specific to single case design to analyze students' retell and fluency scores across conditions. Single case research relies primarily on visual analysis of graphed data. In accordance with What Works Clearinghouse standards (Kratochwill et al., 2010), I graphed quantitative data for each student across both difficulty conditions and determined the consistency of data in each phase (by evaluating level, trend, and variability); determined if the data showed a change between conditions (by evaluating the immediacy of the change when conditions change, the degree of overlap across conditions, and the consistency of data in similar phases); and examined any anomalies. To meet WWC standards for demonstrating an effect, the data must yield three demonstrations of an effect, and "an effect is demonstrated if manipulation of the independent variable is associated with predicted change in the pattern of the dependent variable" (Kratochwill et al., 2010, p. 18). In addition to visual analysis, I calculated the averages for students in each condition and the percent non-overlapping data (Horner & Odom, 2014; Scruggs, Mastropieri, & Casto, 1987). While the WWC acknowledged "there are no agreed-upon methods or standards for effect size estimation" (p. 22), they included both of these methods as acceptable ways of quantifying findings from single case designs.

CHAPTER 4

FINDINGS

Six children participated in this study, and four research questions drove my inquiry. Here, I report the quantitative findings by student because each student responded a bit differently. Single case design studies allow researchers to paint a detailed portrait of each participant. They prevent losing the peculiarities that make students distinct. After presenting each student, I provide a traditional summary of the quantitative findings and then present the qualitative findings for each group. Throughout this section, for the sake of clarity, I refer to the books students read by their titles (especially as several had the same author). Full citations are in Appendix A. The research questions were:

In small-group reading sessions for bilingual third graders reading below grade level,

1. How does text difficulty impact student engagement?
    a. How do students engage with the books?
    b. What comprehension strategies do students use in discussion?
    c. In what ways do students interact to participate in discussions?
2. How does text difficulty impact small-group discussion?
    a. What types of responses (literal, inferential/interpretive, and incorrect) do students share?
    b. How often do students participate in each condition?
3. What is the effect of text difficulty on reading comprehension?
4. What is the effect of text difficulty on reading fluency?

I address these questions out of order to begin with a close-up of each student's response (questions 4 and 3) before moving to group results about engagement and discussion (questions 1 and 2). As a preview, in some areas, text difficulty made a pronounced impact, and in other areas it did not have any effect at all. The students each responded differently, so sometimes as soon as one group demonstrated a pattern, the other group showed the opposite or one student stood out as an exception.

Quantitative Findings

Alyssa

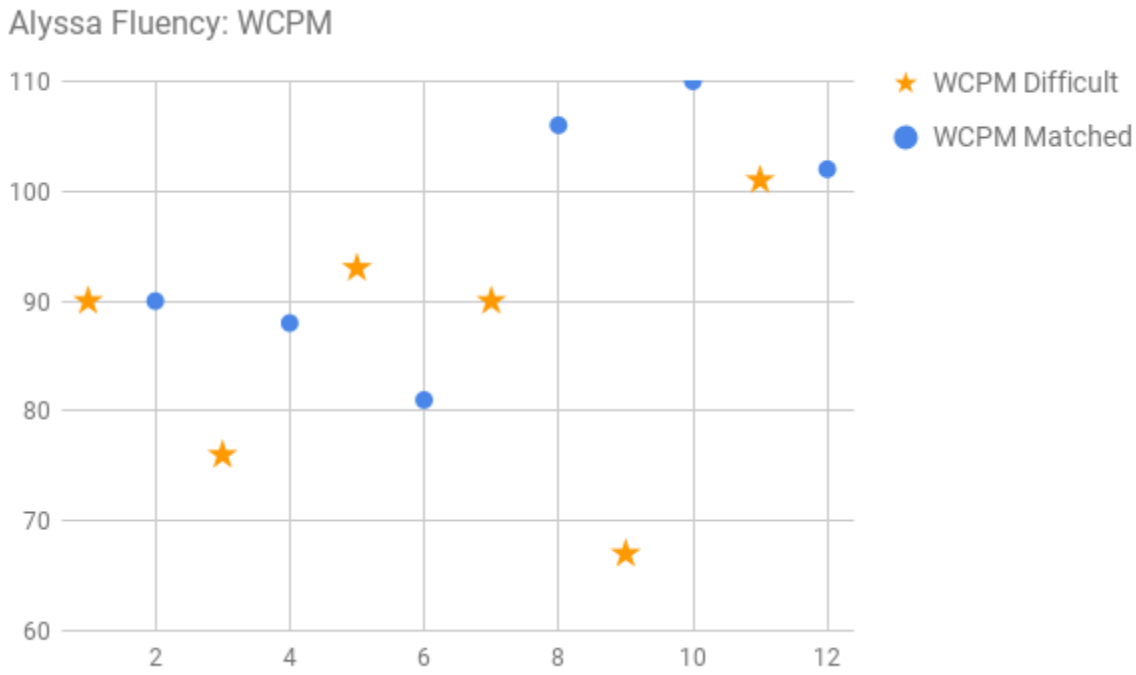
Figure 2. Alyssa with her Favorite Book

Alyssa reported in her exit interview that she liked that "we got to read books." She described the groups as reading and discussing. She identified one book as hard, none as easy, and overall thought the books were in the "middle" as far as difficulty.

When I asked her to pick her favorite book from pairs of two where I combined matched and difficult books (without ever identifying the books as matched or difficult), she picked the difficult books in four of six cases. Figure 2 shows Alyssa with her favorite book, which she liked because a small blurb in the end matter told an interesting story about how the bush baby got its name.

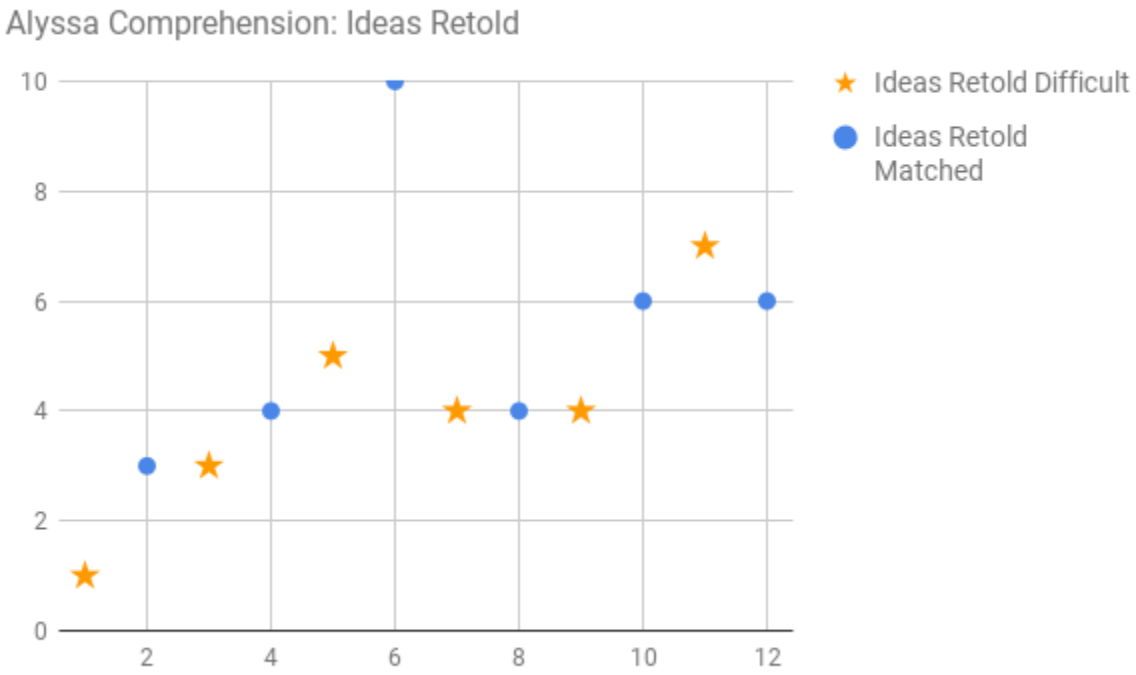
Fluency. As Figure 3 shows, Alyssa varied in how fluently she read. The two books she read the least fluently were difficult, and the three books she read the most fluently were matched. Alyssa read difficult books at an average of 86 words correct per minute (SD=12 words). Within these books she ranged from 67 to 101 words correct per minute. She read matched books at an average of 96 words correct per minute (SD=11 words). Here she ranged from 81 to 110 words correct per minute. (In the spring of third grade, readers at the 50th percentile read 107 words correct per minute (Hasbrouck & Tindal, 2006).) Alyssa's data for difficult books shows almost no trend (slope of the line of best fit, m = .357), and for matched books she read slightly more fluently as the twelve days progressed (m=2.157). Alyssa's graph does not show consistent performance for either level of book or an immediate clear effect when she changed difficulty levels. Only 42% of the data did not overlap with data in the other condition.

Figure 3. Alyssa's Fluency Scores



Comprehension. Text difficulty had minimal impact on Alyssa's comprehension in terms of ideas retold (Figure 4). She retold the most ideas from a matched book and the fewest from a difficult book, but her ten other data points overlap with each other and show no relationship to difficulty. She retold an average of 4 ideas from difficult books (SD=2 ideas, ranging from 1 to 7), and 5.5 ideas from matched books (SD=2.5, ranging from 3 to 10), and only 17% of her data points did not overlap with the other condition. Alyssa showed no substantial trends over time (m=.457 for difficult books and .214 for matched books), and her graph does not reflect consistency within each level of book or a clear effect from text difficulty.

Figure 4. Alyssa's Comprehension Scores



Alyssa's fluency and comprehension graphs do not tell the same story as each other. I did not set out to examine the relationship between fluency and comprehension, but for five of the six students, their graphs made it immediately apparent that the relationship was not as straightforward as the literature suggests it typically is. Fluent reading did not always enable comprehension, nor did disfluent reading impede it for these students. Thus, I report the relationship between fluency and comprehension for each student alongside comprehension data.

Alyssa retold the most ideas for a book that she read with lower fluency, and for the book she read the least fluently, she recalled an average number of ideas. For Alyssa, comprehension and fluency had almost no relationship (r=.15). Taken together, the two

graphs suggest that text difficulty did not make an important difference in Alyssa's reading performance.

Rosa

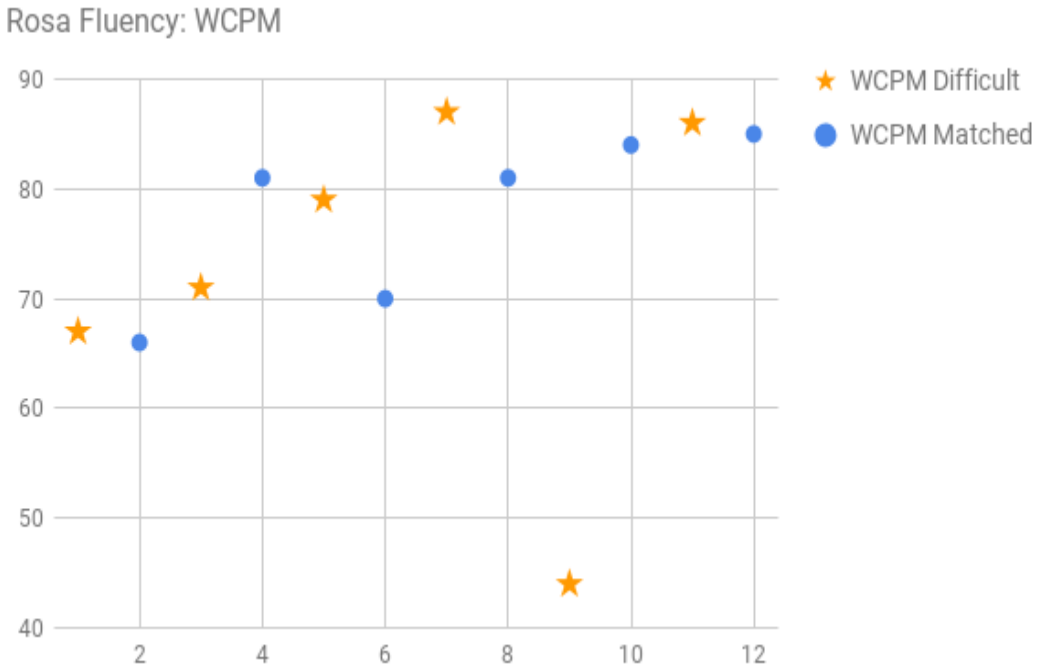
Figure 5. Rosa with her Favorite Book

Rosa described the reading groups as "doing different books." She liked guessing books (like *Looking Closely*), but she did not care for *Fossils* "because it tells you about fossils." When I asked her to pick her favorite book from pairs of two, she picked the matched books over difficult ones in four of six cases. Rosa also liked data collection: both getting to draw while other students stepped aside for their retells and fluency, and getting to do a retell herself. Figure 5 shows Rosa with her favorite book, *Pop!*, which she liked because it tells how to make bubble solution at home.

Fluency. Text difficulty had almost no discernible impact on how fluently Rosa read (Figure 6). The book she read the least fluently was difficult, but otherwise her results were mixed. She even produced her two most fluent readings for difficult books. Rosa read difficult books at an average of 72 words correct per minute (SD=16 words, ranging from 44 to 87 wcpm) and matched books at an average of 78 words correct per minute (SD=8 words, ranging from 66 to 85 wcpm; compared to the grade level norm of

107 wcpm). Like Alyssa, Rosa's difficult readings showed little trend (m=.314), and her matched readings improved by a word or two per day as the groups continued (m=1.642). Rosa's graph does not show consistent and distinct patterns for matched or difficult books, and it decidedly does not indicate an immediate effect of changing the difficulty level of the books. Only 8% of her data did not overlap with data in the other condition, the lowest percent of non-overlapping data of any student in either group.

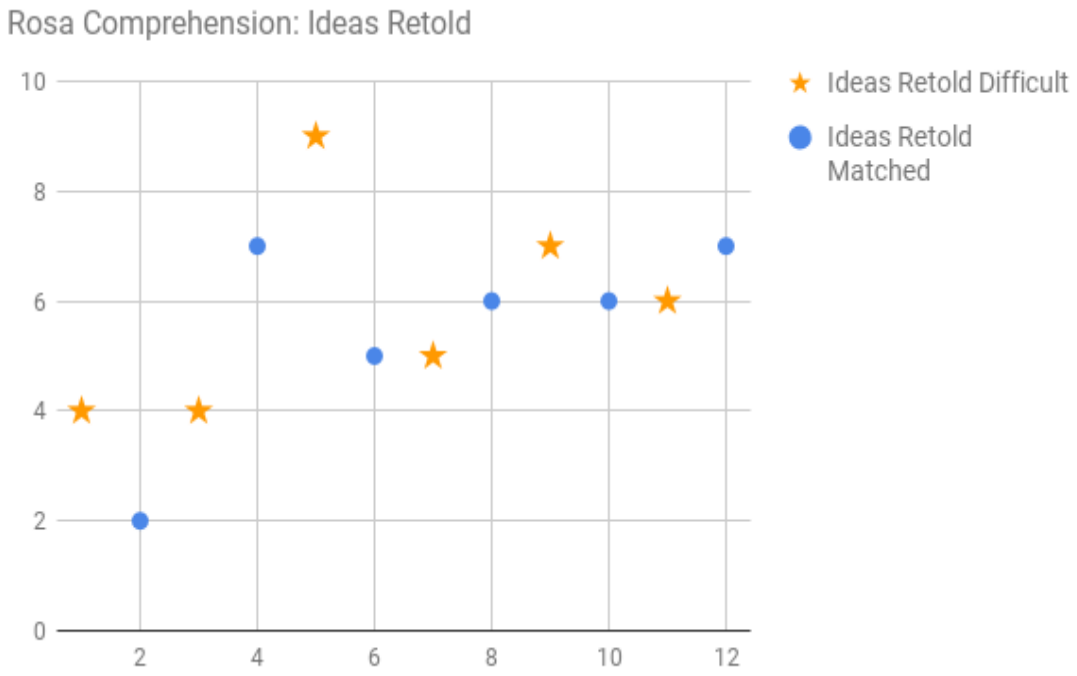
Figure 6. Rosa's Fluency Scores



Comprehension. Just like her fluency scores, Rosa's comprehension scores reflect almost no impact from text difficulty. Paradoxically, she retold the fewest ideas from a matched book, and the most from a difficult book, but otherwise she had mixed results. Rosa retold almost the same number of ideas on average from each group of books: 5.8 ideas from difficult books and 5.5 ideas from matched books. (For difficult

books, SD=2 ideas, ranging from 4 to 9, and for matched books SD=1.9 ideas, ranging from 2 to 7). Figure 7 does not reflect consistency with each level of book, and the slopes of the lines of best do not show a trend over time (m=.214 for difficult books and .328 for matched books). For Rosa, 17% of her comprehension scores do not overlap with the other condition.

Figure 7. Rosa's Comprehension Scores



Rosa's fluency and comprehension initially show no relationship (r=.20). However, she read *Where Are the Night Animals?* with uncharacteristic poor fluency. That day she read 44 words correct per minute when all other days her scores ranged from the high 60s to the high 80s. With this outlier removed, Rosa's data does suggest a moderate positive relationship between fluency and comprehension (r=.64). Both graphs together suggest that text difficulty really did not matter for Rosa's reading performance.

Jack

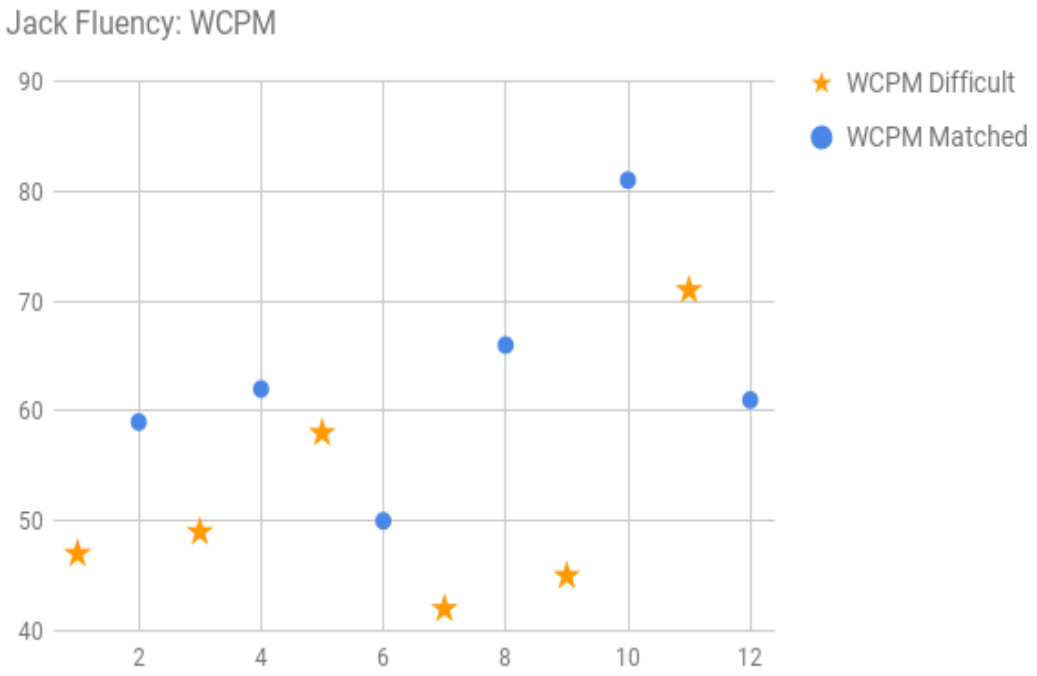
Figure 8. Jack During Discussion

Jack described the reading groups as "fun." He only identified two books as difficult (the two about plants), and he said most of the books were easy. When Jack picked his favorite books from pairs I presented, he picked matched books over difficult books four of six times. Figure 8, a screenshot from the video recording, shows Jack's exasperation when I responded to one of his text-based questions about bears with, "I don't know, Jack. I'm not a bear expert."

Fluency. Text difficulty made more of a difference for Jack than for the other two members of his group. He produced his lowest four readings in difficult texts, and his best reading came from a matched text. On the graph in Figure 9, five of the six highest points represent matched texts, and five of the six lowest points represent difficult texts. Jack read difficult books at an average of 52 words correct per minute (SD=11 words, range from 42 to 71 wcpm) and matched books at an average of 63 words correct per minute (SD=10 words, range from 50 to 81 wcpm; compared to the norm of 107 wcpm). For both matched and difficult books, Jack read about one more word correct per minute

each day over time (m=1.314 for difficult books and 1.185 for matched books). Jack's graph shows that he consistently read matched books better than difficult books, and in most cases (with exceptions on days 5, 6, and 11) he showed a clear effect from difficulty level. 42% of the data did not overlap with data in the other condition.

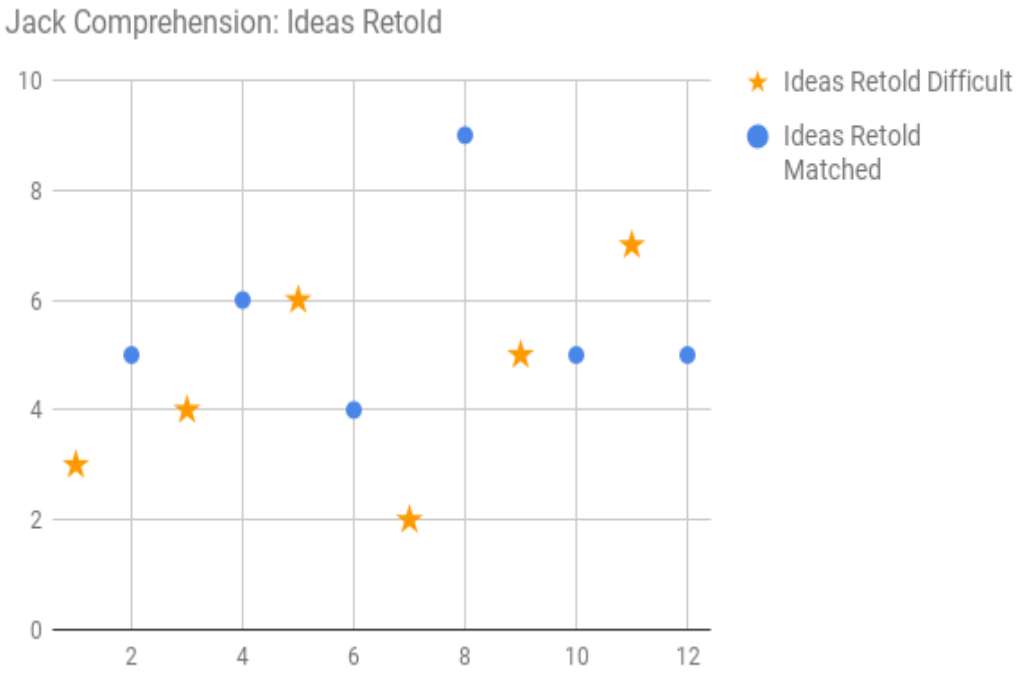
Figure 9. Jack's Fluency Scores



Comprehension. Jack's comprehension scores (Figure 10) suggests a small impact from text difficulty. He retold the most ideas from a matched book, and three of his four lowest scores came from difficult books. In the middle, he has quite a bit of overlapping data. He retold an average of 4.5 ideas from difficult books (SD=1.9 ideas, ranging from 2 to 7) and 5.7 ideas from matched books (SD=1.8 ideas, ranging from 4 to 9). Jack's data do not reflect much trend over time (m=.271 for difficult books and .028 for matched books). Twenty-five percent of his scores do not overlap with the scores

from the other condition. Jack did not consistently retell ideas differently across the difficulty levels, but he did evidence an effect of text difficulty on days 2, 7 and 8.

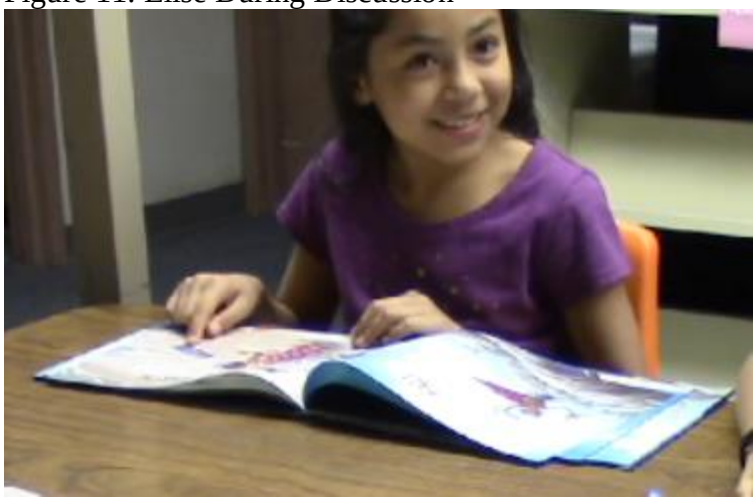
Figure 10. Jack's Comprehension Scores



Jack's fluency and comprehension scores had a moderate positive relationship (r=.64). Jack experienced an effect of text difficulty on his performance more so than the other two members of his group, but the effect, especially on comprehension, was not intense.

Elise

Figure 11. Elise During Discussion

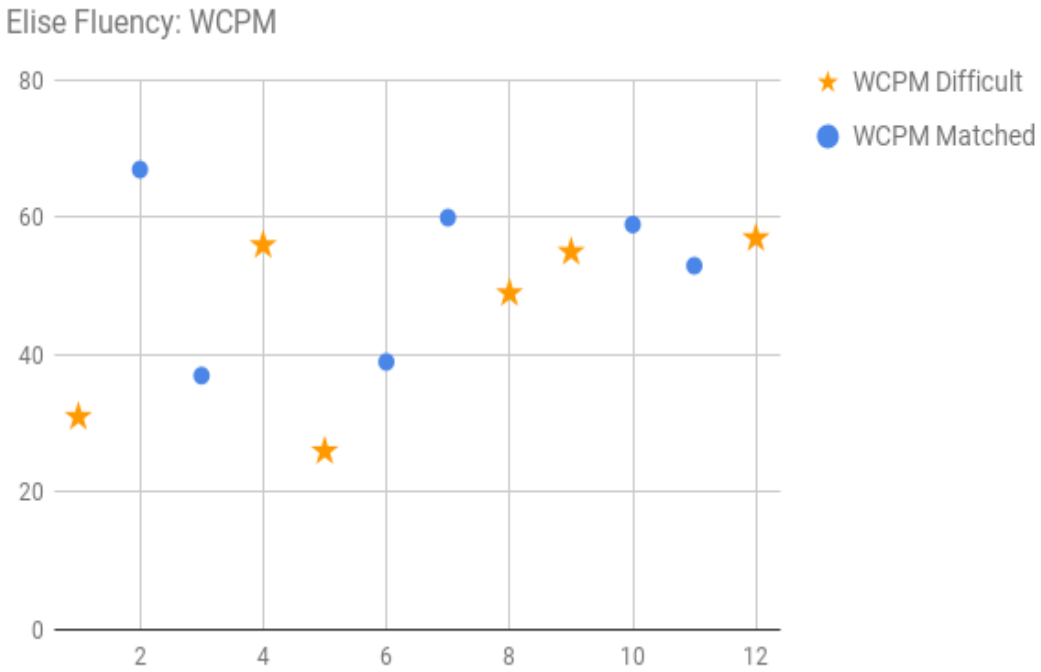


When I asked Elise about the reading groups, she said, "I liked everything." She thought "little words, tiny words, easy words" made books easy. She picked *I am a Frog* as her favorite because it shows "a lot of information about frogs." When I asked her to select her favorite book from pairs, she picked the matched books in four of six cases. Figure 11, a video still shot, shows Elise's face when she realized the illustrations in *Gravity* included pictures of the book itself (copies of *Gravity*) falling through space.

Fluency. Text difficulty did not have a clear impact on Elise's fluency (Figure 12). Her lowest two readings came from difficult books, and her highest three came from matched books. However, the graph shows these data points close together with plenty of overlap across the conditions. Elise read difficult books at an average of 46 words correct per minute (SD=14 words, range from 26 to 57 wcpm) and matched books at an average of 53 words correct per minute (SD=12 words, range from 37 to 67 wcpm; compared to the norm of 107 wcpm). Elise read about 2 correct words more per day in difficult books over time (m=2.258), but readings of matched books showed little trend (m=.374).

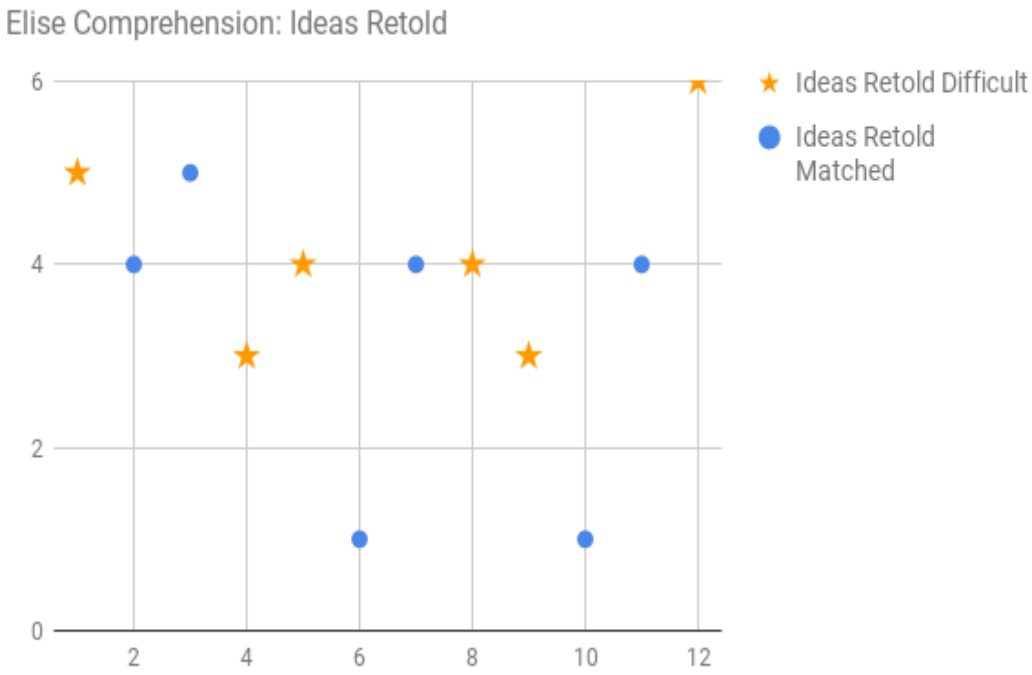
Elise's data reveal neither consistency nor immediate effect from text difficulty. 42% of the data did not overlap with data in the other condition.

Figure 12. Elise's Fluency Scores



Comprehension. Elise's comprehension graph (Figure 13) shows mostly overlapping data, but the non-overlapping points yield a surprise: she retold the fewest ideas with matched texts, and two of her three highest-scored retells came from difficult texts. Elise, on average, retold more after reading and discussing difficult texts—4.2 ideas for difficult texts (SD=1.2 ideas, ranging from 3 to 6) compared to 3.2 ideas for matched texts (SD=1.7 ideas, ranging from 1 to 5). The data do not reveal powerful trends (m=.07 for difficult texts and -.19 for matched texts). Elise did not consistently retell ideas more with matched or difficult books. Overall, just 25% of her comprehension scores did not overlap across conditions.

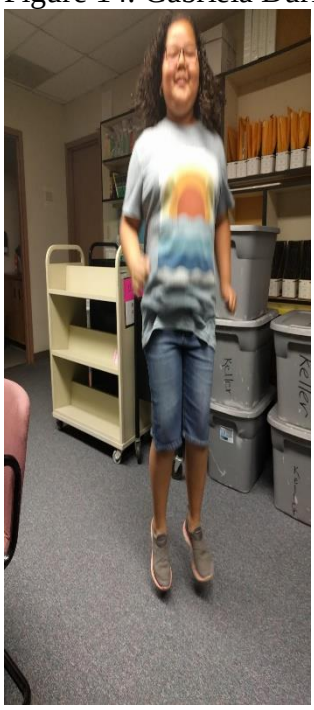
Figure 13. Elise's Comprehension Scores



Plotting Elise's fluency and comprehension together shows almost no relationship whatsoever (r=-.13). The fluency and comprehension graphs together suggest that text difficulty did not make much difference to Elise.

Gabriela

Figure 14. Gabriela During Discussion

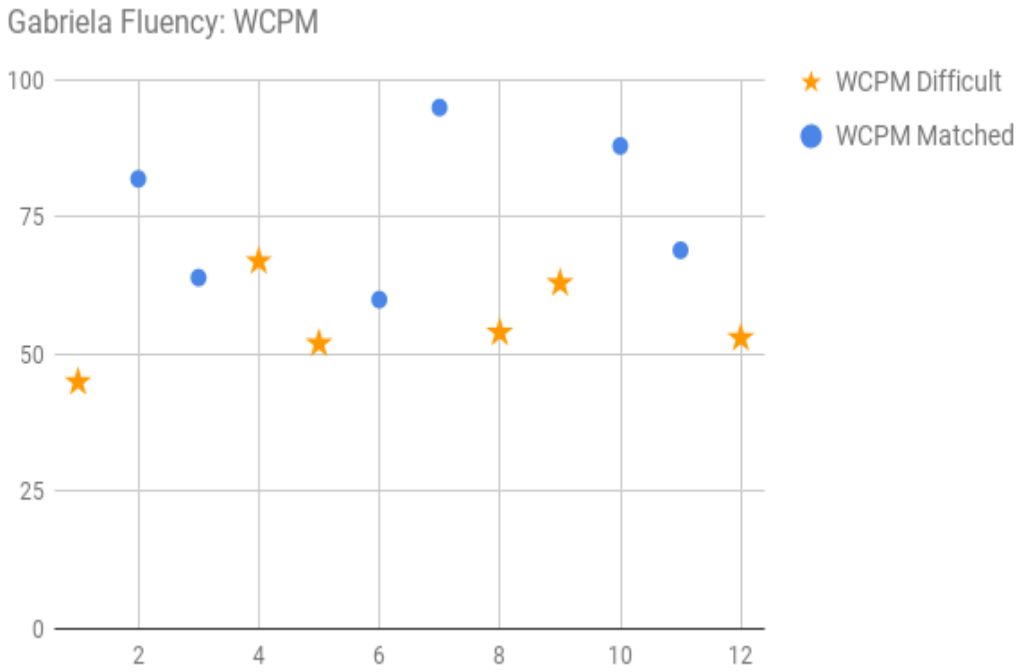


After all the sessions ended, Gabriela reflected, “I liked reading.” She thought easy books “had a little bit of words,” and that “a lot of difficult words” or pages with “a ton of words, and you didn’t know where to start first” made books difficult. She identified *Gravity* as her favorite because “I liked to learn about gravity and stuff in space like that.” When she picked her preferred books from pairs, she picked the matched books over difficult books in five out of six cases. Figure 14 shows Gabriela jumping during a discussion to try to make sense of how the photographer of *Red-Eyed Tree Frog* managed to capture a frog in mid-jump for an image in the book, an artistic and technical feat that fascinated her.

Fluency. Text difficulty made a difference for Gabriela in terms of fluency more than for any other student. Figure 15 shows five of her six highest readings came from matched texts, and five of her lowest came from difficult texts. Gabriela read difficult

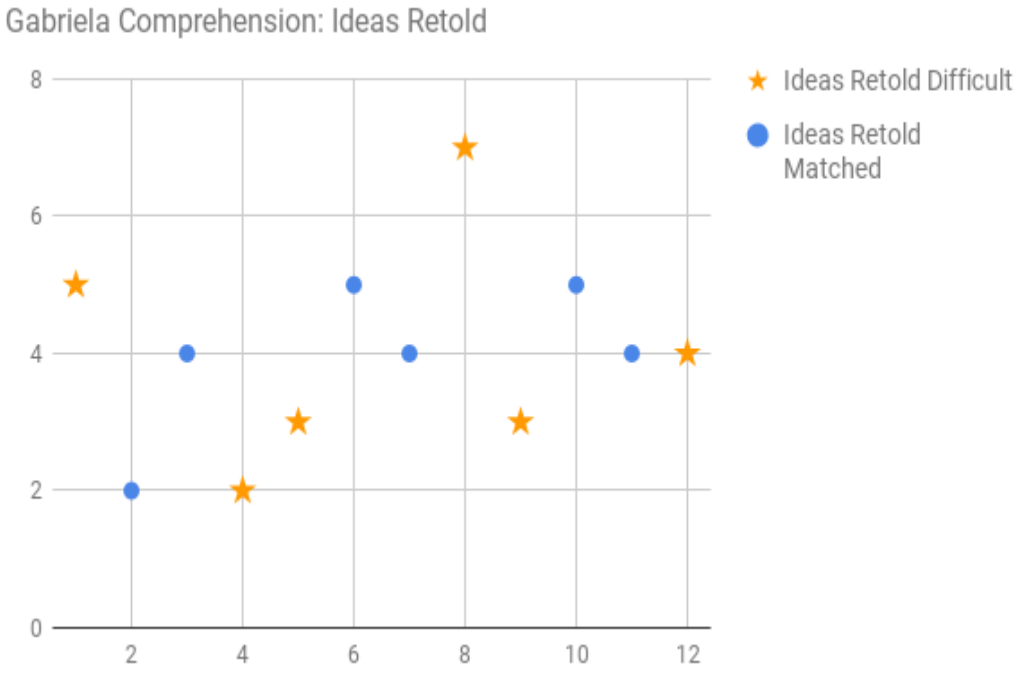
books at an average of 56 words correct per minute (SD=8 words, range 45 to 67 wcpm) and matched books at an average of 76 words correct per minute (SD=14 words, range from 60 to 95 wcpm; compared to the grade level norm of 107 wcpm). Gabriela's data show almost no trending (m=.477 for difficult books and .656 for matched books). Her graph shows that she consistently read matched books more fluently, and it reveals a clear effect of text difficulty on days 1, 2, 7, and 10. Sixty-seven percent of the data did not overlap with data in the other condition. Gabriela's graph shows the highest percent of non-overlapping data among all six students.

Figure 15. Gabriela's Fluency Scores



Comprehension. Gabriela's comprehension scores (Figure 16) show no impact from text difficulty. Only 8% (one data point) does not overlap across conditions. Interestingly, that data point reflects a difficult book for which she retold the most ideas (7); her next highest score was 5 ideas. She retold an average of 4 ideas in both conditions (for difficult books, SD=1.8 ideas with a range of 2 to 7; for matched books, SD=1.1 ideas with a range of 2 to 5). The data show no trending (m=.038 for difficult books and .183 for matched books). Gabriela did not consistently retell ideas more with one difficulty level, and the graph does not show any evidence of effect from text difficulty.

Figure 16. Gabriela's Comprehension Scores



For Gabriela, fluency and comprehension showed little if any correlation (r=-.24). Text difficulty impacted Gabriela's fluency, but not her comprehension.

Sarah

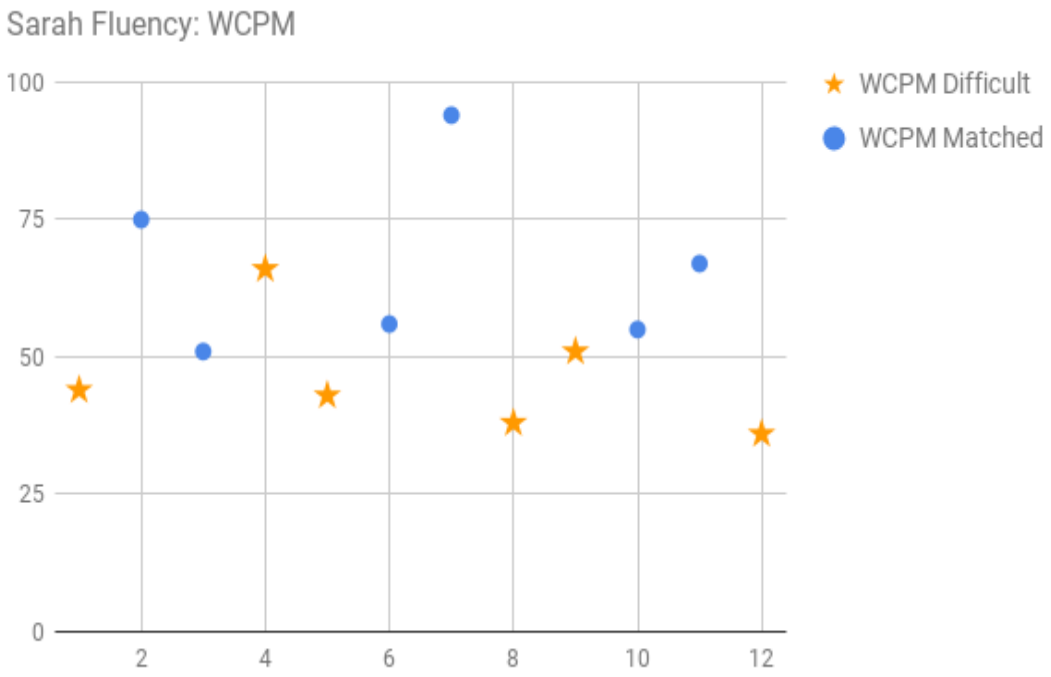
Figure 17. Sarah with her Favorite Book

Sarah shared that she liked our reading groups "because I like the books." Sarah considered books difficult if she had trouble decoding the words "like a second grader." To her, easy books "have a little bit of words and some easy words." She picked *I Am a Frog* as her favorite because she learned specific details about frogs like where they lay their eggs. When presented with the book pairs, Sarah picked her favorite as the matched book (rather than the difficult book) every time, the only student who never identified a difficult book as a favorite. Figure 17 shows Sarah reading her favorite book from the groups.

Fluency. Sarah's graph (Figure 18) shows a glut of overlapping data from matched and difficult books in the middle, but it also reflects some impact of text difficulty for her in terms of fluency. Like Jack and Gabriela, five of her six highest

readings came from matched texts, and five of her lowest came from difficult texts. Sarah read difficult books at an average of 46 words correct per minute (SD=11 words, range from 36 to 66 wcpm) and matched books at an average of 66 words correct per minute (SD=16 words, range from 51 to 94 wcpm; compared to the norm of 107 wcpm). Over the course of the groups, Sarah began to read difficult books slightly less fluently (m=-1.148) and developed no trends with matched books (m=-.045). Sarah's graph makes clear that she consistently read matched books more fluently, and days 1, 2, 5, 7, 8, and 12 suggest clear responses to text difficulty. 58% of the data did not overlap with data in the other condition.

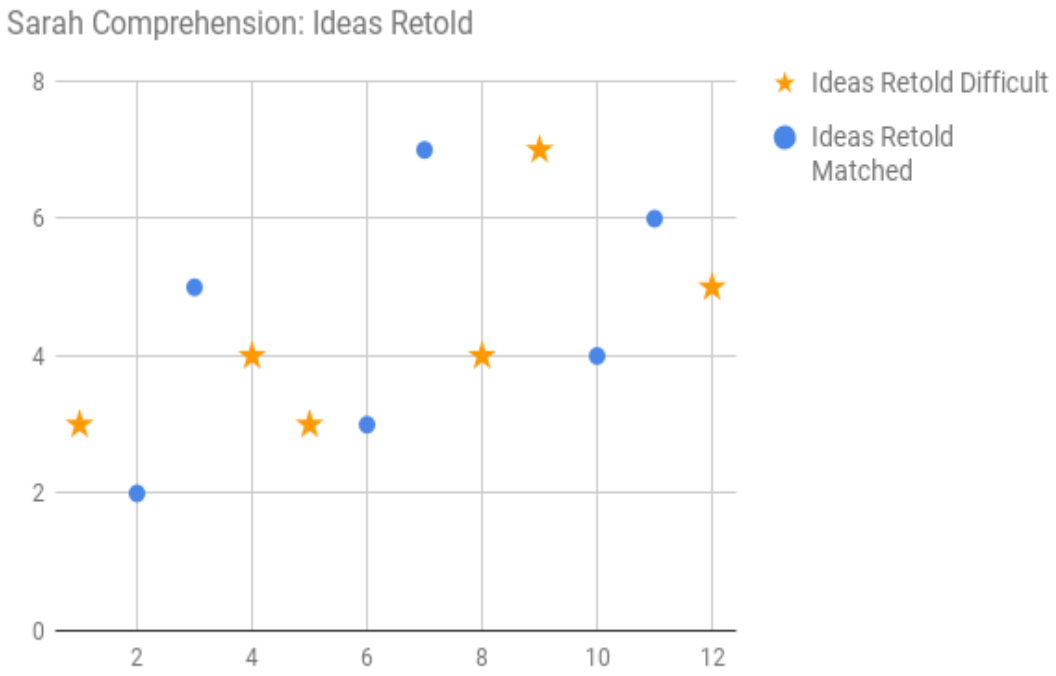
Figure 18. Sarah's Fluency Scores



Comprehension. In contrast to her fluency results, Sarah's comprehension results (Figure 19) show no evidence that text difficulty mattered to her. Only 8% of the data do

not overlap, meaning that 11 out of 12 data points do overlap. She retold an average of 4.3 ideas from difficult books (SD=1.5, ranging from 3 to 7) and 4.5 ideas from matched books (SD=1.9 ideas, ranging from 2 to 7). Her graph shows neither consistency within difficulty levels nor effects upon changing difficulty levels. She had minimal trending (m=.258 for difficult books and .251 for matched books).

Figure 19. Sarah's Comprehension Scores



For Sarah, fluency and comprehension had little if any relationships (r=.271), and text difficulty only had an impact on fluency.

Traditional Summary

Fluency. Each student responded differently to text difficulty in terms of fluency. All the graphs show some overlap between how students read with difficult and matched texts. In some cases, they show a lot of overlap (ranging from 8 to 67% of the data).

Because of this high degree of overlap, none of the graphs show an un-nuanced functional relationship between text difficulty and fluency. Rather, for some students it mattered more, and for others it mattered less.

All the students read on average more words correct per minute with matched books than difficult books. But, all students read at least one difficult book with similar fluency to how they read matched books, and all read at least one matched book with similar fluency to how they read difficult books. The standard deviations and the ranges showed a lot of variation within matched and difficult texts. Students varied by up to 43 words correct per minute with texts from the same difficulty category!

In general, the data did not reflect big trends over time. (In single case design, one does *not* want to see trends within a condition because it would suggest something other than text difficulty making a difference.) Elise improved with difficult texts a little over time, Sarah read them a little worse, Alyssa and Rosa started to read matched texts a little better, and Jack started to read both levels about one word faster. In all cases though, these improvements consisted of 1 or 2 words each day. Gabriela had no change over time.

For Rosa, text difficulty seemed to make almost no difference, and Elise and Alyssa also had mixed results reflecting minimal impact. While Jack, Gabriela, and Sarah did have lots of overlapping data, their graphs showed that they consistently read more fluently with matched books than difficult books. All three performed five of their six most fluent reads with matched texts.

Gabriela and Sarah came from group 2, the group reading the most below grade level, and Jack's average fluency lagged behind his group members (reading an average

of 63 wcpm with matched texts compared to their 96 and 78). These cases suggest that text difficulty makes more of an impact for lower readers, but Elise defies this conclusion. She had the lowest reading profile of all the students (according to her teacher who did not share specific data other than that she recently began receiving special education services for reading), and yet her graph indicates minimal impact of text difficulty on her fluency.

The students in Group 2 seemed to respond to the books similarly. Drawing a line through all their data points results in a similar line across all three graphs. For example, everyone read more on day 2, less on day 3, back up on day 4, down a bit on day 5, up a tad on day 6, up a lot on day 7, down on day 8, and up more by days 9 and 10. The group members do diverge from one another on days 11 and 12. However, the graphs from Group 1 do not all tell the same fluency story. The differences become apparent almost immediately, by day 2. Comparing the second day to the first day, Alyssa read the exact same number of words correct per minute, Jack read more, and Rosa read less. The differences continue, with a few exceptions: everybody struggled on day 9 with one of their lowest readings, and everybody read well on the last three days. These findings may support the idea that text difficulty more reliably predicts students' fluency for lower readers (those in Group 2). They may also suggest that the differences between matched and difficult books become less significant (and so less predictive) as reading levels increase.

Comprehension. Text difficulty did not affect anybody's comprehension consistently. Jack showed some evidence of an effect when changing conditions, but only on 3 of 12 days; otherwise he did not have consistent results either. Generally, after a

brief read aloud, time for independent reading, and a small-group discussion, students retold ideas from difficult books at roughly the same rates as they retold them from matched books. All the students had minimal comprehension data across matched and difficult books that did not overlap (ranging from 8 to 25%), and the ranges of the number of ideas they retold varied considerably even within texts of the same difficulty band.

Students showed little trending across the twelve days of reading groups. None of the slopes of the lines of best fit for matched and difficult books suggested a change of even half an idea over the trajectory of the groups. Interestingly, only Jack (and Rosa with the outlier removed) had a moderate relationship between fluency and comprehension. For the other students, comprehension and fluency had no correlation at all. Appendix D summarizes findings by book, rather than by student.

Qualitative Findings

When I designed the qualitative arm of this mixed-methods study, I meant for the qualitative data to shed light on students' engagement and discussion with texts of different difficulty levels. I thought that students might express frustration with difficult texts or maybe excitement about the challenge. I imagined that the data would paint a portrait of qualitatively different groups: that the ways students used strategies, participated, and responded would vary according to the text difficulty in big and immediately obvious ways. That did not happen. My observations, the research assistant's journals, the students' comments and actions, and the classroom teacher's report all suggested that the students enjoyed coming to the groups regardless of which book we read that day, and they did not change their behavior in response to the difficulty

level of the books. At no point did they evidence that they even noticed that some books presented more of a decoding challenge than others. However, after analyzing the transcripts in depth, some subtle patterns of engagement and participation that differed according to text difficulty became clear. While no one participating in the groups identified these patterns while they were happening, the data illuminates the way discussions developed around texts of different difficulty levels.

Engagement

For the purpose of this study, I defined engagement as students' social interaction and strategy use around texts in an effort to construct meaning. (See Guthrie & Anderson, 1999; Unrau & Quirk, 2014). I report findings related to strategy use and interaction below, but I begin with another category that I developed from reviewing the data. As described in the methods chapter, I created the code *focus of engagement* to capture the points of contact for student engagement: the ideas, pictures, and text.

Focus of engagement: How do students engage with the books? When I designed the reading groups, I expected most of the discussion to center on the text. I planned to ask text-based questions, and I thought students' contributions to the discussions would come clearly from the text. However, I observed early that students engaged with pictures and ideas more than they did the text. Table 4 shows the number of times each student and each group as a whole engaged at each of these points.

Table 4
Focus of Engagement

	Ideas (D)	Ideas (M)	Pictures (D)	Pictures (M)	Text (D)	Text (M)
Alyssa	58	60	22*	14*	16	19
Jack	51	47	21	22	20	21
Rosa	11	13	1	0	14	12
Group 1 totals	120	120	44	36	50	52
Elise	15	11	8	10	9*	3*
Gabriela	50*	38*	9*	31*	15	13
Sarah	66*	47*	15*	27*	25	22
Group 2 totals	131*	96*	32*	68*	49*	38*

Note. The table shows the number of times each student engaged with ideas, pictures, or text through talk in group discussions. Columns marked with (M) refer to discussions about matched texts, and columns marked with (D) refer to discussions about difficult text. Numbers marked with a * indicate a difference of greater than 6 instances of engagement between the matched texts and difficult texts.

Engaging with ideas. Engaging with ideas referred to students talking or asking about ideas introduced in the text without specific reference to a particular text. Examples of this type of engagement include asking general questions, sharing connections, and answering teacher questions without specific reference to the text.

In discussing *Seed, Soil, Sun* (a difficult book) Rosa mentioned that she had learned about using worms to compost from a TV program. While the book discussed the ways that worms return organic matter to the soil, Rosa's intertextual connection (to a media program) showed how she engaged and then extended an idea from the text without specific reference to the words on the page.

Rosa: I saw on TV that worms were eating… That people put like a bucket, and they put soil, then they find worms. They put food. Then, they let them free to go.

Alyssa: So, they're helping them.

Researcher: So, they put the worms in there to do what? Jack, you were saying like a bucket of rotten food?

Jack: Mmhm.

Researcher: Yeah, the worms will eat through the food and turn it back into soil with their droppings.

Alyssa: So, they eat it? The droppings?

Researcher: No, they don't eat their OWN droppings. They eat food and the things that are in the soil. It says they eat "debris," so like pieces of dirt, leaves, rotten food, like Jack was saying.

In this example, I turned students back to the text by drawing their attention to the word debris, but student contributions to the discussion revolved around ideas.

With both matched and difficult books, students frequently engaged with ideas to answer teacher questions. That is, they would respond to me in discussion by stating ideas introduced in the book without specifically referring to the words, as in the transcript about *Penguin Chick* (a matched text) that follows.

Researcher: The father penguin is gone. Why?

Rosa: Because the baby penguin is growing up.

Jack: 'Cause the baby penguin can't get in the dad's patch.

Researcher: He doesn't fit in the brood patch anymore. And, what else? What did

he need the father penguin for?

Alyssa: To protect the baby chick from the egg.

LK: Okay, so he protected him. What else did the parents do?

Jack: The dad kept it in its pouch, where it didn't roll away.

Other times students engaged with ideas by raising questions that the text made them think of. For example, in the same discussion of *Penguin Chick*, Alyssa wondered, "Does the mama sleep in the ocean? Or, does she get out?" and Jack became curious about whether penguins have nostrils. Students also often engaged with ideas when they explained why a particular page captured their attention as when Rosa said, "...the noses are interesting so they can dig. It's like a shovel" while discussing *What Do You Do With a Tail Like This?*

Students of course also combined engaging with ideas with engaging with pictures and text. In the next transcript from a discussion of *Best Foot Forward* (a difficult text for this group), Gabriela pointed to a picture (engaging with pictures), Sarah started reading (engaging with text), and then the group had a conversation that built up several related ideas (engaging with ideas) before the girls chorally read (engaging with text again) to confirm their ideas.

Gabriela: Why are there things right here? (pointing to hard spots below padding on tiger foot)

Sarah: "With the soft, cushioned pads on its feet, a tiger can creep up very quietly on its prey, and then click! out comes the claws!"

Researcher: Okay, so from what you were just reading what does it have? What are these things on its feet?

Sarah: Claws!

Researcher: They're not the claws because the claws are these parts. (flipping back and showing picture)

Gabriela: Cushions.

Researcher: The cushions! It said they have cushioned pads on their feet. Why would they need cushioned pads?

Gabriela: So it feels comfortable where it is.

Researcher: Yeah, it wouldn't want to always be hitting very hard and hurting.

Gabriela: Like their shoes.

Researcher: Yeah. Right. You guys have shoes so you're not always hitting the bottom of your feet. The tiger has its cushioned pads.

Sarah: Or, like if it was too hot and the sun was burning like here, and then if you would step on it, you could burn yourself.

Researcher: Mmhm. One of these other animals had cushioned feet. Do you remember which one?

Elise: This one.

Researcher: Let's see.

Sarah: Elephant!

Researcher: What does it say about the elephant and its feet because it's so heavy?

All 3: (overlapping reading)

Researcher: Yeah. What word do you see that's the same between here and here?

All 3: Cushion.

Sarah: That they have cushion.

When students shared ideas about feeling comfortable, cushions being like shoes, and cushions providing protection from heat, they engaged with ideas in between engaging with pictures and text.

Students in Group 1 engaged with ideas exactly the same number of times with matched texts as with difficult texts. However, students in Group 2 engaged with ideas quite a bit more in difficult texts.

Engaging with pictures. Engaging with pictures referred to students talking or asking about pictures. Common examples include commenting on what animals looked like, expressing amazement at a photo, asking what something in the picture is, answering a teacher question by pointing to a photo, and asking how the artist drew or photographed something.

Sometimes students engaged with pictures superficially as when Alyssa looked at a wasp's nest in *Acorn to Oak Tree* (a matched book) and said, "Dude, that looks awesome!" or Jack blithely commented about the star-nosed mole that its facial features "look like worms." More commonly though, students used pictures as part of their sense-making process. For example, Sarah asked about why birds fly in a v-formation after seeing the illustration in *About Birds* (a matched book), and Elise compared illustrations of birds in the same text. *About Fish* (also matched) has minimal text, so in the following transcript, the students used the pictures to supplement their understanding of how seahorses are born.

Elise: So, the seahorse is here, and the one spray is coming out of his stomach. What's happening there?

Gabriela: Just, they just poked out?

Researcher: They just come out, right? What are they doing? // This is the moment that they're being born. Do they have eggs, or do they just come out that way?

Gabriela: Just come out that way.

In reading *Antarctica* (a matched book), Jack used the pictures to identify a relevant page to answer his and Alyssa's question. After she found the same page, Alyssa also used the pictures to address their question, and Jack used them to disagree with me and substantiate his point that scientists had established a large camp, not a small one, to do their research in Antarctica.

Jack: Do people still live there [Antarctica]?

Alyssa: People don't live there, right?

Researcher: People don't live there, but people are there. So, what are they doing?

Jack: Exploring.

Researcher: Yeah, they're exploring. Sometimes they're scientists. They're studying the land and the ice and the animals.

Alyssa: Like, if they visit over there, do they have to like camp or something?

Researcher: Was there something in the book about that?

Jack: Yeah. Here. (Pointing to the page he already has open)

Alyssa: "Just behind …" Oh, "men build a base camp"

Researcher: So, they don't live there, but they do have a camp for where they stay.

Alyssa: (pointing to picture) That?

Researcher: Yeah, good job finding that. I didn't see that at all. But, yeah, a little

camp back there in the snow.

Jack: It doesn't look little because it has another part right here. (pointing to illustration continuing across page break)

Researcher: Oh. Yeah. It's on both sides. So, you're right. It's not little. There's lots of space there. So, what are the people doing?

This interaction shows how students attended to all the resources in the book while I was hyper-focused on the text. Jack had paid attention to the images in the book (when I had not), and he and Alyssa used them as a resource for addressing their own questions even as I tried to push them back to the text to find where it said people built a base camp.

In other cases, the pictures sparked conversations that provided opportunities to push students back into close reading of the text. In reading *Red-Eyed Tree Frog* (a matched book), Elise became interested in why the frog was photographed sitting by an ant if it was looking for food, but was not actually going to eat the ant.

Elise: I thought the frog going to eat the ant, but it didn't. If the frog's not going to eat the ant, is it just staring at it?

Researcher: Yeah, that's a really good question. Why is the frog even there if it's not going to eat the ant? What do you think?

Gabriela: To get a better look at it.

Researcher: Why does it need a good look?

Gabriela: Because it might not-

Sarah: It might be poisoned or

Gabriela: Its prey.

Elise: (shrugging) Something like that.

Researcher: What do you think happened right after they took this picture?

Elise: Maybe the frog ate it already.

Gabriela: Maybe the frog left.

Researcher: So, those are two different answers, and the book supports one of those ideas more than the other. Do you think the frog ate it, or do you think the frog left based on what it says right here?

Gabriela: Left.

Sarah: Left.

Researcher: Probably left because what does it say?

Sarah & Elise: "The frog is hungry, but it will not eat the ant."

Here the pictures led to a conversation in which students offered multiple explanations and then evaluated them in light of the text.

In some cases, I wondered if students focused on the pictures in order to avoid reading difficult text. For example, in the difficult book *Best Foot Forward*, Group 2 elected to largely ignore the text while letting innumerable questions and comments about the photographs drive the discussions: What are these red things? Do they hurt? Is this a cut? It's weird how it has these holes. Do these open? Are those feet? Why does it have a big one, a small one, and a medium one? Why are there things right there? What are these little things? Why is it a different color? Obviously, the data do not address whether these questions reflect genuine childhood curiosity in response to a compelling book format with close-up photos, a desire to avoid difficult text, or a combination of both. The engagement with pictures for Group 2 with this book did represent a departure from their

normal behavior though: they engaged with pictures twice as often in matched books as in difficult books.

Group 1 engaged with pictures more commonly in difficult texts, but not much more commonly. In contrast, as just noted, Group 2 engaged with pictures over twice as often with matched texts as with difficult texts. Since Group 2 read further below grade level than Group 1, for them matched texts were relatively simple. These books often had significantly less text on a page than the difficult texts, overall contained less text in the whole book, and may have developed less complicated ideas.

Engaging with text. Engaging with text included students talking or asking about one or more exact words in the text. Students engaged in this way when they reread, asked what words meant, quoted, referenced exact numbers from text, answered questions with words from text, or directly referenced text by introducing a paraphrase with "it says…"

Most commonly students engaged with text to answer a question that I asked or when I pushed them to in the course of discussion. For example, in discussing *About Fish*, Gabriela referenced the text to report that fish lay 77 to 190 eggs. Jack referred to the text when he introduced a paraphrase with "it says" and then went on to explain that a diagram in *Penguin Chick* showed winter, spring, and summer in a penguin's growth. Sarah used language from the text when she answered a question about dandelions (in discussing the matched book *Dandelions*) by explaining "they bloom like golden stars." The following transcript from the matched book *I Am a Frog* provides a representative example of how student engagement with text often resulted from me pushing them into the text.

Gabriela: But, the eggs that they lay, do they do like turtles? Do they just leave it? Do they just leave them, or do they stay with them?

Researcher: So, we read something about that. It's actually on this page too. Do they just leave them, or do they stay with them?

Sarah: Some leave them, and some stay with them.

Researcher: Okay. And, tell us how you know that from the book.

Sarah: Because it tells you on "Super Dad" (the heading on a call-out box).

Researcher: What was the first sentence, there, Gabriela? What do you see there?

Gabriela: "Most frogs don't look after their babies."

Researcher: Most frogs don't. And, then, what's the very next sentence? That helps you too.

Gabriela: "They lay eggs, and swim away before they hatch."

Even after the read-aloud and independent reading, basic questions addressed in the text lingered for Gabriela. Sarah answered Gabriela's question and even knew where to find the answer in the text, but it took my prompting to drive Gabriela to attend to the relevant text about her question.

Students also commonly returned to the text when I pushed them to use the language of the books to talk about their ideas. In discussing the matched book *Acorn to Oak Tree*, Alyssa brought up the different color of the top part of the acorn:

Researcher: Yeah, what's that top part called? What'd you guys learn?

Jack: Stem.

Researcher: It has a special name. Oh, it's actually on this page. Page 8 and 9. See if you can find it because the diagram has it.

Rosa: Cops.

Researcher: Say it again.

Rosa: Cops.

Researcher: It's close to cops. It looks like a cup, but look at this word here.

Alyssa: Cup-pool.

Researcher: I think you probably say cup-ule. So, Alyssa was wondering... She knows that the seed turns brown when it gets ripe. She was wondering if the cupule turns brown.

Jack: It does.

Some of the most frequent engagement with text occurred in instances like this one where students collaborated to sound out words as I pushed them back to the text to use the vocabulary of the books to talk about scientific ideas.

When students engaged with text, they frequently read in labored word-by-word reading that caused the rest of the group to lose focus. In this discussion of a matched text (*What Do You Do With a Tail Like This?*), the students became interested in the afterword material that provided much more detail about each kind of animal in the book. Thus, they elected to discuss more difficult material from a few end pages despite having a matched book.

Researcher: Okay, Rosa, tell us about it.

Rosa: "The horned lizard, often called a horny toad lives in Afri- the American Southwest. It is small, 3 to 5 inches in length and covered with sharp spikes. The lizard feeds on ants and other insects and protects itself ..."

(Alyssa and Jack looking around. Rosa's reading is mostly word by word or

phrase by phrase.)

Researcher: Oh, hang on. You skipped a line there. “It protects itself in an unusual way... If threatened, it first tries holding very still…” And, then you can pick up. “If that doesn’t work…”

Rosa: “If that doesn’t work, it puffs itself up with air to make itself larger.”

Researcher: And, then if it still feels threatened, what’ll it do next?

(Jack opens book.)

Alyssa: Its eye will blink.

Researcher: It says, “It will squirt streams of blood from the corner of its eyes.”

Alyssa: Where does it live?

Jack: Desert.

Researcher: Yeah, in the desert. Where did it say?

Alyssa: (surprised) Here in Arizona?

Researcher: Well, what did Rosa just read? (All three look at book.)

Alyssa: Arizona southwest.

Researcher: The American southwest. So, Arizona is in the American southwest. So, maybe. Could be. So, watch out for bleeding lizards. (laughs) It says that it squirts the stream of blood from the corner of its eyes. Why does it do that?

Alyssa: It’s angry.

Jack: To protect.

Researcher: We should probably read that last sentence. It explains. (Rosa is looking behind her. Alyssa and Jack begin reading slowly, and Alyssa in the wrong place.)

Researcher: Where it says, “This probabl-” Yeah, Jack.

Jack: “blood from the corner of its eyes. This probably confuses its predator in time to get away.” (Reading is word by word. Rosa is looking behind. Alyssa is looking at book and researcher.)

While this transcript came from a matched discussion, it tells more about how students respond to difficult text since they chose to discuss the afterword. They evidenced interest in the topic: Rosa introduced an idea, Alyssa asked questions about it, Jack answered them, and they read together to confirm their thinking. But, they struggled. They read word by word, they skipped lines, and they kept losing their place. When their peers read in this way, they lost focus, despite caring about the topic and participating in the discussion: they leaned back or they looked around the room, and they could not answer questions based on what others had just read.

Other times, students read from the text to participate in the discussion, but they did not know how to use the text to move the conversation forward. Rosa seemed shy by nature and made fewer comments and questions than the rest of her group. Reading from the text (both matched and difficult) gave her a way to participate without having to think of something original to say, but it did not inspire group discussion. When asked to share her thinking, she often picked a favorite page, said “it says…” and started reading. Jack and Alyssa did not follow along with her reading or know how to build a discussion around it when she stopped.

So much text engagement in this study arose from my prompting, reflected lower-level responses, or did not stimulate group discussion as reported above. However, sometimes (less commonly) students used the text for purposes that both reflected

comprehension and moved the conversation forward. In this transcript from the difficult book *Seed, Soil, Sun,* Alyssa selectively read to explain something that captured her attention.

> Alyssa: I marked this page because it's cool when the cow takes some of its milk and like in the milk it has plants, grass, and corn and soil beans [sic].
>
> Jack: Does a cow eat and then-
>
> Alyssa: And, "when you drink milk from a cow, you are drinking a food made of what the cow eats."
>
> Researcher: You don't usually think milk is a plant right because it doesn't grow on a tree or anything, but Alyssa really understood this page because it says really milk *does* come from plants because the cow can't make any milk if it doesn't do what?
>
> Alyssa: Eat.
>
> Researcher: Eat plants, right? And, the cow eats plants. And, Alyssa read all those plants it eats.

In the ensuing discussion (not reported here), Alyssa and Jack co-developed this idea further. Alyssa's use of the text here clarified her original claim and showed Jack that her idea came from the book.

In one final example of using the text to move conversation forward and develop comprehension, Jack used text to clarify a point of confusion for Alyssa in discussing *Looking Closely in the Rain Forest*, a difficult book. Alyssa marked a page that piqued her interest so the group could discuss it. That page discussed how banana plants have leaves so big that some people use them as umbrellas.

Alyssa: Like, the peel a banana, they make an umbrella out of it.

Researcher: Oh, they do make an umbrella from something from that tree. It's not the peel though. Read that again. (Alyssa silently looks back at text. Jack says the leaves.) Show her where you found that.

Jack: (pointing) Oh, here. "The people use banana leaves as umbrellas."

Students sometimes used texts to develop points and support ideas, but most commonly they used text to answer questions, make basic observations, find book language when I sent them back to the text, or read in a way that derailed other group members. These common uses occurred with both matched and difficult texts although disfluent reading that caused students to lose interest happened more with difficult texts. While students seemed eager to talk, they did not seem eager to read (or listen to anybody else read) as part of the discussion.

Summary. As Table 4 shows, the students varied in whether text difficulty correlated with different types of engagement. Parametric statistics do not apply to these data to provide useful cut-offs for determining whether the differences in engagement between matched and difficult texts matter. Lacking clear statistical guidelines, I have chosen to indicate differences of 6 or greater with an asterisk since 6 represents half of the twelve-day study.

As a whole, the results indicate much less talk about text than expected for text-based discussion groups and suggest that students may benefit from support that pushes them to connect their discussion more strongly to the text. The students in Group 1 talked mostly about ideas, then text, and lastly pictures, and this pattern did not change according to the difficulty level of the books. However, the students in Group 2, who read

further below grade level than Group 1, did engage differently with books of different levels. While they did not all respond the same way, as a group, they also talked mostly about ideas, but they talked much more about pictures in matched books and more about texts in difficult books. While they talked mostly about ideas with both levels of books, difficult books generated talk about ideas over twice as often as matched books did.

Strategies: What comprehension strategies do students use in discussion? I coded for these comprehension strategies in the students' discussions: inferencing, asking questions, research, personal connections, gesture, fix-up/decoding strategies, textual connections, background knowledge, genre use, cognates, and summary. Because of formatting constraints, the number of times students used each strategy for difficult and matched books is reported across two tables, Tables 5 and 6. I did not model or intentionally teach strategy development in the discussion groups, so these strategies reflect the ones that students used independently without instruction or ones that they used when I asked something that created an opportunity for them to use a strategy. Students' most common strategies included asking questions and making inferences, but I include these strategies in subsequent sections because of how they overlapped with other codes. I discuss inferring later when I describe students' types of responses, and I include asking when explaining students' interaction patterns. Here I discuss all the other strategies students adopted.

Table 5
Most Common Comprehension Strategies

	Infer (D)	Infer (M)	Ask (D)	Ask (M)	Research (D)	Research (M)	P. Connect (D)	P. Connect (M)	Gesture (D)	Gesture (M)	Fix-up (D)	Fix-up (M)
Alyssa	24*	30*	8*	20*	0*	15*	2	3	5*	13*	6	6
Jack	21*	28*	10*	18*	0*	18*	1	2	5	8	5	5
Rosa	1	6	0	1	0	5	1	4	2	3	2	1
Group 1 totals	46*	64*	18*	39*	0*	38*	4	9	12*	24*	17	14
Elise	6	10	9	8	1	0	4	0	1	1	1	1
Gabriela	29	21	21*	28*	0	10*	7	10	2	5	1	1
Sarah	38*	23*	32*	18*	4	3	9	12	3	2	3	5
Group 2 totals	73*	54*	62*	54*	5	13*	20	22	6	8	5	7

Note. The table shows the number of times each student implemented comprehension strategies out loud in the discussions. Columns marked with (M) refer to discussions about matched texts, and columns marked with (D) refer to discussions about difficult text. Numbers marked with a * indicate a difference of greater than 6 instances of engagement between the matched texts and difficult texts. P. Connect stands for personal connection.

Table 6
Less Common Comprehension Strategies

	Text Connect (D)	Text Connect (M)	BK (D)	BK (M)	Genre (D)	Genre (M)	Cognates (D)	Cognates (M)	Summary (D)	Summary (M)
Alyssa	7*	1*	3	2	0	0	1	0	0	0
Jack	6*	0*	6	8	0	0	0	0	0	0
Rosa	2	1	0	0	0	0	0	0	2	0
Group 1 totals	15*	2*	9	10	0	0	1	0	2	0
Elise	0	1	0	0	0	0	0	0	0	0
Gabriela	0	5	1	2	2	6	0	2	0	0
Sarah	1	4	1	2	3	1	0	4	0	0
Group 2 totals	1*	10*	2	4	5	7	0*	6*	0	0

Note. The table shows the number of times each student implemented comprehension strategies out loud in the discussions. Columns marked with (M) refer to discussions about matched texts, and columns marked with (D) refer to discussions about difficult text. Numbers marked with a * indicate a difference of greater than 6 instances of engagement between the matched texts and difficult texts. BK stands for background knowledge.

Research. Students participated in research when they used peritextual material (afterwords, authors' notes, glossaries, etc.) to learn more information about a topic introduced in the main text. All instances of research for Group 1 (and most for Group 2) occurred with matched texts. For example, Elise prompted her group members to turn to the back of the matched text *About Birds* to learn more about the v-formation. Gabriela turned to the recipe at the end of *Dandelions* to learn what part of that plant people can eat. In talking about a specific spider in *Spiders,* Elise used the photo glossary at the end of the book to identify the name of a spider species. Alyssa took an interest in the bush baby (a small, African, nocturnal primate discussed in the matched book *What Do You Do with a Tail Like This?*), and during independent reading she looked up more information in the back of the book which she later shared through paraphrasing and reading during the group discussion. Her sharing prompted other group members to look up details not reported in the main text of the book in the afterword material.

Whether or not students could use research as a strategy depended on if the book contained additional information beyond the main text. In Group 1, three of the difficult books contained extra features and four of the matched books did. In Group 2 (who used lower level books), only two of the difficult books but five of the matched books had extra information after the main text. Authors who wrote books targeting higher reading levels made less use of peritextual features because they simply included more detail in the main text. However, authors who wrote simple texts at low levels often included additional information in glossaries, activity/experiment/recipe instructions, photo indices, afterwords, and variously titled additional information sections (find out more,

notes for parents and teachers, a closer look, did you know?). Thus, students mostly researched with matched texts.

Personal connection. Students made personal connections by connecting the text to their experiences. They did this evenly across difficult and matched texts, but Group 2 did it much more than Group 1 with both levels of text. Group 2 often recounted personal narratives that the text reminded them of, and they did not always demonstrate skill at tying these back to the text or making explicit the ways in which their personal experiences supported their understanding of the book.

> Gabriela: Oh, I have gotten bit by an ant before, and it really hurt.
>
> Sarah: Oh yeah.
>
> Gabriela: On the next day, I was still playing around a tree, and there was a little sand castle that they were making, and I didn't know, and on accident, I probably stepped on it, and some ants got in my shoe and bit my-
>
> Researcher: So, what does that experience tell you about... It says, "The frog is hungry, but it will not eat the ant." Why will it not eat the ant?
>
> Elise: It bites.
>
> Gabriela: It's not its prey.
>
> Researcher: It's not its prey because it bites. It doesn't want to have an experience like you where you got bit.

In this transcript, I interrupted Gabriela and drove her back to the text and asked her to connect her experience to the book. When students in Group 2 shared personal connections, they often seemed to me tangential and in need of tying back to the text.

However, for the students, sharing these stories provided an enjoyable way of connecting the books back to their lives.

Gabriela: Can I tell something about the tarantula?

Researcher: What?

Gabriela: Because my cousin when I came yesterday... Yesterday, I saw his bedroom because I always do my homework there, and I found a tarantula. Because he adopted one. And, it's so scary.

Researcher: Some people do keep them as pets, yes. The guy who took the pictures kept some as pets.

Sarah: One time I went to the zoo, and-

Researcher: You've seen them?

Sarah: Yeah, and like a girl was like, Want to touch it? And, I'm like, No, thanks.

This discussion of the difficult book *Spiders* allowed the students to contextualize the reading to their experiences. I interrupted Sarah, fearful of the group becoming derailed on a long story about the zoo, but their prior experiences may have served to make the book more interesting and relevant to them. The data contain many examples of Group 2 telling stories of various degrees of relevance, but all in some way inspired by the text.

In contrast, the students in Group 1 made their personal connections much more briefly. For example, while reading the matched book *Pop! A Book About Bubbles*, Rosa shared, "I have a big bubbles, a can, and if we waste them ...where it comes with a stick, we put some more." In the same discussion, students read about how some homemade bubble recipes include cornstarch, and Jack and Rosa each shared how their families used cornstarch at home, but both contributions consisted of only one line. When Group 1 read

Acorn to Oak Tree, Alyssa and Rosa had questions about one of the tools used to make things from the oak tree, and Jack explained that his dad had one: "It's this thing that looks like a knife, but it isn't. But, the thing is flat (sliding hands along each other). It's to smooth wood."

The differences between the groups suggest that the students reading closer to grade level (Group 1) knew better how to make personal connections tightly connected to the text. They also adhered better to the norms of school, that teachers usually prefer students to stay on topic and avoid long, personal stories while discussing text. Group 2 however found great value in their stories, and they did move discussions forward despite their lack of apparent focus on text. The results point to the need for more research on the role of personal stories in supporting comprehension in text-based discussion.

Gesture. Students used gesture in their talk about text to define/exemplify terms; estimate measurements; show how things work; and to support, extend, or take the place of words. Sometimes their gestures reflected common childhood behaviors, and other instances likely arose from their background as bilingual people who frequently used a range of linguistic and nonlinguistic tools to make their points.

For example, in discussing the matched text *Acorn to Oak Tree,* Alyssa and Jack used gesture to show 20 inches, the amount an oak tree can grow in one year. Alyssa stretched her hands out to show her estimate of 20 inches, and Jack moved his fingers slowly across the table in one-inch increments. Other times, students simply talked with their hands as people often do. For example, in discussing the difficult book *Looking Closely in the Rain Forest,* Alyssa explained the meaning of of squirrel monkeys "chattering" by opening and closing her hands a few times and saying, "They talk a lot."

As bilingual people, students used gestures to clarify, extend, and sometimes substitute for words. For example, students often used gestures to define words. When asked what fins are, Sarah put her hands to her sides and laughed as she flopped them around as if she were a fish with fins. Gabriela explained the word "leaping" as "doing the dolphins, like that" while she moved her hand in a wave pattern, indicating a fish leaping in and out of the water. Later, she explained to Sarah and Elise what a den was by picking up the book and tenting it to create a den. Sometimes, students compensated for a minimal answer by adding a gesture. While reading *Pop!*, students talked about other ways besides blowing to make bubbles, and Jack simply said, "Running," but he held his hand up as if holding a bubble wand and running into the wind as discussed in the text.

Other times, students used gesture to demonstrate comprehension and participate silently. As they discussed what it meant to hinge and unhinge (because a snake in the matched book *What Do You Do With a Tail Like This?* could unhinge its jaws to swallow large prey), Rosa opened and closed her mouth widely to show hinging, but she never said anything. In fact, roughly one-fourth of instances of gesture referred to silent participation like Rosa's. Gesture silently taking the place of words occurred evenly across matched and difficult texts, but in general the students in Group 1 used gesture twice as often with matched texts. The data do not suggest a reason why. Students in Group 2 used gesture evenly with both levels of text.

Fix-up. I applied the code fix-up when students worked to troubleshoot an unknown word or received my support to decode an unknown word. Surprisingly, instances of fixing-up occurred evenly for both groups across matched and difficult texts. However, they occurred more with Group 1 who read more advanced texts. Students did

not exhibit a variety of fix-up strategies for figuring out unknown words, and I did not employ many strategies for supporting them with decoding either. Most commonly, students guessed words or stumbled over them. Usually when they faltered, I supplied the unknown word to keep the group momentum going. For example, in reading from the difficult book *Fossils*, Jack read a caption commenting that "there's dinosaurs, tyrannosaurus rex, stegosaurus, tri-, tri-" before I gave him "triceratops." As discussed previously, when students employed labored reading with many miscues, they lost the attention of the group. The data reflect that rather than focus on ways to support students with fix-up strategies, I chose to keep the group moving by providing the word so the discussion and reading could progress.

In one exception, I helped Gabriela and Sarah figure out the word "startle." While reading the difficult text *What Do You Do When Something Wants to Eat You?*, Gabriela did not understand how having a blue tongue helped the blue-tongued skink survive.

> Gabriela: But, how does it protect it...by licking it? (sounding incredulous)
>
> Researcher: That is a good question. How does having a blue tongue keep it safe? / We might have to go back to the page before and see what it said about the blue tongue.
>
> Sarah: It said, "the blue tongue stick (Researcher: The blue-tongued skink...that's the name of the animal) sticks…"
>
> Researcher: Okay, this word that Sarah's on, this is the key. The blue-tongued skink does what...?
>
> Gabriela: It stings it.
>
> Researcher: It doesn't say sting. Double check what word that makes. Do you

recognize a part of it? / Cover up the -les, and what word do you have?

Sarah: Start.

Researcher: Start. And, then look at the -les now. Start-

Sarah: -les.

Researcher: Startles.

In this interaction, the students returned to the text because I encouraged them to so they could answer their own question, but they had trouble reading it. When Sarah miscued "skink" as "stick," I simply gave her the word as I often did. (The text read: "The blue-tongued skink sticks out its tongue and startles…") Then, Gabriela guessed "stings," reflecting the students' common tendency to guess words they could not read. Somewhat uncharacteristically, I provided prompting to look at word parts as a decoding strategy. Even though I did not offer exemplary decoding support in this instance (because I put the word parts together for Sarah after she read them), I provided more decoding help than usual. Because multiple students focused on the word and it played a key role in a meaning-oriented discussion, I showed more willingness to help them figure it out than when one student read text in a labored way while others drifted off.

Textual connection. Students made textual connections when they connected the texts to other texts, including multimodal texts like television shows or movies. In fact, students exclusively connected the texts to television and movies without prompting; they made connections to other written texts only when I prompted them.

In discussing the matched book *What Do You Do With a Tail Like This?*, Rosa shared a connection about eggs that an animal buries in a hole that she learned about on *Wild Kratts* on PBS Kids. Sarah also made connections to movies while discussing the

matched books *About Birds* (which she related to *Happy Feet*) and *About Fish* (which she connected to *Finding Nemo*).

Connections to other texts occurred when I asked students to make them. For example, while reading the matched book *Dandelions*, students became interested in how the wind carries dandelion seeds on tufts and what would eventually cause them to fall down and perhaps sprout. Students worked together to develop the idea that the seeds would come down when the wind stopped, and then I asked them what we had read about in a different book that would make the seeds come down. Gabriela excitedly responded, "Gravity!," the title of a book we had read several days earlier.

On another occasion, Jack became confused by the pictures on the back of the difficult book *Animal Hair*.

> Jack: I forgot. What's taller? What's bigger? What's the biggest bear? The grizzly bear or the polar bear?
>
> Researcher: Did we read... Are the polar bears the largest bear? (getting polar bear book from previous day) Did we read that before?
>
> Alyssa: I think so.
>
> Researcher: "The polar bear is the biggest bear in the world" (from polar bears book). So, you're right. That's a really good question. So, look at this picture, and notice what he's saying. He's saying the grizzly bear looks bigger than the polar bear. Why do you think?

Jack almost made this textual connection by himself because the difference between the drawing in *Animal Hair* and the text in *Polar Bears* produced the cognitive dissonance

leading to this question. However, I made the discrepancy explicit and returned to the text in the earlier book.

Students in Group 1 made textual connections more often with difficult texts (13 more times), and students in Group 2 made them more with matched texts (9 more times). Because I prompted so many of these connections, the differences may reflect more about which books happened to relate to others than they suggest a relationship between text difficulty and making connections.

Background knowledge. Students used background knowledge when they stated a relevant fact or previous learning connected to the text. Personal connections (reported above) also represent a kind of background knowledge, but I distinguished the two in this study by considering personal connections as deriving from students' experiences and often resulting in narratives shared in discussions while considering background knowledge to represent factual information that students briefly reported. In the following transcript from the difficult book *Seed, Soil, Sun,* Jack shared background knowledge about cows as ruminants.

Jack: Oh! I know. I know something about cows.

Researcher: What do you know about cows?

Jack: Something nasty.

Researcher: Well, do you want to tell us or not?

Jack: That they eat their food, they swallow their food, and then they swallow it back up (gesturing hand coming up along neck), and they eat.

Alyssa responded to this information with a "Yuck!" and the discussion moved on. Jack often had tidbits of background knowledge like this one that he shared with the group

while we read and discussed. On another occasion, Jack and Alyssa used their background knowledge to object to the name of "moth orchid" that they learned about in the difficult book, *Looking Closely in the Rain Forest*. They complained that moths are gray, so really the orchid should be a butterfly orchid since it had bright colors. They reported their background knowledge about the insects' colors, when they are active, and why they have the colors that they do to substantiate their claim.

Students used background knowledge evenly across difficult and matched texts in both Groups 1 and 2. Because I considered personal and textual connections a different category (and thus not part of the code background knowledge), background knowledge—when defined as reporting of factual information—did not figure prominently as a strategy for students.

Genre use. Attention to genre as a strategy for making meaning of the text occurred exclusively in Group 2 as students marveled at features of the natural world and wondered if they could possibly be "real." Gabriela looked at a picture in *About Birds* and asked, "Do these birds really lay blue eggs?" While talking about *What Do You Do When Something Wants to Eat You?*, Elise asked, "In real life, is there like a real fish that flies?" Elise, a child growing up in an urban desert, also wanted to know while reading *I Am a Frog* if frogs "really" sit on lily pads. In each of these discussions, I asked them to think about genre, and they always admitted the books were "real" or nonfiction. Group 2 used genre to remind each other that they read nonfiction texts evenly across difficult and matched books.

Cognates. Students used cognates when they connected an unknown English word to a Spanish word. I prompted all use of cognates; the data have no examples of

students relying on cognates spontaneously. I once prompted the students in Group 1 to think about cognates when they read the difficult book *Fossils* and learned that scientists use dental picks to clean and excavate fragile fossils. The students did not know what a dental pick was, and I asked them to think about the Spanish word for tooth (diente) to help them see the connections between diente, dentist, and dental.

> Researcher: Yeah, they're using picks. It says they're using dental picks. Does dental sound like a word that you know?
>
> (Jack and Rosa shake heads no.)
>
> Researcher: What about dentists?
>
> (Alyssa nods yes.)
>
> Researcher: Or in Spanish, how do you say tooth?
>
> Researcher & Alyssa: Diente.
>
> Researcher: Diente, dentist, dental. So a dental pick...have you been to the dentist before when they put that teeny tiny little pick in your mouth and they clean off your teeth?

I did most of the talking in this episode, and the transcript does not suggest that students took up cognates as a way to help them understand. However, in a prompted discussion in Group 2 about a matched book, students participated more fully in a word-conscious discussion where they learned about a cognate, made personal connections, and explored multiple meanings of a word.

> Researcher: To go to Sarah's question about the katydid, it says in the back, "Katydids are often big and have lots of spines, so they are hard to swallow." What are spines?

Sarah: Like, they're like, have pointy stuff.

Researcher: Yeah. Does spine... it kind of sounds like a Spanish word... espinas?

Elise & Sarah: Yeah.

Sarah: Cause one time, I was playing with my friend, and espinas, spines, got in my spine, and they were like stuck like this, and my friend had to help me take it off.

Gabriela: Yeah, my dog also gets them.

Sarah: Mostly her hair was full of spines.

Researcher: In English we have two meanings for spines. One is these little espinas that you see right here. But, also a spine is ...you all have a spine down your back. Right down your back, this is your spine.

Gabriela: That's why we can move our back.

Researcher: Yeah. And, books have a spine too. This (picking up to show) is called the spine of the book. So, that's lots of meanings for the word spine.

In this episode, Sarah used both the English and Spanish words together, and the group continued to talk about their experiences with spines for several turns. Sarah's initial contribution used "spines" in two senses (getting spines on her spine), which led to a discussion of the multiple meanings of the word.

However, these two discussions (of diente/dental and espinas/spines) represent the only discussions of cognates across all the groups. So, while cognates may represent a potentially powerful tool for bilingual students to comprehend informational texts, these students never used them independently, and I had mixed success on the occasions that I

focused their attention on them. Because cognates only became salient when I introduced them, students' using them did not have any relation to text difficulty.

Summary. Students did not summarize with the rare exception of Rosa, who summarized only twice. I earlier described how Rosa seemed more timid by nature and often used reading text without comment as a way to participate in discussions. Once when she read and then fell silent, I asked her to elaborate on why that text struck her, and she produced a perfect summary of what she had just read, but that did not drive the discussion forward. Another time, instead of reading she picked a page to talk about and summarized it. Both of these instances occurred with difficult texts.

Students tended to talk about small portions of text, specific pictures, and ideas that led them to questions and connections. Their ways of engaging with the books did not lead them to summarize, and I did not emphasize summarizing as a strategy for comprehending and thinking about texts.

Strategies summary. Overall, the strategies reported in this section did not differ according to text difficulty. (A difference associated with difficulty did exist for both inferring and asking questions, discussed in subsequent sections.) Students applied personal connections, fix-up strategies, background knowledge, genre use, and summarizing evenly across both matched and difficult texts. The differences noted here reflect differences inherent in the texts, such as some texts having features that made research possible and others lacking those features. Other differences had to do with which texts related to others or when I brought up a cognate. The students' strategy use did not suggest a response to text difficulty.

Interaction: In what ways do students interact to participate in discussions? I analyzed how students interacted with each other and me because I considered interaction a key component of engagement, and I wanted to understand how discussion dynamics changed with more difficult text. Students interacted in these ways: answering the teacher (me), building on previous ideas, asking questions, introducing new topics, answering their own and each other's questions, disagreeing and agreeing with each other, and providing peer coaching. While text difficulty did impact how frequently students made each of these moves, students applied a few of them consistently. In all the groups, the most common move was to answer the teacher, and the least common move was to provide peer coaching. In most groups, asking questions and building on previous comments figured prominently, but students seldom bothered to agree or disagree with each other aloud.

As part of my flexible plan for each group, I prepared literal and inferential questions for the group to discuss, but each discussion began with the topics students wanted to address. During independent reading, they marked what they wanted to talk about in the book, and we began with these comments and questions. If the students exhausted all the ideas they wanted to talk about, then I turned to my questions. When students shared their own comments, I often directed them to the text or asked them questions to encourage them to elaborate. Thus, many interactions involved a student answering me. When students introduced what they wanted to talk about (often as a question), other students would take up the issue, and so asking questions, building on a previous topic, introducing new ideas, and answering peer/self all occurred frequently. Other than insisting on basic school norms of turn-taking and listening to other speakers,

I did not teach or emphasize any interaction strategies, which may explain why students infrequently agreed, disagreed, or provided peer-coaching.

For both groups, I coded more interactions with matched texts than with difficult texts. At first, this result came as a surprise since discussions with matched texts and difficult texts took the same amount of time. Since students did not obviously withdraw and fail to participate with one type of text, the number of interactions should be roughly the same. However, upon reviewing the data I found that I spoke more (and sometimes longer) with difficult text, and I did not code my own turns since I focused the analysis on student interactions. My "extra" turns with difficult text usually took two forms. First, I often explained or elaborated an idea that a student brought up from the text. And, secondly, I frequently followed-up on student contributions with comments like "So, what is that called?" "Why was that interesting?" or "Tell us more about that." I did these things with matched texts too, but the data show that I did them more with difficult text. Thus, I unconsciously took more of the speaking turns with difficult texts, leading students to interact more with matched texts, including to answer me (the most common move in all groups) more with matched texts.

Text difficulty impacted the groups differently. It did not have an effect on how often students agreed, disagreed, or provided peer-coaching. Every other interaction differed for both groups (but not always in the same way) according to text difficulty. Both groups had more instances of answering me and answering themselves with matched texts. However, Group 1 asked more questions with matched texts, and Group 2 asked more with difficult texts. Additionally, Group 1 introduced new ideas and built on each other's ideas more often with difficult texts, but Group 2 did the exact opposite by

introducing and building ideas more with matched texts. Tables 7 and 8 reflect these differences.

Table 7
Most Common Interactions

	Answer teacher (D)	Answer teacher (M)	Build(D)	Build (M)	Ask (D)	Ask (M)	Intro (D)	Intro (M)
Alyssa	70*	89*	30*	20*	8*	20*	14*	7*
Jack	60	62	37*	26*	10*	18*	13	10
Rosa	7*	14*	14*	5*	0	1	11	9
Group 1 totals	137*	165*	81*	51*	18*	39*	38*	26*
Elise	11	8	7	4	9	8	1	4
Gabriela	46*	72*	14*	42*	21*	28*	2	4
Sarah	51*	64*	27*	38*	21*	28*	2	4
Group 2 totals	108*	144*	48*	84*	62*	54*	4*	12*

Note. The table shows the number of times each student implemented an interaction out loud in the discussions. Columns marked with (M) refer to discussions about matched texts, and columns marked with (D) refer to discussions about difficult text. Numbers marked with a * indicate a difference of greater than 6 instances of interaction between the matched texts and difficult texts.

Table 8
Less Common Interactions

	Answer peer/self (D)	Answer peer/self (M)	Disagree (D)	Disagree (M)	Agree (D)	Agree (M)	Coach (D)	Coach (M)
Alyssa	2*	9*	1	0	0	2	0	1
Jack	2*	11*	0	3	1	2	1	1
Rosa	0	0	1	0	0	0	0	0
Group 1 totals	4*	20*	2	3	1	4	1	2
Elise	6	5	0	0	2	0	1	0
Gabriela	4	8	3	1	1	1	0	0
Sarah	8	11	4	2	1	3	0	0
Group 2 totals	18*	24*	7	3	4	4	1	0

Note. The table shows the number of times each student implemented an interaction out loud in the discussions. Columns marked with (M) refer to discussions about matched texts, and columns marked with (D) refer to discussions about difficult text. Numbers marked with a * indicate a difference of greater than 6 instances of interaction between the matched texts and difficult texts.

Diverging patterns: Group 1 (asking with matched and building/introducing with difficult). Across the twelve days, Group 1 asked over 20 more questions with matched texts than with difficult texts. In the interaction that follows, Jack and Alyssa engaged the ideas and pictures in the matched book *Antarctica* through a rapid succession of questions. This page showed skuas (a predatory bird) raiding a penguin rookery for its eggs.

Alyssa: Those are a kind of birds?

Researcher: Those are a kind of birds, right? So, what did you mark about this page? What's interesting on it?

Alyssa: Are the eggs already like, are they like already big? Or, are they still like, squishy?

Researcher: When a bird lays an egg, does the egg get any bigger? Or, does the egg stay the same size the whole time?

Jack: Same size.

Researcher: The egg stays the same size. What's getting bigger?

Alyssa: The baby.

Researcher: The baby inside the egg. So, the actual egg stays the same size, but on the inside, the baby's developing. So, it's already big.

Jack: (interrupting, waving arms) I got a question.

Researcher: Yeah?

Jack: Why does a bird eat another bird?

Researcher: Hmm.

Alyssa: Penguins are birds?

Researcher: Penguins are birds. What makes them birds?

Alyssa: A chicken is a bird.

Researcher: They have (flapping, mimicking Alyssa)... What is this?

Alyssa: Uhh, wings.

Researcher: What else makes them a bird?

Alyssa: Feathers. Beaks.

Researcher: Yeah. Why does a bird eat another bird? I guess it gets the nutrients it needs.

Jack: (waving hands) I have a question! How is its mouth that wide? (looking at

picture of skua's open mouth)

Alyssa: I was going to ask something. What was it? Oh, do they ...when they're like in the egg, do they already have fur?

Jack: No.

Researcher: Do you remember when we read *Penguin Chick*? Is this the one where it has the diagram...and it shows...

Alyssa: Yeah.

Researcher: So, when they first come out of their egg, what do they have?

Jack: Little bit.

This discussion illustrates the ways that matched texts generated interacting around questions. Students rapidly asked questions and responded to each other and me as we engaged their wonderings.

When Group 1 read difficult texts, they asked fewer questions, but they more commonly introduced new ideas and built on them. When they had matched texts, they stayed on one topic for longer often asking many questions about it, and thus they introduced fewer ideas. In the transcript that follows, students discussed the difficult book *What If You Had Animal Hair?* Rosa introduced the idea about the Arctic fox having white hair in the winter, and students built up an idea about how the animal's hair changes with the seasons as they referred to the picture. It showed one fox, but half of the illustration showed its white hair against a winter landscape while the other half of its body showed its brown hair against a summer background.

Rosa: It says that an Arctic fox hair is snow white in the winter.

Researcher: Okay. What about that? What did you guys think about that? Alyssa marked the same page.

Jack: (raising hand) I know! That the hair from this side is brown, and (gesturing) that side is white.

Researcher: Why is that? Why did the illustrator make it that way?

(Jack gasps as if about to answer.)

Alyssa: It's because she got the the fur from this, and then she got her regular hair. (referring to person in illustration) No, it's because its fur...

Jack: (waving hand) I know. I know. It's because in the winter, Arctic foxes are white, and in the spring or summer, it's brown.

Researcher: How can you tell that from this page?

Jack: Cause right there (pointing to book), it has trees that have leaves, and the other one doesn't.

Researcher: Yeah. That's a good thing to notice about the background, how each side is different. What about the pictures of the fox on this page?

Alyssa: This one's white and black. And this one matches with that one. I was going to say because her hair changes every season.

This interaction exemplifies how students introduced and built up new ideas with difficult texts. In this discussion, the students did not ask any questions, but they elaborated their thinking before moving on to a different idea.

Why did Group 1 ask questions more with matched texts and introduce and build ideas more with difficult texts? Any answer to this question can only apply to Group 1

since Group 2 had the exact opposite interaction pattern. These transcripts suggest that longer conversations on one topic, which happened more often with matched texts for Group 1, created more space for questions. Students in Group 1 asked questions as a way of having a conversation, not necessarily to clarify confusion. They asked conversational questions like "Is it bigger than this?" (while stretching out arms) or "Is this table wood?" (while reading *Acorn to Oak Tree*). In contrast, when Group 1 talked about difficult books, they introduced and built ideas quickly before moving to new ones; they did not often develop extended discourse from one idea, creating space to ask questions about it. The transcript above from *Animal Hair* is not long, but it is one of the longest instances in the data of the students interacting about the same idea from difficult texts.

Diverging patterns: Group 2 (asking with difficult and building and introducing with matched). Group 2 defied the patterns established in Group 1. They asked questions more with difficult texts, and they introduced and built ideas more with matched texts. They asked questions to clarify concepts they did not understand in difficult text. In the difficult book *What Do You Do When Something Wants to Eat You?*, they learned about how the glass snake, when attacked by a predator, can release its tail to escape and later grows a new one. The students labored to wrap their minds around this fascinating concept and then to explore related pressing issues such as what if it were attacked on the head instead of the tail and how often it can protect itself in this way before it runs out of tail.

Elise: How did it... Did they actually cut the tail?

Researcher: So, what does it say happens to the tail? You might need to read that text to see.

Elise: So, it breaks in half?

Researcher: Yeah, it does. What does it say it does? The tail breaks into ...

Elise: “small...

Researcher: “small wriggling pieces.” What does wriggling mean? ///

Gabriela: It means like, broke in very, very little pieces.

Researcher: That’s what small means, but what does it mean that they’re wriggling? Do you know the word wiggling? Wriggling is like wiggling.

Sarah: That they’re like wiggling. That it keeps on wiggling.

Researcher: Yeah, but what is wiggling? What’s it doing?

Gabriela: (gasp) It’s not dead.

Sarah: It’s dead?

Gabriela: No, it’s not dead. It just keeps going.

Researcher: When it wiggles, wiggling means it moves. So, if it breaks, it’s just moving a lot. (Shows hand flopping around on table) So, some kind of creature is on its tail, and its tail, when that happens, is going to break. And, these pieces are going to do what?

(Elise puts hand out and mimics researcher’s flopping hand from earlier.)

Researcher: Flop around, right? And, wiggle and wriggle. What will happen to the animal that’s trying to kill it when it sees its tail doing that?

Sarah: He’s going to go away.

Gabriela: Eat it.

Researcher: Why do you think it might go away? //

Sarah: Because, // it broke in half. Because it broke in half.

Researcher: Because it broke in half. It might be confused. Or, maybe it might... When you said eat it, Gabriela, what part were you thinking it would eat?

(Gabriela points to broken tail pieces.)

Researcher: Yeah, it might eat part of this, but it won't eat the snake. The snake will still escape. You marked this page too. What were you thinking, Sarah?

Sarah: That like... I was thinking the same thing with Elise because how ...How is going to break? Does it get a new tail?

Researcher: Like, grow a new tail?

Sarah: Yeah.

Researcher: Yeah.

Gabriela: Those (pointing to snake that lost its tail) survive?

Researcher: Mmhm. This part of the animal survives, and that part breaks off. This is how it protects itself when it's being attacked.

Gabriela: What happens when they step on its head?

Researcher: (repeats question) I don't know. It has to break part of the tail.

Sarah: How 'bout if another predator sees it, and then it has its tail and all that stuff, and he just capture it and he breaks it, and how about like if it breaks again?

Researcher: If it breaks... like right after this (pointing to picture of broke tail) if another animal attacks it? He probably can't do it too often, right? He can probably just do it one time and wait for it to grow back. But, if it happens again right away, he's probably...

Sarah: Going to eat it?

Researcher: I guess so.

Gabriela: Or dead.

The students collaborated to understand the text ("When it is grabbed by the tail...its tail breaks into many small, wriggling pieces.") and the illustration. Their questions arose from a combination of fascination and confusion, and they represented a way of trying to understand the text. These questions differed from the conversational questions of Group 1.

In a discussion of a matched text, *About Fish*, Group 2 introduced and developed ideas related to the pufferfish. Elise knew they puffed up to protect themselves, so she found it strange that one in the illustration was puffed up and the other right by it was not. If one was in danger, they both were in danger and should both protect themselves.

Elise: It's so weird because this one has spikes, and this one's not.

Researcher: Yeah. Why do you think that is?

Sarah: (gasps, raises hand) //

Researcher: What do you think?

Gabriela: Maybe because its spikes are camouflaging it.

Researcher: Okay, so right here they're blending in. Is this the same kind of fish? (tapping both fish on the page)

Gabriela & Elise: Yes.

Researcher: Yeah. So, why this this one puffed up, and this one's not? What do you think?

Sarah: Because there's the... It's the ... That means that there's danger.

Researcher: Okay. So, do you think that this one maybe sees some danger that this

one doesn’t see?

Gabriela: Yes.

Researcher: Could be. If this one maybe sees it in a minute, what will he do?

Elise & Sarah: He’ll puff up. (showing with hands)

Researcher: He’ll puff up too. But, Elise, you think it’s strange that they're so close to each other and only one of them’s puffing. Yeah. Good observation.

Gabriela: (pointing to book) Maybe that’s their enemy.

Researcher: Maybe. And, you think he hasn’t seen yet?

Gabriela: No.

This transcript shows Group 2 introducing and building on ideas about a matched text that they understood. They did not ask questions, but they used their background knowledge and prior learning from the text to develop an idea about why one fish puffed when the other one did not.

When Group 2 asked questions with matched books, they did not arise from not understanding the text. They asked questions about pictures (what is that?), that arose from childhood curiosity (This reminded me...do bees blink?), and that arose from confusion about genre (wait...is this real?). They asked these questions about difficult text too, but difficult text also generated the kinds of questions reflected in the transcript above: questions that betrayed that they did not understand the text and that they were puzzling through complicated ideas. For Group 2 then, difficult texts generated more questioning, and matched texts provided an opportunity to use their understanding of the texts to introduce and build on ideas in a discussion.

Discussion

To analyze student discussion beyond engagement, I collected data about the ways students responded in discussion and how often they participated.

Response. Student responses (contributions to the discussion) fell into the categories of literal or interpretive/inferential. Literal responses involved restating a fact or idea from the text, asking a "right there" question answered by the text, or asking a basic factual question even if the text did not contain the answer. Interpretive and inferential responses included comments and questions that extended beyond the information provided by the text. Some student responses—both inferential and literal—also received the code *incorrect* if the student made a statement or claim that either contradicted or lacked support from the book. Incorrect responses occurred infrequently (ranging from, on average, a low of 1.5 per discussion in group 2's matched discussions to a high of 3.1 per discussion in Group 1's matched discussions). Incorrect responses appeared evenly across matched and difficult discussions for both groups. By far, both groups had more literal responses than interpretive or inferential ones. Both groups also had more literal responses in discussing matched books than difficult ones, but the difference was not great. It likely reflects the reality discussed earlier that students had more speaking interactions with matched books, so it makes sense that their most common response occurred more with the books where they spoke more. The differences between interpretive responses were not enormous either: students made them about 20 times more in one condition than the other, which works out to an average of three more interpretive comments or questions in each discussion. Group 1 made interpretive

responses more with matched texts, and Group 2 did the opposite and inferred and interpreted more with difficult text. Table 9 summarizes the responses in each group.

Table 9
Types of Responses

	Incorrect (D)	Incorrect (M)	Interpretive (D)	Interpretive (M)	Literal (D)	Literal (M)
Alyssa	10	13	24*	30*	60*	74*
Jack	5	6	21*	28*	67	68
Rosa	0	0	1	6	24*	19*
Group 1 totals	15	19	46*	64*	151*	161*
Elise	0	0	6	10	21*	14*
Gabriela	6	7	29*	21*	38*	62*
Sarah	6	2	38*	23*	60	59
Group 2 totals	12	9	73*	54*	119*	135*

Note. The table shows the number of types of responses for each student. Columns marked with (M) refer to discussions about matched texts, and columns marked with (D) refer to discussions about difficult text. Numbers marked with a * indicate a difference of greater than 6 instances of response between the matched texts and difficult texts.

Group 1: More interpretive/inferential responses with matched texts. Group 1 produced the same kind of interpretive comments with matched and difficult texts. For example, in discussing the difficult book *What If You Had Animal Hair*, Jack and Alyssa inferred that some animals need double coats because they live in cold climates, and those in warm habitats need less fur.

Jack: But, this one lives in the hot.

Researcher: It lives in a hotter place, so maybe it doesn't need it, you're saying?

Jack & Alyssa: Yeah.

Researcher: How do the double coats help the animals?

Alyssa: So they can be warm.

Rosa: Warm.

Jack: It's because the habitat they live in.

Researcher: Mmhm. The habitat they live in is what?

Jack: Cold.

This inference unfolded over several lines and involved all the students. Other times, students simply made one-line inferences and interpretations. In discussing the difficult book *Looking Closely in the Rain Forest*, after reading that squirrel monkeys chatter in the trees to each other, Alyssa said, "I wonder how they understand each other." While discussing the matched book *Antarctica* about people coming to drill in the Arctic, Jack noticed the book did not explain why people wanted to drill for oil in the first place, and he asked a question that moved beyond the literal when he wondered aloud, "How is oil popular?" Group 1 made inferential and interpretive comments and questions in similar ways whether discussing matched or difficult books, but they made them more with matched books. They may have found it more challenging to access difficult books beyond a literal level of response.

Group 2: More interpretive/inferential responses with difficult texts. Group 2 inferred and interpreted more with difficult text. For example, the difficult book *Spiders* prompted students to ask and answer many inferential questions. They wanted to know if

a person could run a finger through a spider web and make the spider fall. This question prompted students to return to the text and read about spider silk, share their own stories about spiders in their cars, and ultimately infer that the "silk safety lines" that the author described as "stronger than steel" and "stretchy" would make it difficult for a person to knock a spider off its line. Students also engaged in interpretive talk when they worked together to define dens and speculate about why a spider would need a den (for warmth, they thought).

Near the end of the discussion, the students became interested in a photo of a spider wrapping up its prey in silk to eat it. They wanted to know both if the spider would eat its own silk when it later ate the prey and what the specific prey was. Elise decided to look up the prey in the photo index while Sarah kept trying to decide if the text suggested the spider would eat its own silk. Elise found that the photo index only contained spiders and not their prey. Sarah inferred that based on the scope of the book, the author and photographer had judged that identifying the spiders mattered more. She explained, "[Spiders are] more important. The book is about spiders" and directed our attention to the cover of the book (with the large title *Spiders* and a photo of one) to make her point. Gabriela reiterated this idea and said, "I think that spiders are more important...Spiders are more important because that's what they're talking about, only spiders, not the prey."

Later, Sarah remembered the question about spiders eating their own webs, and Elise responded that the spider just ate the prey without removing the web. She interpreted that there was no point to wrapping up the prey just to unwrap it: "Because if he's putting the web on that, then why, if he's going to take it off, is he putting the web on that?" The text supported this inference by describing a spider wrapping an insect in

silk, oozing digestive juices onto it, and then sucking the prey into its stomach. When I reread this text aloud, Gabriela and Sarah made slurping noises, and the group agreed with the final inference that spiders do not unwrap their prey before eating it.

Group 2 engaged in inferential talk about matched books too, but less often. For example, while discussing *About Birds* Sarah wanted to know where birds could hide in the event of natural disasters, and Gabriela and Elise speculated possible ideas about this question. Gabriela wanted to know how the photographer in *Red-Eyed Tree Frog* got a particular picture "so perfectly," and that question led the students to suggest strategies the photographer might have used. These examples from matched texts show how Group 2 inferred about pictures or related ideas, but their interpretive talk had less grounding in the text with matched texts. Because Group 2's matched texts fell significantly below grade level, they often had limited text per page and developed fewer ideas over the course of the entire book. Group 2 thus experienced a more dramatic difference between their matched and difficult texts than Group 1 did, and for them the simple, matched texts did not provide much raw material to develop inferences and interpretation about.

Participation. To gauge student participation, I counted the number of times students contributed to the discussions in matched and difficult texts. Table 10 shows these results.

Table 10
Number of Discussion Contributions

	Contributions to Difficult Discussions	Contributions to Matched Discussions	Summary
Alyssa	118	153	35 more in matched
Jack	128	139	11 more in matched
Rosa	33	35	2 more in matched
Group 1 totals	279	327	48 more in matched
Elise	36	28	8 more in difficult
Gabriela	89	137	48 more in matched
Sarah	123	128	5 more in matched
Group 2 totals	248	293	45 more in matched

Most students did not have a big difference between the number of times of they participated in matched or difficult discussions. Sarah and Rosa participated just about the same with each level of text. Jack made 11 more contributions with matched texts, and Elise made 8 more with difficult texts, but these differences work out to less than 2 statements per discussion on average. Alyssa and Gabriela stood out from the others because they both made many more contributions with matched texts. On average, Alyssa made six more contributions to matched discussions than difficult ones, and Gabriela averaged making 8 more contributions to a matched discussion than a difficult discussion.

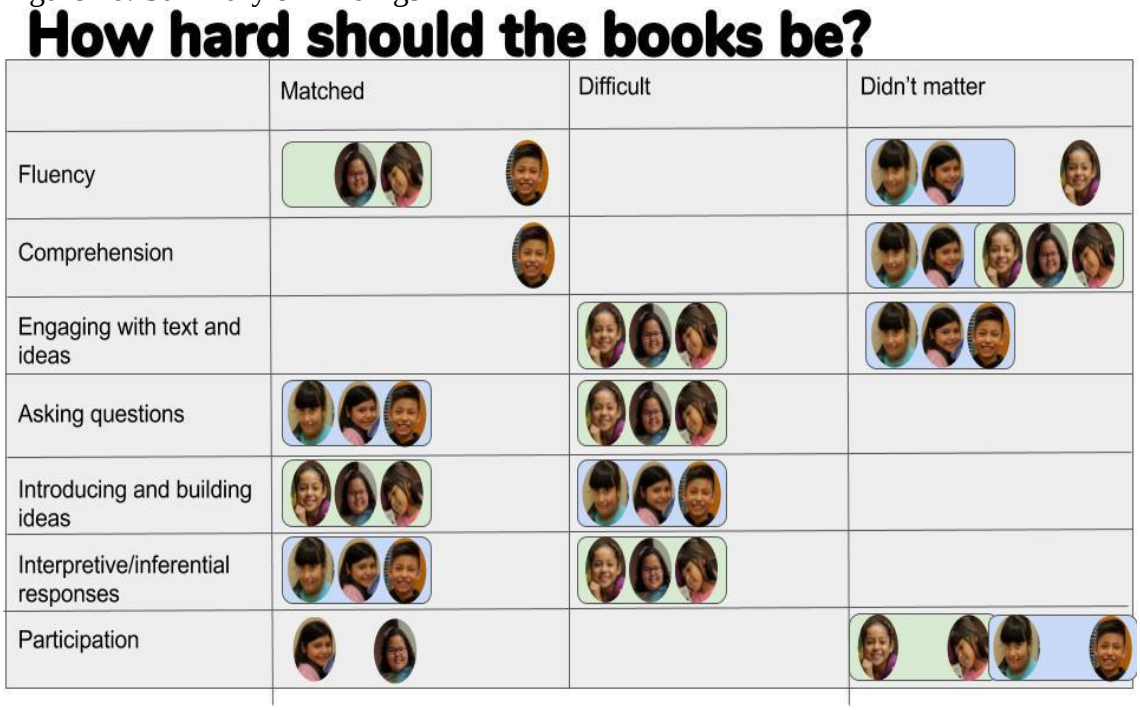
These participation patterns showcase the ways in which students responded differently to text difficulty: for one third of them, it did not matter (Sarah and Rosa); for another third of them, it barely mattered (Jack and Elise); and, for the last third of them,

matched texts encouraged participation much more than difficult texts did (Alyssa and Gabriela). For Alyssa and Gabriela, the accessibility of matched texts facilitated more participation, but for the other students, the social space of the discussion seemed to provide enough support to produce roughly even participation patterns.

Summary of All Findings

In this study, I collected data on students' fluency, comprehension, engagement, and discussion in small-group reading of matched and difficult books. I aimed to address a broad question about small-group reading for bilingual third-grade students reading below grade level: how hard should the books be? The answer, it turns out, is it depends. It depends on the students—their personality and reading proficiency. And, it depends on what outcomes matter. The results for fluency, comprehension, engagement, and discussion did not align. For some outcomes, matched books produced better results, and for some difficult books did. In some cases, Group 1 had results that contradicted Group 2. Figure 20 summarizes the results.

Figure 20. Summary of Findings



Note. The students featured on the blue background were Group 1, and Group 2 has a green background. (Group 2 read further below grade level than Group 1.)

Half the students read more fluently in matched books, and for the other students, the difficulty did not matter. Students of lower reading proficiency benefited from matched books in terms of fluency.

Difficulty did not impact any student's comprehension except Jack's. For Jack, this impact was slight and in favor of matched books.

Group 2 engaged the text and ideas, asked questions, and evidenced inferential thinking more with difficult books. I suspect their matched books did not offer enough content for them to apply these higher-level thinking skills in matched discussions. However, Group 1 (for whom the difference between matched and difficult books was less striking) demonstrated all these higher-level thinking skills more with matched

books. For them, rather than offer access to exciting new content, difficult books presented some barriers that matched books did not.

Most students participated evenly across conditions, but Alyssa and Gabriela both made more contributions when discussing matched texts.

CHAPTER 5

DISCUSSION

In this study, text difficulty had less of a clear effect on students' reading outcomes than expected. When it did have an effect, it varied by student and with each outcome. In summary, the findings included the following:

1. For engagement,
    a. Students engaged with pictures, ideas, and texts. Group 1 engaged at each of these points roughly evenly, but Group 2 engaged much more with texts and ideas in difficult books and with pictures in matched books.
    b. Students used a variety of common comprehension strategies, and these did not vary with text difficulty.
    c. The two groups showed opposite interaction patterns. Group 1 asked questions with matched texts and built up and introduced ideas with difficult texts. In contrast, Group 2 asked more questions with difficult texts, but introduced and built new ideas more with matched texts.
2. For discussion,
    a. Group 1 had more inferential/interpretive responses with matched texts, and Group 2 had more inferential and interpretive responses with difficult texts.
    b. Most students participated evenly regardless of the difficulty of the text under discussion. However, Gabriela and Alyssa talked more when discussing matched texts.

3. For fluency, half of the students benefited from matched texts. The other half (after modeling and discussion) could read difficult texts with similar fluency to how they read matched texts.
4. For comprehension, text difficulty did not matter for anyone except Jack, and for him it only had an effect on 3 of 12 days.

Thus, both matched and difficult texts both had benefits and drawbacks for these students. These differences likely related to reading proficiency, group personality, and serendipity. Group 1 had deeper conversations (with more inferencing and interpreting) when they discussed matched texts. For them, the difficult texts seemed a bit beyond reach to inspire rich conversations. Two of the students in Group 2 benefited from matched texts for fluency, but the group as a whole had higher level conversations when they used difficult text.

The finding that most strikes me in this study is the way that Group 2, those students needing the most support to access grade-level texts, benefited in important ways from talking about difficult books. This difference in engagement across difficulty levels for Group 2 suggests that difficult books supported talk about texts and ideas more than matched books, which contained less text and developed fewer ideas, and thus pictures instead served as the focus of engagement. Perhaps they found the ideas more complex and worth engaging in difficult books when compared with simple matched texts. Interestingly, students in both groups applied the strategy of "research" (or using peritextual material to learn more) with matched texts, perhaps reflecting their eagerness to engage ideas beyond what simple matched texts sometimes represented.

In both groups, the amount of engagement with text was lower than I expected. That finding suggests that these groups of students needed extra support (beyond that offered in this study) on text-based discussion. They needed instruction in how to use the text to drive conversation and in how to respond to someone else using the text, and they needed supportive social spaces that allowed them to practice these skills. They also suggest something interesting about fluency: fluency did not seem to make a big difference for comprehension (as reported with the quantitative findings), but it did make a difference in how well students could use text in discussions because disfluent reading tended to thwart conversation or require a lot of adult support to work through.

Because the two groups developed almost opposite interaction patterns according to text difficulty, it seems likely that both their different reading levels and their group personalities impacted the interactions. Unfortunately, the data do not speak to determining exactly why Group 1 and Group 2 responded differently—whether reading proficiency, group personality, some combination, or something else entirely. However, over time each group developed a personality and "ways of being" during the course of discussion that I described in the findings. Possibly if these same students had been grouped with other students or met a few months earlier or later, different interaction patterns would have developed. At the time of this study and with the children who participated, difficult texts pushed Group 1 to introduce and build on ideas and Group 2 to ask questions while matched texts did just the opposite.

Relation to Existing Research

This study fits well in the existing research on text difficulty because it—like the current literature—does not settle the issue. In this study, students did do well (in some ways) when reading difficult texts, which other researchers have also found (D. Fisher & Frey, 2014b). Allington et al. (2015) criticized other researchers (Morgan et al., 2000; Stahl & Heubach, 2005) for advocating difficult texts because their studies offered students a lot of support, more than what Allington et al. considered available in typical classrooms. But, this study confirms those criticized findings: students can engage with difficult texts when provided with ample support. Yet, in the Morgan et al. study, in many of the areas they evaluated, the researchers did not find significant differences across groups reading different levels of text. Their study, like this one, showed that whether research finds difficult text advantageous depends on the outcomes the researchers choose to evaluate.

Other studies, like this one, failed to find a difference in comprehension outcomes when students read either matched or difficult text (O'Connor et al., 2002, 2010). The first of these studies though, did find a slight advantage for fluency when students read matched texts, a finding that corresponds to the results for half of the students in this study. Hiebert (2005) also found a fluency advantage when students read texts with fewer rare and multisyllabic words, which, while not exactly the definition of "matched," still favors easier texts over difficult ones.

This study contributed something new to the research base by examining engagement and discussion in relation to text difficulty. Previous studies reported qualitative findings about engagement anecdotally (Morgan et al., 2000; Stillman &

Anderson, 2017) or with older students (Feng et al., 2013; D. Fisher & Frey, 2014c; Fulmer et al., 2015; Lyons & Thompson, 2012; Thomason et al., 2016). I could not locate any other studies that specifically and systematically addressed student engagement, interaction, participation, and discussion when talking about texts of different difficulty levels. This study contributes the finding that, with support, students eagerly participated in discussions about both kinds of texts, but their level of thinking and patterns of interactions did vary according to text difficulty as described above. The small-group support in this study did help students avoid the shutting down that other researchers have observed when young bilingual students read complex texts (Stillman & Anderson, 2017).

Limitations

Small Sample

The limitations associated with small sample size that usually apply to single case design and qualitative case studies apply to this study too: I cannot say that the students represented a broader population or that the results reported here would apply to other students in other contexts...or even to these students in a different context or at another time in their educational career. This limitation became especially evident in analyzing the interaction patterns: each group developed its own personality and ways of talking about texts, and student characteristics impacted group dynamics. Because of this problem of generalizability, Phillips (2014) described education as an "extremely hard but softer" (p. 9) domain of research when contrasted with the so-called hard sciences. He described learning as involving real students who "have or have not consumed breakfast...live in neighborhoods with or without frequent gun violence...are attracted by

(or clash with) the personality of their teacher" (p. 10). Thus, he suggested that "just about all the variables are relevant" in education research, making work that zooms in on the educational experiences of students "difficult or impossible to generalize to the other almost infinite number of settings where these variables do indeed vary" (p. 11).

This limitation leaves researchers working with small groups to make what Bassey (1999) called "fuzzy generalization" (p. 52) and Yin (2014) described as analytic generalizations. These conclusions refer to lessons learned from a study, working hypotheses, and principles believed to apply to other situations. However, they imply an invitation for other researchers and practitioners to make judgments about their own contexts to determine the extent to which findings from one study might apply more broadly. The idea of generalization arises largely from a quantitative approach to research, and puts the responsibility on researchers to determine what population their sample represents. However, Hitchcock (2010) suggested instead the idea of transferability in which consumers of research determine if the characteristics of a sample and context of a study make its findings relevant to their needs.

In the same way that the small sample size indicates a limitation of the study, in another sense, it became a strength. I collected a considerable amount of data from each participant, and that data contributed to a detailed portrait of each student. Large-scale quantitative work often loses sight of these particulars by averaging scores across a large number of participants and failing to report much beyond scores.

Narrow Scope

This study's results do not address many questions about text difficulty. For example, this study has no implications for early literacy and the kinds of texts that work

best for beginning readers because the students in the sample had been in school for four years. Schools usually consider third grade the end of early childhood and the beginning of the intermediate elementary grades. Furthermore, instructional emphasis in third grade often shifts from learning to read to reading to learn (Chall, 1983); while this dichotomy is somewhat mythical, the fact remains that students beyond third grade typically experience a shift in the type of reading instruction they receive (Sanacore & Palumbo, 2008). The study narrowly addressed text difficulty for small-group reading: it does not shed light on recommended text difficulty for independent reading, read-alouds, shared reading, partner reading, the books librarians should steer students to for extracurricular reading, or whether anyone should steer students in any of these contexts.

This study also does not suggest anything about students' long-term literacy development after a diet of matched or difficult text for an extended period of time. I premised this study on the assumption that strong daily performance would likely lead to literacy growth, but the study did not test that assumption or extend across a significant enough time frame to measure reading improvement. Comparing students over several months as they work in small groups with different levels of texts merits its own research project. Similar projects could explore the relationships between reading progress and text difficulty levels in other contexts beyond small groups. While little research exists in this area, the studies that we have show mixed findings just like the study here. The field could benefit from more work, but the existing research base suggests that text difficulty may not be so significant for predicting student performance and growth. I find in the end that I agree with Topping et al. (2008) that students can experience growth when reading

from a broad range of levels and that finding the exact right level for a student may not be too important.

This study also did not elucidate a strong instructional protocol for small-group reading with bilingual students. While developing a protocol was not the purpose here, four findings highlight the weaknesses of the protocol that I implemented. First, an analysis of both the retells and the discussion transcripts show that the students did not reach high levels of comprehension of the texts. Their lack of comprehension became apparent in two ways: they did not talk about the texts as a whole, but only focused on little details or facts they found interesting, and they often asked basic questions betraying that they did not understand what they had just read. Second, the students read less than I expected. They relied more on the read-alouds and the discussion than on their independent reading. They often rushed through the independent reading and instead used that time to mark pages they wanted to talk about. The transcripts show that their attention to text happened most when I pushed them into the text, but they did not spend much time reading or talking about the words on their own. Third, the protocol used here gave no explicit attention to language or culture. Science texts are not culturally neutral, and I did not foreground culture in deciding which books to use. I did not draw on the students' strengths as fluent speakers of Spanish either. Using their home language may have better activated their background knowledge and facilitated connections and comprehension, and it might have helped them process Spanish-English cognates, which appear frequently in informational texts (Lubliner & Hiebert, 2011). Finally, the interactions showed that despite my attempts to foster a discussion, our groups still relied

heavily on the initiate-response-evaluate sequence that researchers have criticized for failing to stimulate the kind of dialog that really fosters learning (Cazden, 1988).

The findings also showed that students seldom independently used comprehension or decoding strategies meaningfully, and my lesson plans gave no particular attention to strategies. I did not foreground strategies because I wanted to implement a consistent protocol for each book. Since texts vary in which strategies readers need to marshal to comprehend them well, it would have been difficult to create equivalent discussions of different books while emphasizing comprehension strategies. However, the results obviously may have been different if I had taught strategies. The strategies that best assist striving bilingual readers in making meaning of complex grade-level text certainly deserve research.

Problems with Retells

Finally, I measured comprehension with a retell, and many researchers have recognized the weakness of retells for assessing comprehension, especially for bilingual students. While quantitative work has established a reliable correlation between the number of words and/or ideas that children retell and other standardized measures of comprehension (L. S. Fuchs, Fuchs, & Maxwell, 1988; Reed & Vaughn, 2012), it remains unclear what theoretical view of reading supports assessing comprehension in this way. Notwithstanding, retells have become a common feature of comprehension research, particularly because they work for any text, students can complete them quickly (Reed & Vaughn, 2012), and they rely less on decoding skills than other common comprehension assessments (Keenan et al., 2008).

Yet, retelling challenges students because they have to 1) remember the text, 2) decide what to include, 3) decide how to organize this information, and 4) determine how to say all this information (Goldman, O'Banion Varma, & Sharp, 1999). Obviously, these challenges become even more pressing for students attempting them in a second language in which they are still developing proficiency. Retells do not always predict reading level or growth for bilingual students (Kieffer, 2012; Reed & Vaughn, 2012). And, research has identified the fairly evident conclusion that bilingual students do better on retells after they have had extended time to study the topic of the text and develop the language to talk about it (Faggella-Luby, Griffith, Silva, & Weinburgh, 2016).

Indeed, background knowledge becomes a serious issue in evaluating comprehension (A. C. Miller & Keenan, 2009) and especially for students retelling in a second language since background knowledge includes the oral language needed to talk about the topic. This study did not attempt to "control" for background knowledge; I agree with Burns et al. (2015) that it would be "almost impossible" (p. 443) to do so within the context within which I worked.

Finally, for retellings to give useful information, students need to learn how to do them. One evaluation rubric classifies a retelling as "very cohesive and complete" if it contains "all the main ideas and supporting details, sequences material properly, infers beyond the text, relates text to own life, understands text organization, summarizes, gives opinion of text and justifies it, and may ask additional questions" (Moss, 2004, p. 717). No student in this study, which ultimately included 144 retells, ever came close to this description of quality retelling. Indeed, many of their retells (about informational science text) began with "the story was about…," a beginning that reflects some basic confusion

about how genre should influence retelling. Because of the limited time I had to work with students and remove them from their regular class, I did not take time prior to the study to build their retelling skills to ensure the best possible retellings once the study began. This weakness no doubt compromised an already suspect measure of comprehension.

Classroom Implications

Because this study produced such complicated findings (see Figure 20), it does not lend itself to immediately straightforward classroom implications. This study did not, as I anticipated originally, answer the simple question I brought from my years as a classroom teacher: how hard should the books be during small-group reading? Instead, it showed that this question is not simple. How hard the books should be, it turns out, depends on the student and the purpose for reading.

How Hard for What?

Teachers convene small reading groups for many purposes, and they may struggle to honor all these purposes simultaneously. After the release of the National Reading Panel's report and the implementation of Response to Intervention, reading groups categorized as intervention groups targeting specific skills prevailed in elementary schools and the reading research literature (M. S. Hall & Burns, 2017). But, teachers continued to value sociocultural-inspired practices like literature circles and discussion groups too (Moses, Ogden, & Kelly, 2015). Can a teacher target a student's oral reading fluency while simultaneously encouraging participation in discussion with interpretive responses? Perhaps, but even if the teacher can reconcile the seemingly disparate worldviews that prioritize each of those goals, the texts that best facilitate them may not

be the same. For example, in this study, two of the three students in Group 2 needed matched texts to support their fluency, but they needed difficult texts to inspire interpretive responses.

Lampert (1985) described teachers as dilemma managers who "bring many contradictory aims to each instance of [their] work" (p. 181). This characterization may especially apply in small-group reading instruction. Johnson (2017) noted that teachers have many reasons for having students discuss texts, and they often struggle to balance all those reasons at once. She particularly observed teachers struggle with the tension between comprehension and conversation skills. Teachers had a hard time letting students lead discussions when they felt students missed key details in comprehending texts. They found some students "were not yet skilled in sustaining respectful conversations with peers, particularly when the text under discussion was challenging" (p. 334). Balancing multiple purposes with small-group reading already challenges teachers; finding appropriate texts to facilitate so many purposes at once only complicates the matter further.

In this study, I considered comprehension, fluency, discussion, and engagement. Even with only four outcomes, the results contradicted each other. Yet, small-group reading produces other important outcomes as well. What is the "right" level for promoting conversation skills, love of reading, curiosity, critical thinking, appreciation of author's craft, imagination, writing, critical consciousness, or building content knowledge (Hoffman, 2017)? Hoffman explained that the historical preoccupation with finding "just right" levels in elementary schools may stem from a "level [that] is 'just right' for the

teacher to promote certain kinds of reading strategies" (p. 268), but not necessarily to support these other important goals in a child's literacy development.

This study did not produce evidence to suggest that teachers should use matched or difficult texts all the time for small groups. Rather, the findings indicate that different levels of text support different purposes, and teachers must use their judgement to provide a variety of texts for students across and beyond the school day. For example, perhaps a student like Gabriela should spend some time each day in a small group with matched texts to build her fluency while also participating in a literature circle that exposes her to more challenging texts that cause her to ask questions and engage deeply with ideas.

Because this study, like many other studies of text difficulty for elementary readers, had mixed or understated findings, I suggest that text difficulty is not the critical factor for intermediate elementary readers. Neither matched texts nor difficult texts represent a panacea for students reading below grade level, but both have their place in promoting different aspects of literacy. Hoffman (2017) explained that matching texts to readers matters most when teachers follow formulaic scripts that limit their professional decision-making in teaching reading. However, if the teacher has the professional knowledge and freedom to offer a variety of supports to a student, he suggested that the level of the text matters much less.

How Hard for Whom?

This study also highlighted how students experienced text difficulty differently. Even within the same group, the text one student read the most fluently, another student

of roughly the same reading level who participated in the same discussion would read at only an average rate. The same disparity applied to comprehension.

The helpful concept of "individualized text complexity" (Fitzgerald et al., 2015, p. 37) invites teachers to consider students' background knowledge and interests in connection with the concepts in texts. It also meshes well with the tri-dimensional framework for evaluating text complexity in the Common Core State Standards, in which one of the three areas of analysis is reader and task factors (Council of Chief State School Officers, 2010). In fact, the standards argue that assessing student motivation, knowledge, and experiences under this domain happens best by "teachers employing their professional judgment, experience, and knowledge of their students" (Appendix A, p. 4). Yet, the fact remains that each student has different interests, background knowledge, and motivation. If teachers could perfectly know all of their students in all of these ways, it would complicate, rather than simplify, matching books to readers.

Burns et al. (2015) explained the difficulties associated with leveling authentic books because of the way that successful reading depends on so many individual factors like background knowledge that become difficult to take into account for each student. Indeed, teachers who choose to use authentic books or other texts about the real world will have this "problem": students do already know things. Thus, a text deemed to have a particular level based on linguistic features alone may prove difficult or easy depending on how much the student already knows about the topic, has read similar style texts before, or really cares to read the one in front of them now.

These considerations do not lend themselves to quantifying or rules of thumb about matching texts to readers, but they do suggest that teachers make allowances for

background knowledge and interest when they consider which texts to teach in small groups.

The Small-Group as a Site for Access to Grade-Level Content

I earlier noted the limitations of the protocol in this study, chiefly that it limited students' actual reading as they relied more on modeling and discussion. However, the protocol did provide students reading below grade level and in a second language with opportunities to explore and discuss grade-level text and content. This strength is significant because often students reading below grade level and/or in a second language do not have access to rich and interesting texts with opportunities to discuss them. Students who do not read at grade level often receive workbook activities that drill skills in isolation and provide no reading of connected text (Allington, 2013; Gambrell, Wilson, & Gantt, 1981; Knapp, 1995), and students learning to read in a second language frequently have the same experience (DaSilva Iddings, 2005; Valdés, 2001).

Additionally, many schools track emerging bilingual students into programs that limit their access to grade-level content and focus instead on English. This tracking occurs with devastating consequences for secondary students who find it forecloses them from taking the courses they need to graduate (Lillie et al., 2010; Valdés, 2001). But, it also occurs in elementary schools when students in English as a second language classes miss core content like social studies, science, and math (Jimenez-Silva, Gomez, & Cisneros, 2014; Lillie, Markos, Hornberger, & Baker, 2014; Rios-Aguilar, González-Canche, & Moll, 2010; Umansky & Reardon, 2014) or even at the classroom level when teachers group emerging bilingual students together and provide them lower quality instruction (DaSilva Iddings, 2005).

Wong Fillmore (2014) explained at length how this instructional arrangement deprives students learning English of the texts and experiences they most need to accelerate their second language literacy. She critiqued the common English as a second language focus on "grammatical structures and vocabulary divorced from content" that centralized "decoding skills" while "scant attention is given to reading for understanding or learning" (p. 625). She suggested that "more complex materials are in fact precisely what [students learning English] have needed, and lack of access to such materials is what has prevented them from attaining full proficiency in English to date" (p. 624). She objected to giving second language readers "brief, watered-down oral" versions of texts or "simplified versions limited to simple sentences and high-frequency vocabulary" (p. 624) and protested that these texts "carry so little substance that there is little content to be gained from reading them" (p. 626). Notwithstanding, she concluded that harder texts without supports would be "disastrous" (p. 626)!

This study illuminated several supports that allowed students reading in a second language to work with the grade-level text that Wong Fillmore argued for. When students in small homogenous groups listened to a brief read-aloud that modeled reading the text fluently and participated in a student-driven discussion, they discussed and began to understand grade-level texts. Researchers in English as a second language have long emphasized the importance of comprehensible input—or making language understandable to students (S. K. Baker et al., 2014; Echevarria et al., 2016; Krashen, 1981; National Academies of Sciences Engineering and Medicine, 2017). This study illuminated simple practices already familiar to most elementary teachers that can make grade-level informational texts comprehensible for students reading in a second language.

Simple Texts Have Less to Talk About

Wong Fillmore's (2014) advocacy for grade-level content connects directly to this issue of simple texts. In this study, the difference between matched and difficult texts was less pronounced for Group 1, who benefited from matched texts in terms of having inferential conversations and asking questions. But for Group 2, a pronounced difference existed between the two levels of text. Group 2 needed difficult texts to have inferential discussions, ask questions, and engage with texts and ideas. They struggled to engage text and ideas when their matched books contained little text and few ideas, often only a few words or a sentence to a page. In order to talk about text and ideas, they needed books that had more text and developed more ideas.

This finding corresponds with a core component of close reading that the passage students read has to merit multiple reads and have ideas worth talking about (D. Fisher & Frey, 2012). Similar ideas prevail in instructional conversations (Goldenberg, 1992) and collaborative reasoning (Zhang & Dougherty Stahl, 2011), two small-group reading protocols demonstrated successful with emerging bilingual students. Collaborative reasoning specifically requires a challenging text that raises controversies, has unresolved issues, and leads students to take multiple points of view. Students as young as first grade have better conversations and do more inferring when they talk about substantive texts with some level of ambiguity (Kelly & Moses, 2018).

Classroom teachers might provide a variety of supports to enable students reading below grade level to participate in discussions of difficult texts. Students might first experience the text as a read-aloud, in shared reading, with a partner, through audio, or in

a supported reading group. However, students' ability to independently decode the text should not preclude them from participating in discussions of texts at grade level.

Multimodal Texts Lead to Multimodal Engagement

When teachers use picture books and other multimodal texts (texts that combine multiple modes of presentation like words, photos, diagrams, etc.), they should expect children to engage with all the modes. Like many teachers, I knew less about how images and texts work together than I did about text on its own (Painter, 2013), and I pushed children to focus on text even when their natural points of entry to the discussion came through the visual. Yet, scholars recognize that in today's visual and information-saturated world, proficient reading requires more than decoding and comprehending printed words (Kress & Van Leeuwen, 2006).

Picture books, including the informational ones used in this study, contain both pictures and words that readers use to construct meaning. Both the pictures and the words matter, and readers modify their understanding of the one based on the other (Sipe, 1998). Readers who only attend to the text do not fully understand the book (Arizpe & Styles, 2002; Serafini, 2010; Sipe, 2008). Teachers who plan text-based discussions of multimodal texts should expect engagement around the visual as much as around the printed words.

Other researchers have found, as in this study, that in many reading experiences, children become "more attuned to the illustrator's rather than the author's craft" (Martinez, Roser, & Dooley, 2003, p. 224). Particularly for bilingual students, viewing images serves as an inclusive literacy practice (though not one valued on standardized assessments) that leads to critical thinking, meaning construction, engagement and

motivation, and positive literate identities (Moses, 2013). Images in informational texts serve a variety of purposes for bilingual readers, including helping them understand content, prompting discussions with their peers, pushing them to read the text to get more information, and complementing what they have read in the text already (Moses, 2015). In her study, Moses found that monolingual English-speaking first graders sometimes ignored images in books, but across a year of intense data collection, she never observed a bilingual child pass over the images in informational texts. She found that for these children, the images led to conversations that drew out their background knowledge, extended their vocabulary, and supported their content learning.

The transcripts in this study show that I really wanted students to engage with the text. When they talked about images and ideas, my responses conveyed that I valued these contributions less than their comments that they firmly rooted in the words on the page. Yet, the existing research base suggests that I should not have been surprised by their engagement with pictures, nor should I have discounted it.

Rosenblatt (2013) described the aesthetic response of a reader who "pays attention to—savors—the qualities of the feelings, ideas, situations, scenes, personalities, and emotions that are called forth and participates in the tensions, conflicts, and resolutions of the images, ideas, and scenes as they unfold" (p. 933). While this description sounds more applicable to literary texts, Rosenblatt emphasized that "aesthetic" refers to the stance or response of the reader and not the text itself. She explained that the thoughts the text calls up for the reader or the "'evocation,' and not the text is the object of the reader's 'response' and 'interpretation,' both during and after the reading event" (p. 933). In this study, students discussed the images and the ideas evoked

by the books as a whole more often than the printed words on the page. This practice turns out to be normal and the way even adult readers often interact with books. Rosenblatt explained that when students talk about their "evocations" from the text it serves as "a powerful means of stimulating growth in reading ability and critical acumen" (p. 948). While these evocations can drive attention back to the author's words, they have value for their own sake.

Research Implications

This study highlighted several needs in the reading research literature. Certainly, it did not settle questions around text difficulty. Larger and longer studies that compare students' responses to different levels of text would help teachers theorize better how to match texts and readers to accelerate reading growth. Most of the existing studies of text difficulty do not match texts to readers after assessing the students' reading levels (Morgan et al., 2000; Stahl & Heubach, 2005); instead, they compare giving (for example) second graders second-grade text versus third-grade text regardless of the students' reading levels. Better matching procedures that take into account a child's current reading level as in this study would improve larger studies. However, the studies we have, including this one, do not indicate that finding the ideal text-to-reader match will resolve the challenges some students face with reading, and so more research into this area may produce limited practical value. In this section, I suggest other areas of research that this study indicated are needed.

Measuring Individual and Qualitative Responses to Instruction

This study stands out in the literature on small-group reading because of the way that it highlighted individual students' responses to instruction. Both looking at students

at the individual level and looking at qualitative outcomes as well as quantitative ones represent important contributions of this work.

The existing research makes clear that reading strengths and challenges vary greatly across students (Buly & Valencia, 2002), and thus students who struggle with reading need different emphases in their small-group support. The research base has established that individualized instruction works (Connor et al., 2011, 2013; Reis, Eckert, McCoach, Jacobs, & Coyne, 2008) and students learning English should have access to it (Richards-Tutor, Baker, Gersten, Baker, & Smith, 2016). In fact, when small-group instruction has not worked, researchers speculated that the failure occurred when scripted curricula took the place of an individualized problem-solving approach that targets the unique needs of each student (Gilbert et al., 2013).

In light of this understood need for tailored instruction, researchers have emphasized the need for reading research to highlight how individual students respond. In discussing the strange marriage between alternating treatment designs and socially-oriented views of literacy learning, Neuman and McCormick (1995) explained that this at first odd combination allows researchers to move beyond asking which method works best to asking which methods work best for particular students in given contexts. Paris (2005) also called for analyzing reading instruction research "for individuals, not groups" (p. 199) because every child has a somewhat different zone of proximal development leading to different results depending on the students' academic starting places.

The analysis of individual-level data in this study made important findings clear. For example, text difficulty did not affect comprehension data for any student except Jack. Jack was also the only member of his group for whom text difficulty made a

difference in fluency. Looking at groups as a whole or at averages scores would lead to the conclusion that text difficulty did not matter for Group 1 in terms of comprehension and fluency. But, looking at individual scores shows that conclusion does not apply to Jack. At the classroom level, this understanding might affect which small groups a teacher places Jack in or which books those groups read. It would not serve Jack well to make whole-class decisions based on averages or group responses because he responded differently than his group.

The attention to qualitative responses also added value to this study. The current literature suggests that students' experience and participation provide important complements to quantitative markers of their reading progress (Moses & Kelly, 2017; Turkan & DaSilva Iddings, 2012). If the analysis ended after evaluating fluency and comprehension scores, then the data only suggest that text difficulty does not matter much, but that for some students matched books support fluency better than difficult ones. In other words, the fluency and comprehension data show no particular advantage for difficult books. But, the analysis of the discussion and responses do. Particularly for Group 2, engaging with text and ideas, asking questions, and giving interpretive or inferential responses all rose with difficult books. For Group 1, difficult books gave them the opportunity to introduce and build up new ideas more. None of these advantages appear in the quantitative data, and so the qualitative data add important insights into students' reading experiences and performance in the groups.

Thus, this study played out what the existing research base suggests: individual responses matter, and qualitative outcomes matter. Future reading research, particularly

large-scale projects, should build in case studies of individual students to understand responses at the individual level and to have a fuller picture of the students' experience.

The Relationship Between Fluency and Comprehension in Second Language Reading

I did not have any research questions about the relationship between fluency and comprehension at the start of this study. However, when I put students' fluency and comprehension graphs side by side, I noticed that they did not align in the way I expected. High levels of fluency did not promote corresponding high levels of comprehension nor did low levels impede it. Thus, this study raised questions for future research to explore about how tightly linked fluency and comprehension are for bilingual students.

The National Reading Panel's report identified fluency as a core component of reading instruction (National Institute of Child Health and Human Development, 2000). Since then, classroom teachers have targeted fluency during the literacy block and schools have weighed students' oral reading fluency (measured in timed one-minute probes) in determining whether students need reading intervention (Gersten et al., 2008). Teachers willingly devoted so much instructional energy to fluency because the research base suggested that fluency related to—if not also facilitated—comprehension. Researchers found correlations between fluency and comprehension (Reschly, Busch, Betts, Deno, & Long, 2009). Theoretically, they identified support for the relationship too: students who can read with automaticity free up their mental resources to focus on the meaning of the text (L. S. Fuchs, Fuchs, Hosp, & Jenkins, 2001).

But, some problems with this thinking have become clear. Fuchs et al. (2001), who defended fluency as an indicator of reading competence, made clear that they did not include higher level comprehension skills such as the "capacity to analyze literature" or "learn new information from complicated expository text" (p. 241) as a component of reading competence.

This exclusion of higher order thinking raises the question of what they even considered "reading competence" and what exactly oral reading fluency did correlate with in their studies if not the ability to think deeply about and learn from texts. Most people would probably agree that fluency for its own sake has little value (Pikulski & Chard, 2005).

Other studies have identified potential weaknesses in the link between comprehension and fluency as well. Applegate, Applegate, and Modla (2009) studied "good readers" (p. 512) and found that teachers often defined this group based on fluency alone even when the so-called good readers did not comprehend well. They learned that "for many...students...the freed-up resources that result from automaticity and fluency do not necessarily or automatically flow toward comprehension" (p. 519). Another study profiled students who failed their statewide reading assessment and found that students fail these tests for a variety of reading-related (and no doubt, some not reading-related) reasons (Buly & Valencia, 2002). Eighteen percent of students who failed read quickly and accurately, but did not comprehend; over 60% of the students in this group were learning English and may have lacked the English oral language to know what they read. Another group (also 18% of the sample) stumbled over words, but used slow reading and context to figure out the general meaning of the passage. Thus, in that study for over a

third of the students who failed the state assessment, the link between comprehension and fluency failed to hold up.

Some conceptual problems exist with this link as well. Paris et al. (2005) identified the obvious, that reading fast does not make children understand better and reading slowly does not necessarily make them understand worse. They identified other potential conceptual challenges too. Students who spend so much time thinking about accuracy (and thus do produce a fluent reading) might have no resources left to comprehend. Younger readers may have fewer well-developed comprehension strategies. Students who try to read fast might actually miscue more. And, some readers just feel uncomfortable reading aloud, especially in a testing environment.

Paris (2005) also identified some statistical problems with how researchers have analyzed fluency. Fluency is a constrained skill, or one that children master. They do not just read faster and faster each year forever. Because they reach asymptotic levels (or get as fast as they should), there comes a point in their reading development when fluency does not predict anything anymore. Fluency scores also usually represent little variation (when researchers present passages that children can read with 95% accuracy or higher), so they do not produce a normal distribution to which parametric statistics best apply (S. G. Paris et al., 2005). Because fluency quickly changes from a skill students do not have at all to one where they have reached ceiling levels, Paris suggested that the strong correlations only appear briefly in early childhood. He cautioned against considering it an "enduring individual difference variable" (S. G. Paris, 2005, p. 184).

So, the relationship between fluency and comprehension is not straightforward. Nor is it strictly linear (Nese et al., 2013), and the correlation between the two does

decrease as children get older—and better—as readers (Kim & Wagner, 2015). Students with reading disabilities may slow down and read as slow as they need to in order to facilitate understanding (Walczyk et al., 2007). O'Connor (2017) explained how supporting students to make fluency gains takes a lot of instructional time. She suggested teachers should know when a student is reading fast enough to facilitate comprehension: at what point does reading more fluently not help students with reading problems understand any better? She found a stronger link between comprehension and fluency for typical than for "poor" readers (p. 1), and she also found a ceiling—lower than grade-level norms—above which improving fluency did not lead to comprehension benefits. She expected the same might apply to students reading in a second language; fluency has some value, but after achieving a basic acceptable rate, spending instructional time on fluency may not help bilingual students to comprehend any better (personal communication, November 25, 2017; see also García & Godina, 2017). Research has yet to explore this question specifically.

Students who learn to read in English as a second language often achieve comparably to their monolingual English-speaking peers on word-level tasks (Lesaux & Geva, 2006). For some of these students, their English word skills are high, but English word knowledge is low (Mancilla-Martinez & Lesaux, 2011). They know how to decode the words, but they have not yet had educational experiences that help them develop the oral language and vocabulary to know what the words mean. These students would read fluently without a coherent idea of what they read. On the other hand, as bilingual people, some students have experience making meaning in environments where they do not understand every word. They bring a strength of knowing how to coordinate multiple

sources of information like context and illustrations to understand the ideas of a text even when they have not perfectly pronounced each word. These students fit the profile of "word stumblers" who still use context to know what a passage says (Buly & Valencia, 2002).

The students in this study demonstrated multiple patterns: sometimes they read fluently and could not retell much, sometimes they did not read fluently but could still retell an average number of ideas, and other times they read and retold about the same. They did however illustrate that the link between comprehension and fluency was not as tight for them as the research literature has described for other populations. The weak relationship may have to do with their status as older readers, readers reading below grade level, children reading in a second language, or some combination of these factors.

This study suggests the need to better understand what role fluency has for bilingual readers. A study like O'Connor's (2017) that seeks to identify the point at which increased fluency instruction has little benefit for bilingual students would help teachers prioritize instructional needs. Additional research should explore what strategies bilingual students use to understand texts when they do not know every word and how to support bilingual students who do fluently read texts that they do not understand.

Small Groups that Accomplish Multiple Purposes for Bilingual Students

This study highlighted the need for educational research that works towards designing small-group reading instruction that accomplishes multiple purposes for bilingual students. Particularly, it accentuated the tension between having dynamic conversations while engaging actively with books and actually reading those books independently. The students in this study engaged a lot and read a little. Hoffman (2017)

criticized this distinction when he asked, “What will it take for us to recognize that students exploring books that are beyond their ‘just right’ levels are engaged in real reading?” (p. 268). He found that bilingual students engaged in practices he described as “co-reading” (p. 270) in which they collaborated during inquiry projects to construct meaning from various text features in informational books. Hoffman vigorously defended this type of engagement with challenging material and suggested that teachers and researchers want to dismiss it because they do not understand the reading that is happening and have no idea how to support it.

Aukerman et al. (2017) described a process similar to Hoffman’s “co-reading,” something they termed “intercomprehending.” These researchers observed emergent bilingual students collaborating to make meaning in text-based discussions, and they argued for a collaborative and social view of comprehension that currently does not pervade literacy theory, assessment, or classroom practice. Like Hoffman, they described the social meaning-making practices of bilingual students as “underrecognized” (p. 484). They advocated that understanding group reading processes would help researchers and teachers reposition bilingual students as “the competent, thoughtful textual meaning makers we believe them to be, rather than as struggling readers” (p. 484). They suggested that these group processes matter particularly for bilingual students who work with linguistically and culturally unfamiliar and otherwise challenging texts.

The field may well lack this understanding of how students learn from complex texts. Yet, Hiebert (2017) pointed out that “if the kids don’t read the text, that’s called *listening*.” She suggested that it disadvantages democracy and students themselves when they cannot independently access texts. Even scholars who wish to foreground the social

context of reading would no doubt agree. When Moses (2015) described the ways that images support bilingual students in constructing meaning, she explained that images can prompt students to seek access to the written language. She then described a group of bilingual students asking a monolingual native English speaker what the text said about a photo they had discussed. This positive peer-support should occur in classrooms, and can serve as a step on the way to independent reading. Yet, ultimately all students will need to read the text on their own.

Much reading research comes from one of these two camps: students should have positive literacy experiences in supportive classroom communities or students should build fluency and comprehension with texts they read on their own. The dichotomy has not helped the field. Students should do both. They should have engaging conversations, become acquainted with a variety of literature and informational texts, and they should feel a part of a literate classroom community. They should develop in this way while they learn to really read and understand texts. They may have a richer literate life, but will certainly have more positive educational outcomes across their school experiences, if they know how to decode and comprehend texts on their own.

I observed early in this study that students engaged directly with the text less than I wanted them to, and I saw that they did not get the full power of their independent reading time. I wanted to change the protocol to deepen their interaction with the text, but I had committed to the alternating treatment design and could not make changes. This experience underlines the need for more work, particularly design-based research (Reinking & Bradley, 2008), around developing small-group instructional protocols that serve multiple purposes for bilingual students. I suspect that bilingual students can have

interpretive responses and engaging discussions around texts that they really read on their own, but in this study, they did not. Design-based research permits researchers to make modifications in response to data during the study, and so it would help develop instruction that meets multiple goals at once.

Design-based research would also allow researchers to explore ways to foreground culture and language in small groups for bilingual students. In this study, the protocol did not invite students to make connections between what they knew in Spanish and what they read in English or to see connections between the two languages. More explicitly attending to issues of language and culture could potentially increase student engagement, help them develop new strategies for reading (García & Godina, 2017), and push them into the text. A protocol should support the comprehension of bilingual students in text discussions by pushing them into the text with independence, using their funds of knowledge including their home language, and developing a true dialog.

These multiple purposes in tension illuminate the need to retheorize what it means for small groups to "work" and then develop measures that help teachers evaluate how students respond to their instruction. I already described the limitations of the retell measure for comprehension, but teachers need something like it that they can implement quickly and with any text. Maze tasks, oral and written retells, and sentence verification tasks all provide some useful measures of comprehension (Marcotte & Hintze, 2009), but they may not overcome the limitations noted in this study without introducing other significant drawbacks. In addition to fast and versatile comprehension assessment, teachers need consistent ways to measure students' engagement and discussion. Teachers will not, for example, transcribe reading groups to code for instances of inferential talk.

The field needs some other quick and classroom-based assessments to help teachers gauge the quality of discussion and student participation.

Grouping

This study's findings that in some ways students reading below grade level benefited from reading difficult texts suggests they could participate in heterogenous small groups. In such groups, they could read and discuss alongside more proficient readers while they work with grade-level text. Teachers often group students by ability in order to provide relevant instruction to students with similar needs (Ford & Opitz, 2008; Fountas & Pinnell, 1996), but these groups become problematic when students begin to identify some groups as the "low" groups. Researchers have recommended grouping students of mixed proficiencies to avoid stigma and have strong reading and English language models in each group (S. K. Baker et al., 2014; Zweirs, O'Hara, & Pritchard, 2014).

Yet, this study did not compare different grouping strategies. In fact, it may be that students participated and engaged difficult texts as well as they did because the groups did not have more proficient readers who answered questions first or took over reading. Perhaps difficult text worked because the students had similar access to it and approached the discussion from roughly similar reading proficiencies. More research could address this question directly.

Teacher Education

Finally, research needs to tie work on small-group reading for bilingual students to teacher and paraprofessional education and professional development. It does no good to develop a protocol that supports multiple reading purposes like discussion and

engagement while also accelerating progress in areas like comprehension and fluency if teachers do not implement it widely in schools.

In many cases, teachers do not create opportunities for bilingual students to participate in meaning-focused discussions of connected texts (Allington, 2013; DaSilva Iddings, 2005; Valdés, 2001). A story from my research journal during this study illustrates this point. While I set up the recording equipment and books one day for this study, I began talking with an instructional assistant who also occasionally used the space to tutor students. She expressed to me that she was so glad I was coming to the school to work with students on reading because they really needed it. She then described a fifth-grade student that she was working with. She was a child who could "read anything" but "didn't understand it at all," so they were "working on phonics with her, like silent e." This vignette illustrates many problems with the education this child and others in the school received, but I retell it here because it shows that, in defiance of common sense, schools do not always treat comprehension challenges with comprehension instruction. Indeed, many teachers and paraprofessionals have a hard time envisioning meaning-oriented reading support for students who struggle even if comprehension is the identified problem area.

Deficit views of what students can handle combined with years of policies emphasizing intervention in basic reading skills have constrained what educators can imagine offering for students reading below grade level. One research team reflected on their work with superintendents, administrators, and teachers of Latinx immigrant students and described that "caring experienced educators" frequently explained that their students "could not handle dynamic, agentic learning experiences" because they lacked

sufficient English vocabulary (Adair, Colegrove, & McManus, 2017, p. 309). In my own work, I have observed how pre-service teachers struggle to visualize literacy instruction for students learning English that extends beyond teacher-directed whole-group activities focused on basic skills (Kelly, 2017). For these teachers, the state accountability context defined their vision of what instruction for bilingual students should look like.

Thus, even if this study had developed a robust multipurpose small-group reading protocol for bilingual students, it would not represent a real contribution to the research literature without also exploring what it takes to help teachers and paraprofessionals implement it. As the field continues to grow in understanding small-group reading support, attention to teacher education (pre-service, in-service, and paraprofessional) will determine the impact of the work.

Conclusion

I undertook this study after several years of teaching reading in primary classes in the early years of implementing the Common Core State Standards. With the new push for even young students to read complex texts, I wanted to expand the research base about how these children do when given such tough texts. I regret that I cannot distill the findings to one simple and clear statement. My best effort still extends across at least two sentences:

> Students reading close to grade level benefited from matched texts. But, students reading below grade level, while they benefited from matched texts in terms of fluency, had richer discussions with difficult (ie., grade-level) texts.

My sociocultural orientation led me to elevate the findings about discussion and engagement above the findings about fluency. Careful readers of this entire study will

understand that even these two sentences fail to include many of the subtleties and exceptions that became clear across different children and for different outcomes.

In other words, this study showed that students need both matched and difficult texts for different purposes. Its findings do not suggest that teachers should select one level of text and only use that level in small groups. The wide variation and nuanced findings of this study highlight how providing appropriate small-group reading instruction requires a considerable amount of professional expertise. Teachers have to know their students' reading proficiencies and needs as well as their background knowledge and interests. They have to know the variety of texts available for children to read and know how to align these texts with what children need next in their reading instruction. Finally, they have to know how to evaluate children's understanding of and interactions with texts in the moment in order to document growth and determine next steps. The findings from this study suggest that pinning down the exact right reader-to-text match is probably not the vanguard of reading research in the next decade, but that students should have access to a variety of text levels, with appropriate supports as needed, and teachers should know how to use various levels of texts for many different purposes to facilitate students' reading growth.

REFERENCES

Adair, J. K., Colegrove, K. S.-S., & McManus, M. E. (2017). How the word gap argument negatively impacts young children of Latinx immigrants' conceptualizations of learning. *Harvard Educational Review*, *87*(3).

Afflerbach, P. P., & Cho, B. (2008). The classroom assessment of reading. In *Handbook of reading research* (Vol. 4). London: Routledge.

Alexander, P. A. (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist*, *47*(4), 259–280. https://doi.org/10.1080/00461520.2012.722511

Alexander, P. A., & Fox, E. (2013). A historical perspective on reading research and practice, redux. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 3–46). Newark, DE: International Reading Association.

Allington, R. L. (2009). If they don't read much...30 years later. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 30–54). New York: Guilford Press.

Allington, R. L. (2013). What really matters when working with struggling readers. *The Reading Teacher*, *66*(7), 520–530. https://doi.org/10.1002/TRTR.1154

Allington, R. L., McCuiston, K., & Billen, M. (2015). What research says about text complexity and learning to read. *The Reading Teacher*, *68*(7), 491–501. https://doi.org/10.1002/trtr.1280

Almasi, J. F., McKeown, M. G., & Beck, I. L. (1996). The nature of engaged reading in classroom discussions of literature. *Journal of Literacy Research: JLR*, *28*(1), 107–146.

Amendum, S. J., Conradi, K., & Hiebert, E. H. (2017). Does text complexity matter in the elementary grades? A research synthesis of text difficulty and elementary students' reading fluency and comprehension. *Educational Psychology Review*. https://doi.org/10.1007/s10648-017-9398-2

Amendum, S. J., Conradi, K., & Liebfreund, M. D. (2016). The push for more challenging texts: An analysis of early readers' rate, accuracy, and comprehension. *Reading Psychology*, *37*(4), 570–600. https://doi.org/10.1080/02702711.2015.1072609

Amendum, S. J., & Fitzgerald, J. (2010). Reading instruction research for English-language learners in kindergarten through sixth grade, the last twenty years. In A.

McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disability research* (1st ed., pp. 373–391). New York: Routledge.

Anderson, R. C., Chinn, C., Waggoner, M., & Nguyen, K. (1998). Intellectually stimulating story discussions. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning*. New York: The Guilford Press.

Applegate, M. D., Applegate, A. J., & Modla, V. B. (2009). "She's my best reader; she just can't comprehend": Studying the relationship between fluency and comprehension. *The Reading Teacher, 62*(6), 512–521. https://doi.org/10.1598/RT.62.6.5

Arizpe, E., & Styles, M. (2002). *Children reading pictures: Interpreting visual texts*. London: Routledge.

Aukerman, M., Schuldt, L. C., Aiello, L., & Martin, P. C. (2017). What meaning-making means among us: The intercomprehending of emergent bilinguals in small-group text discussions. *Harvard Educational Review, 87*(4), 482–511. https://doi.org/10.17763/1943-5045-87.4.482

Au, K. H.-P. (1980). Participation structures in a reading lesson with Hawaiian children: Analysis of a culturally appropriate instructional event. *Anthropology & Education Quarterly, 11*(2), 91–115.

Avalos, M. A., Plasencia, A., Chavez, C., & Rascón, J. (2007). Modified guided reading: Gateway to English as a second language and literacy learning. *The Reading Teacher, 61*(4), 318–329. https://doi.org/10.1598/RT.61.4.4

Baker, S. K., Geva, E., Kieffer, M. J., Lesaux, N., Thompson, S. L., Morris, J., … Russell, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school* (IES Practice Guides No. NCEE 2014-4012). What Works Clearinghouse.

Bassey, M. (1999). *Case study research in educational settings* (1st ed.). Philadelphia: Open University Press.

Beaver, J. (2012). *Developmental reading assessment* (2nd ed.). Upper Saddle River, N.J.: Celebration Press.

Beck, I. L., McKeown, M. G., Sandora, C., Kucan, L., & Worthy, J. (1996). Questioning the author: A yearlong classroom implementation to engage students with text. *The Elementary School Journal, 96*(4), 385–414.

Begeny, J. C., Daly, E. J., & Valleley, R. J. (2006). Improving oral reading fluency through response opportunities: A comparison of phrase drill error correction with

repeated readings. *Journal of Behavioral Education*, *15*(4), 229–235. https://doi.org/10.1007/s10864-006-9028-4

Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young readers. *Reading Research Quarterly*, *45*(4), 388–404. https://doi.org/10.1598/RRQ.45.4.2

Bernhardt, E. (2009). Increasing reading opportunities for English language learners. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 190–209). New York: Guilford Press.

Betts, E. A. (1946). *Foundations of reading instruction, with emphasis on differentiated guidance*. New York: American Book Company.

Bliss, S. L., Skinner, C. H., & Adams, R. (2006). Enhancing an English language learning fifth-grade student's sight-word reading with a time-delay taped-words intervention. *School Psychology Review*, *35*(4), 663–670.

Braunger, J., & Lewis, J. P. (2006). *Building a knowledge base in reading* (2nd ed.). Newark, DE: International Reading Association & National Council of Teachers of English.

Brickhouse, N. W. (2001). Embodying science: A feminist perspective on learning. *Journal of Research in Science Teaching*, *38*(3), 282–295.

Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology*, *88*(1), 18–37. https://doi.org/10.1037/0022-0663.88.1.18

Brown, S., & Kappes, L. (2012). Implementing the Common Core State Standards: A primer on "close reading of text." The Aspen Institute. Retrieved from www.aspeninstitute.org/education

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research: QR*, *6*(1), 97–113. https://doi.org/10.1177/1468794106058877

Buly, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis*, *24*(3), 219–239. https://doi.org/10.3102/01623737024003219

Bunch, G. C., Walqui, A., & Pearson, P. D. (2014). Complex text and new common standards in the United States: Pedagogical implications for English learners. *TESOL Quarterly*, *48*(3), 533–559.

Bundick, M. J., Quaglia, R. J., Corso, M. J., & Haywood, D. E. (2014). Promoting student engagement in the classroom. *Teachers College Record*, *116*, 1–34.

Burkins, J. M., & Croft, M. M. (2010). *Preventing misguided reading: New strategies for guided reading groups*. Newark, DE: International Reading Association.

Burkins, J. M., & Yaris, K. (2014). Break through the frustration: Balance vs. all-or-nothing thinking. *Reading Today*, *32*(2), 26–27.

Burns, M. K. (2007). Reading at the instructional level with children identified as learning disabled: Potential implications for response-to-intervention. *School Psychology Quarterly: The Official Journal of the Division of School Psychology, American Psychological Association*, *22*(3), 297–313. https://doi.org/10.1037/1045-3830.22.3.297

Burns, M. K., Pulles, S. M., Maki, K. E., Kanive, R., Hodgson, J., Helman, L. A., … Preast, J. L. (2015). Accuracy of student performance while reading leveled books rated at their instructional level by a reading inventory. *Journal of School Psychology*, *53*(6), 437–445. https://doi.org/10.1016/j.jsp.2015.09.003

Carrison, C., & Ernst-Slavit, G. (2005). From silence to a whisper to active participation: Using literature circles with ELL students. *Reading Horizons*, *46*(2), 93–113.

Cazden, C. B. (1981). Performance before competence: Assistance to child discourse in the zone of proximal development. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, *3*(1), 5–8.

Cazden, C. B. (1988). *Classroom discourse: the language of teaching and learning*. Portsmouth, NH: Heinemann.

Cervetti, G. N., & Hiebert, E. H. (2015). The sixth pillar of reading instruction: Knowledge development. *The Reading Teacher*, *68*(7), 548–551. https://doi.org/10.1002/trtr.1343

Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin (Ed.), *Vygotsky's educational theory in cultural context* (pp. 39–64). Cambridge: Cambridge University Press.

Chall, J. (1983). *Stages of reading development*. New York: McGraw Hill.

Chamot, A. U. (2005). The Cognitive Academic Language Learning Approach (CALLA): An update. In P. A. Richard-Amato & M. A. Snow (Eds.), *Academic success for English language learners: Strategies for K-12 mainstream teachers* (pp. 87–102). Pearson.

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

Clay, M. M. (2001). *Running records*. Portsmouth, NH: Heinemann.

Clay, M. M., & Cazden, C. B. (1990). A Vygotskian interpretation of Reading Recovery. In L. C. Moll (Ed.), *Vygotsky and education: instructional implications and applications of sociohistorical psychology* (pp. 206–222). New York: Cambridge University Press.

Coleman, D., & Pimentel, S. (2012). Revised publishers' criteria for the Common Core State Standards in English language arts and literacy, grades K-2. National Governors Association Center for Best Practices.

Collins, M. F. (2016). Supporting inferential thinking in preschoolers: Effects of discussion on children's story comprehension. *Early Education and Development*, *27*(7), 932–956. https://doi.org/10.1080/10409289.2016.1170523

Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, *24*(8), 1408–1419. https://doi.org/10.1177/0956797612472204

Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J. R., Lundblom, E., Crowe, E. C., & Fishman, B. (2011). Effective classroom instruction: Implications of child characteristics by reading instruction interactions on first graders' word reading achievement. *Journal of Research on Educational Effectiveness*, *4*(3), 173–207. https://doi.org/10.1080/19345747.2010.510179

Corbin, J., & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: Sage.

Council of Chief State School Officers. (2010). Common Core State Standards. National Governors Association Center for Best Practices, Washington D.C.

Creswell, J. W., & Clark, V. P. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945.

Cunningham, J. W. (2013). Research on text complexity: The Common Core State Standards as catalyst. In S. B. Neuman & L. B. Gambrell (Eds.), *Quality reading*

instruction in the age of Common Core Standards (pp. 136–148). Newark, DE: International Reading Association.

DaSilva Iddings, A. C. (2005). Linguistic access and participation: ELLs in an English-dominant community of practice. *Bilingual Research Journal, 29*, 165–183.

DaSilva Iddings, A. C., Risko, V. J., & Rampulla, M. P. (2009). When you don't speak their language: Guiding English-language learners through conversations about text. *The Reading Teacher, 63*(1), 52–61. https://doi.org/10.1598/RT.63.1.5

de Jong, E. J., & Harper, C. A. (2005). Preparing mainstream teachers for English-language learners: Is being a good teacher good enough? *Teacher Education Quarterly, 32*(2), 101–124.

Dennis, M. S., Sorrells, A. M., & Falcomata, T. S. (2016). Effects of two interventions on solving basic fact problems by second graders with mathematics learning disabilities. *Learning Disability Quarterly, 39*(2), 95–112. https://doi.org/10.1177/0731948715595943

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *The SAGE handbook of qualitative research* (3rd ed). Thousand Oaks, CA: Sage Publications.

Didden, R., Prinsen, H., & Sigafoos, J. (2000). The blocking effect of pictorial prompts on sight-word reading. *Journal of Applied Behavior Analysis, 33*(3), 317–320.

Dreher, M. J. (2000). Fostering reading for learning. In L. Baker, M. J. Dreher, & J. T. Guthrie (Eds.), *Engaging young readers: promoting achievement and motivation* (pp. 68–93). New York: Guilford Press.

Duke, N. K. (2000a). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35*, 202–224.

Duke, N. K. (2000b). For the rich it's richer: Print experiences and environments offered to children in very low and very high-SES first grade classrooms. *American Educational Research Journal, 37*, 441–478.

Duke, N. K., & Billman, A. K. (2009). Informational text difficulty for beginning readers. In E. H. Hiebert & M. Sailors (Eds.), *Finding the right texts: What works for beginning and struggling readers* (pp. 109–128). New York: Guilford Press.

Duke, N. K., Purcell-Gates, V., Hall, L., & Tower, C. (2006). Literacy activities for developing comprehension and writing. *The Reading Teacher, 60*(4), 344–355.

Dunn, W. E., & Lantolf, J. P. (1998). Vygotsky's zone of proximal development and Krashen's i + 1: Incommensurable constructs; incommensurable theories. *Language Learning*, *48*(3), 411–442.

Dunst, C. J., Williams, A. L., Trivette, C. M., Simkus, A., & Hamby, D. W. (2010). Relationships between inferential book reading strategies and young children's language and literacy competence. *CELLReviews*, *5*(10), 1–10.

Echevarria, J., Vogt, M., & Short, D. J. (2016). *Making content comprehensible for English Learners: The SIOP model* (5th ed.). New York City, NY: Pearson.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262–300.

Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, *44*(2), 414–448. https://doi.org/10.3102/0002831207302175

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading*, *17*(6), 435–452. https://doi.org/10.1080/10888438.2013.774005

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (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), 1–44. https://doi.org/10.1080/19345740802539200

Erickson, F. (1986). Qualitative methods in research on teaching. In *Handbook of research on teaching* (3rd ed., pp. 119–161). New York: Macmillan.

Faggella-Luby, M., Griffith, R. R., Silva, C., & Weinburgh, M. H. (2016). Assessing ELLs' reading comprehension and science understandings using retellings. *Electronic Journal of Science Education*, *20*(3), 150–166.

Fang, Z. (2016). Teaching close reading with complex texts across content areas. *Research in the Teaching of English*, *51*(1), 106–117.

Fang, Z., & Schleppegrell, M. J. (2010). Disciplinary literacies across content areas: Supporting secondary reading through functional language analysis. *Journal of Adolescent & Adult Literacy*, *53*(7), 587–597.

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review, 20*(3), 586–592. https://doi.org/10.3758/s13423-012-0367-y

Ferron, J., Foster-Johnson, L., & Kromrey, J. D. (2003). The functioning of single-case randomization tests with and without random assignment. *Journal of Experimental Education, 71*(3), 267–288.

Fisher, C., Berliner, D., Filby, N., Marliave, R., Cahen, L., & Dishaw, M. (2015). Teaching behaviors, academic learning time, and student achievement. *Journal of Classroom Interaction, 50*(1), 6–24.

Fisher, D., & Frey, N. (2012). Close reading in elementary schools. *The Reading Teacher, 66*(3), 179–188. https://doi.org/10.1002/TRTR.01117

Fisher, D., & Frey, N. (2014a). Closely reading informational texts in the primary grades. *The Reading Teacher, 68*(3), 222–227. https://doi.org/10.1002/trtr.1317

Fisher, D., & Frey, N. (2014b). Close reading as an intervention for struggling middle school readers. *Journal of Adolescent & Adult Literacy, 57*(5), 367–376. https://doi.org/10.1002/jaal.266

Fisher, D., & Frey, N. (2014c). Student and teacher perspectives on a close reading protocol. *Literacy Research and Instruction, 53*(1), 25–49. https://doi.org/10.1080/19388071.2013.818175

Fisher, D., Frey, N., Anderson, H. L., & Thayre, M. C. (2014). *Text-dependent questions, grades K-5: Pathways to close and critical reading*. Thousand Oaks, CA: Corwin.

Fisher, D., Frey, N., & Lapp, D. (2012). *Text complexity: Raising rigor in reading*. Newark, DE: International Reading Association.

Fitzgerald, J., Hiebert, E. H., Bowen, K., Relyea-Kim, E. J., Kung, M., & Elmore, J. (2015). Text complexity: Primary teachers' views. *Literacy Research and Instruction, 54*(1), 19–44. https://doi.org/10.1080/19388071.2014.954086

Florit, E., Roch, M., & Levorato, M. C. (2014). Listening text comprehension in preschoolers: a longitudinal study on the role of semantic components. *Reading and Writing, 27*(5), 793–817. https://doi.org/10.1007/s11145-013-9464-1

Foorman, B. R. (2009). Text difficulty in reading assessment. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 231–247). New York: Guilford Press.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on

reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology, 31*(1), 1–29. https://doi.org/10.1016/j.cedpsych.2004.11.003

Ford, M. P., & Opitz, M. F. (2008). A national survey of guided reading practices: What we can learn from primary teachers. *Literacy Research and Instruction, 47*(4), 309–331. https://doi.org/10.1080/19388070802332895

Ford, M. P., & Opitz, M. F. (2011). Looking back to move forward with guided reading. *Reading Horizons, 50*(4), 225–240.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2005). *Leveled books, K-8: Matching texts to readers for effective teaching* (1st ed.). Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2012). Guided reading: The romance and the reality. *The Reading Teacher, 66*(4), 268–284. https://doi.org/10.1002/TRTR.01123

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 369–394). Mahwah, NJ: Erlbaum.

Frey, N., & Fisher, D. (2013). *Rigorous reading: 5 access points for comprehending complex texts*. Thousand Oaks, CA: Corwin.

Fry, E. (1968). A readability formula that saves time. *Journal of Reading, 11*(7), 513–578.

Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014). What is intensive instruction and why is it important? *TEACHING Exceptional Children, 46*(4), 13–18. https://doi.org/10.1177/0040059914522966

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239–256. https://doi.org/10.1207/S1532799XSSR0503_3

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20–28. https://doi.org/10.1177/074193258800900206

Fulmer, S. M., D'Mello, S. K., Strain, A., & Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology, 41*, 98–110.

Gambrell, L. B., Wilson, R. M., & Gantt, W. N. (1981). Classroom observations of task-attending behaviors of good and poor readers. *The Journal of Educational Research, 74*(6), 400–404.

Gamson, D. A., Lu, X., & Eckert, S. A. (2013). Challenging the research base of the Common Core State Standards: A historical reanalysis of text complexity. Educational Researcher, 42(7), 381–391. https://doi.org/10.3102/0013189X13505684

García, G. E., & Godina, H. (2017). A window into bilingual reading: The bilingual reading practices of fourth-grade, Mexican American children who are emergent bilinguals. *Journal of Literacy Research, 49*(2), 273–301. https://doi.org/10.1177/1086296X17703727

Gelzheiser, L., Hallgren-Flynn, L., Connors, M., & Scanlon, D. (2014). Reading thematically related texts to develop knowledge and comprehension. *The Reading Teacher, 68*(1), 53–63. https://doi.org/10.1002/trtr.1271

Gersten, R., & Baker, S. K. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children, 66*, 454–470.

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: A practice guide* (No. NCEE 2007-4011). Washington, D.C: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide* (No. NCEE 2009-4045). Washington, D.C: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Gickling, E. E., & Armstrong, D. L. (1978). Levels of instructional difficulty as related to on-task behavior, task completion, and comprehension. *Journal of Learning Disabilities, 11*(9), 32–39.

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013). Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, *48*(2), 135–154. https://doi.org/10.1002/rrq.45

Gilbertson, D., & Bluck, J. (2006). Improving responsiveness to intervention for English-language learners: A comparison of instructional pace on letter naming rates. *Journal of Behavioral Education*, *15*(3), 131–147. https://doi.org/10.1007/s10864-006-9013-y

Glasswell, K., & Ford, M. P. (2010). Teaching flexibly with leveled texts: More power for your reading block. *The Reading Teacher*, *64*(1), 57–60. https://doi.org/10.1598/RT.64.1.7

Goldenberg, C. (1992). Instructional conversations: Promoting comprehension through discussion. *The Reading Teacher*, *46*(4), 316–326.

Goldman, S., O'Banion Varma, K., & Sharp, D. (1999). Children's understanding of complex stories: Issues of representation and assessment. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 135–159). London: Routledge.

Goncu, A., & Gauvain, M. (2012). Sociocultural approaches to educational psychology. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (Vol. 1, pp. 125–154). Washington, D.C: American Psychological Association.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement. Retrieved from http://dibels.uoregon.edu

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix providing multilevel analyses of text characteristics. *Educational Researcher*, *40*(5), 223–234.

Greene, J. C. (2007). Mixing methods on purpose and designing mixed methods studies. In *Mixed methods in social inquiry* (pp. 95–137). San Francisco, CA: Jossey-Bass.

Güler, Ö. O., & Özmen, R. R. (2010). Using the brief experimental analysis to determine the effective reading comprehension strategy in story comprehension of students with mental retardation. *International Online Journal of Educational Sciences*, *2*(3), 930–932.

Guthrie, J. T., & Anderson, E. (1999). Engagement in reading: Processes of motivated, strategic, knowledgeable, social readers. In J. T. Guthrie & D. E. Alvermann (Eds.), *Engaged reading: Processes, practices, and policy implications* (pp. 17–45). New York, NY: Teachers College Press.

Guthrie, J. T., Schafer, W. D., & Chun-Wei, H. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *The Journal of Educational Research, 94*(3), 145–163.

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, *92*(2), 331–341. https://doi.org/10.1037/0022-0663.92.2.331

Guthrie, J. T., Wigfield, A., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In *Handbook of research on student engagement.* (pp. 601–634). New York, NY: Springer Science + Business Media.

Gutiérrez, K., & Lee, C. D. (2009). Robust informal learning environments for youth from nondominant groups: Implications for literacy learning in formal schooling. In L. M. Morrow, R. Rueda, & D. Lapp (Eds.), *Handbook of Research on Literacy and Diversity*. New York: The Guilford Press.

Gyovai, L. K., Cartledge, G., Kourea, L., Yurick, A., & Gibson, L. (2009). Early reading intervention: Responding to the learning needs of young at-risk English language learners. *Learning Disability Quarterly*, *32*, 143–162.

Halladay, J. L. (2012). Revisiting key assumptions of the reading level framework. *The Reading Teacher*, *66*(1), 53–62. https://doi.org/10.1002/TRTR.01093

Hall, M. S., & Burns, M. K. (2017). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology*. https://doi.org/10.1016/j.jsp.2017.11.002

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, *59*(7), 636–644.

Hatcher, P. (2000). Predictors of Reading Recovery book levels. *Journal of Research in Reading*, *23*(1), 67–77. https://doi.org/10.1111/1467-9817.00103

Hawkins, R. O., Hale, A., Sheeley, W., & Ling, S. (2011). Repeated reading and vocabulary-previewing interventions to improve fluency and comprehension for struggling high-school readers. *Psychology in the Schools, 48*(1), 59–77. https://doi.org/10.1002/pits.20545

Hiebert, E. H. (2005). The effects of text difficulty on second graders' fluency development. *Reading Psychology, 26*(2), 183–209. https://doi.org/10.1080/02702710590930528

Hiebert, E. H. (2017). *New directions in text complexity from kindergarten and beyond: Evidence from three studies*. Presented at the Literacy Research Association, Tampa, FL.

Hiebert, E. H., & Mesmer, H. A. E. (2013). Upping the ante of text complexity in the Common Core State Standards: Examining its potential impact on young readers. *Educational Researcher , 42*(1), 44–51. https://doi.org/10.3102/0013189X12459802

Hirsch, E. D. (2006). The case for bringing content into the language arts block and for a knowledge-rich curriculum core for all children. *American Educator, 30*(1).

Hitchcock, J. (2010). Single-case designs and qualitative methods: Applying a mixed methods research perspective. *Mid-Western Educational Researcher, 23*(2), 49–58.

Hoffman, J. V. (2017). What if "just right" is just wrong? The unintended consequences of leveling readers. *The Reading Teacher, 71*(3), 265–273. https://doi.org/10.1002/trtr.1611

Horner, R. H., & Odom, S. L. (2014). Constructing single-case research designs: Logic and options. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: methodological and statistical advances* (pp. 27–51). Washington, DC: American Psychological Association.

Jimenez-Silva, M., Gomez, L., & Cisneros, J. (2014). Examining Arizona's policy response post Flores v. Arizona in educating K–12 English language learners. *Journal of Latinos and Education, 13*(3), 181–195. https://doi.org/10.1080/15348431.2013.849600

Johnson, E. M. (2017). Balancing comprehension and conversation: How elementary teachers manage multiple purposes for text discussions. *Teaching and Teacher Education, 66*, 325–337. https://doi.org/10.1016/j.tate.2017.05.005

Johnston, S. S., Buchanan, S., & Davenport, L. (2009). Comparison of fixed and gradual array when teaching sound-letter correspondence to two children with autism who use AAC. *Augmentative and Alternative Communication , 25*(2), 136–144. https://doi.org/10.1080/07434610902921516

Kamil, M. L. (2008). How to get recreational reading to increase reading achievement. *57th Yearbook of the National Reading Conference*, 31–40.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed). New York: Oxford University Press.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral

comprehension. *Scientific Studies of Reading*, *12*(3), 281–300. https://doi.org/10.1080/10888430802132279

Kelley, M., & Clausen-Grace, N. (2006). R5: The sustained silent reading makeover that transformed readers. *The Reading Teacher*, *60*(2), 148–156. https://doi.org/10.1598/RT.60.2.5

Kelly, L. B. (2017). Preservice teachers' developing conceptions of teaching English learners. *TESOL Quarterly*, *52*(1), 110-136. https://doi.org/10.1002/tesq.375

Kelly, L. B., & Moses, L. (2018). Children's literature that sparks inferential discussions. *The Reading Teacher*. http://doi.org/10.1002/trtr.1675

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, *101*(4), 765–778. https://doi.org/10.1037/a0015956

Kieffer, M. J. (2012). Early oral language and later reading development in Spanish-speaking English language learners: Evidence from a nine-year longitudinal study. *Journal of Applied Developmental Psychology*, *33*(3), 146–157. https://doi.org/10.1016/j.appdev.2012.02.003

Kim, Y.-S. G., & Wagner, R. K. (2015). Text (oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from grades 1 to 4. *Scientific Studies of Reading*, *19*(3), 224–242. https://doi.org/10.1080/10888438.2015.1007375

Kinginger, C. (2001). i + 1 ≠ ZPD. *Foreign Language Annals*, *34*(5), 417–425.

Klingner, J. K., Artiles, A. J., & Barletta, L. M. (2006). English language learners who struggle with reading. *Journal of Learning Disabilities*, *39*, 108–128.

Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *The Elementary School Journal*, *96*(3), 275–293.

Klubnik, C., & Ardoin, S. P. (2010). Examining immediate and maintenance effects of a reading intervention package on generalization materials: Individual versus group implementation. *Journal of Behavioral Education*, *19*(1), 7–29. https://doi.org/10.1007/s10864-009-9096-3

Knapp, M. S. (Ed.). (1995). *Teaching for meaning in high-poverty classrooms*. New York: Teachers College Press.

Koslin, B. I., Zeno, S., & Koslin, S. (1987). *The DRP: An effective measure in reading*. New York College Entrance Examination Board.

Krashen, S. D. (1981). The "fundamental pedagogical principle" in second language teaching. *Studia Linguistica, 35*(1), 50–70.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S. D., & Williams, C. (2012). Is self-selected pleasure reading the cure for the long-term ELL syndrome? *NABE Perspectives,* 26.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education: RASE, 34*(1), 26–38. https://doi.org/10.1177/0741932512452794

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. What Works Clearinghouse. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kratochwill, T. R., & Levin, J. R. (2014). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: methodological and statistical advances* (pp. 53–89). Washington, DC: American Psychological Association.

Kress, G. R., & Van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). New York: Routledge.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1), 3–21. https://doi.org/10.1037/0022-0663.95.1.3

Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review, 55*(2), 178–195. https://doi.org/10.17763/haer.55.2.56142234616x4352

Lapp, D., Grant, M., Moss, B., & Johnson, K. (2013). Students' close reading of science texts: What's now? What's next? *The Reading Teacher,* 109–119. https://doi.org/10.1002/trtr.1191

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation* (1st ed.). Cambridge, England: Cambridge University Press.

Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Review of Educational Research, 83*(3), 432–479. https://doi.org/10.3102/0034654313480891

Lennon, C., & Burdick, H. (2004). The Lexile framework as an approach for reading measurement and success. Retrieved from http://www.lexile.com

Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4-6-year-old children. *Reading Research Quarterly*, *47*(3), 259–282. https://doi.org/http://dx.doi.org/10.1002/rrq.020

Lesaux, N., & Geva, E. (2006). Synthesis: Development of literacy in language-minority students. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth* (pp. 53–74). Mahwah, NJ: Lawrence Erlbaum Associates.

Leslie, L., & Caldwell, J. S. (2016). *Qualitative Reading Inventory-6* (6th ed.). Upper Saddle River, NJ: Pearson.

Lillie, K. E., Markos, A., Estrella, A., Nguyen, T., Trifiro, A., Arias, M. B., … Pérez, K. (2010). *Policy in practice: The implementation of Structured English Immersion in Arizona* (The Civil Rights Project: Proyecto Derechos Civiles). UCLA. Retrieved from http://civilrightsproject.ucla.edu

Lillie, K. E., Markos, A., Hornberger, N. H., & Baker. (2014). The four-hour block: SEI in classrooms. In S. C. K. Moore (Ed.), *Language policy processes and consequences: Arizona case studies* (pp. 133–155). Tonawanda, NY: Multilingual Matters.

Lively, B. A., & Pressey, S. L. (1923). A method of measuring the "vocabulary burden" of textbooks. *Educational Administration and Supervision Including Teacher Training*, *9*, 389–398.

Lubliner, S., & Hiebert, E. H. (2011). An analysis of English–Spanish cognates as a source of general academic language. *Bilingual Research Journal*, *34*(1), 76–93. https://doi.org/10.1080/15235882.2011.568589

Lyons, W., & Thompson, S. A. (2012). Guided reading in inclusive middle years classrooms. *Intervention in School and Clinic*, *47*(3), 158–166. https://doi.org/10.1177/1053451211423814

Lysynchuk, L. M., Pressley, M., & Vye, N. J. (1990). Reciprocal teaching improves standardized reading-comprehension performance in poor comprehenders. *The Elementary School Journal*, *90*(5), 469–484.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*(5), 1544–1560. https://doi.org/10.1111/j.1467-8624.2011.01633.x

Mantzicopoulos, P., & Patrick, H. (2010). "The seesaw is a machine that goes up and down": Young children's narrative responses to science-related informational text. *Early Education and Development, 21*(3), 412–444.

Marcotte, A. M., & Hintze, J. M. (2009). Incremental and predictive utility of formative assessment methods of reading comprehension. *Journal of School Psychology, 47*(5), 315–335. https://doi.org/10.1016/j.jsp.2009.04.003

Martinez, M., Roser, N. L., & Dooley, C. (2003). Young children's literary meaning making. In N. Hall, J. Larson, & J. Marsh (Eds.), *Handbook of early childhood literacy* (pp. 222–234). Thousand Oaks, CA: Sage.

Martínez-Roldán, C. M., & López-Robertson, J. M. (1999). Initiating literature circles in a first-grade bilingual classroom. *The Reading Teacher, 53*(4), 270–281.

Martin, J. R., & Rose, D. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School*. Bristol, CT: Equinox Publishing.

Mathes, P. G., & Fuchs, L. S. (1993). Peer-mediated reading instruction in special education resource rooms. *Learning Disabilities Research & Practice, 8*(4), 233–243.

Matson, J. L., Turygin, N. C., Beighley, J., & Matson, M. L. (2012). Status of single-case research designs for evidence-based practice. *Research in Autism Spectrum Disorders, 6*(2), 931–938. https://doi.org/10.1016/j.rasd.2011.12.008

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (3rd edition). Thousand Oaks, CA: SAGE Publications, Inc.

Mayer, R. E. (2012). Information processing. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (Vol. 1, pp. 85–99). Washington, D.C: American Psychological Association.

McCarthey, S. J. (2001). Identity construction in elementary readers and writers. *Reading Research Quarterly, 36*(2), 122–151.

McCarthey, S. J., Hoffman, J. V., & Galda, L. (1999). Readers in elementary classrooms: Learning goals and instructional principles that can inform practice. In J. T. Guthrie & D. E. Alvermann (Eds.), *Engaged reading: Processes, practices, and policy implications* (pp. 46–80). New York, NY: Teachers College Press.

McKeown, D., Brindle, M., Harris, K. R., Graham, S., Collins, A. A., & Brown, M. (2016). Illuminating growth and struggles using mixed methods: Practice-based professional development and coaching for differentiating SRSD instruction in writing. *Reading and Writing*, *29*(6), 1105–1140. https://doi.org/10.1007/s11145-016-9627-y

McNamara, D. S., & Kendeou, P. (2011). Translating advances in reading comprehension research to educational practice. *International Electronic Journal of Elementary Education*, *4*(1), 33–46.

Mellard, D. L., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Practice*, *24*(4), 186–195.

Mendéz, L. I., Crais, E. R., Castro, D. C., & Kainz, K. (2015). A culturally and linguistically responsive vocabulary approach for young Latino dual languages learners. *Journal of Speech, Language, and Hearing Research*, *58*, 93–106.

Menon, S., & Hiebert, E. H. (2011). Instructional texts and the fluency of learning disabled readers. In A. McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disability research*. London: Routledge.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.

Mesmer, H. (2015). Stretching elementary students in complex texts: Why? How? When? In E. H. Hiebert (Ed.), *Teaching stamina and silent reading in the digital-global age* (pp. 79–99). Santa Cruz, CA: TextProject, Inc.

Mesmer, H. A. E., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, *47*(3), 235–258. https://doi.org/10.1002/RRQ.019

Miller, A. C., & Keenan, J. M. (2009). How word decoding skill impacts text memory: The centrality deficit and how domain knowledge can compensate. *Annals of Dyslexia*, *59*(2), 99. https://doi.org/10.1007/s11881-009-0025-x

Miller, S. M. (2003). How literature discussion shapes thinking: ZPDs for teaching/learning habits of the heart and mind. In A. Kozulin (Ed.), *Vygotsky's educational theory in cultural context* (pp. 39–64). Cambridge: Cambridge University Press.

Moll, L. C. (Ed.). (1990). Introduction. In *Vygotsky and education: instructional implications and applications of sociohistorical psychology* (pp. 1–27). Cambridge: Cambridge University Press.

Morgan, A., Wilcox, B. R., & Eldredge, J. L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *The Journal of Educational Research*, *94*(2), 113–119. https://doi.org/10.1080/00220670009598749

Morris, D. (2015). Preventing early reading failure: An argument. *The Reading Teacher*, *68*(7), 502–509. https://doi.org/10.1002/trtr.1346

Morrow, L. M., & Young, J. (1997). A family literacy program connecting school and home: Effects on attitude, motivation, and literacy achievement. *Journal of Educational Psychology*, *89*(4), 736–742.

Moses, L. (2013). Viewing as a cultural tool in the construction of meaning with expository texts for young bilinguals. *Journal of Language and Literacy Education*, *9*(2), 72–93.

Moses, L. (2015). The role(s) of image for young bilinguals reading multimodal informational texts. *Language and Literacy*, *17*(3), 82–99.

Moses, L., & Kelly, L. B. (2017). The development of positive literate identities among emerging bilingual and monolingual first graders. *Journal of Literacy Research*, *49*(3), 393–423. https://doi.org/10.1177/1086296x17713291

Moses, L., Ogden, M., & Kelly, L. B. (2015). Facilitating meaningful discussion groups in the primary grades. *The Reading Teacher*, *69*(2), 233–237. https://doi.org/10.1002/trtr.1392

Moss, B. (2004). Teaching expository text structures through information trade book retellings. *The Reading Teacher*, *57*(8), 710–718.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, *101*(3), 740–764. https://doi.org/10.1037/a0015576

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Nastasi, B. K., & Schensul, S. L. (2005). Contributions of qualitative research to the validity of intervention research. *Journal of School Psychology*, *43*(3), 177–195. https://doi.org/10.1016/j.jsp.2005.04.003

National Academies of Sciences Engineering and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures*. Washington, DC: The National Academies Press.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (No. NIH Publication No. 00-4769). Washington, D.C: U.S. Government Printing Office.

National Reading Recovery Project. (1995). *Reading Recovery booklist: Suggested levels for books*. Wellington, New Zealand: Department of Education.

Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*. Retrieved from achievethecore.org

Nese, J. F. T., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across grades 1–8. *Journal of School Psychology*, *51*(5), 625–642. https://doi.org/10.1016/j.jsp.2013.05.006

Nes Ferrara, S. L. (2005). Reading fluency and self-efficacy: A case study. *International Journal of Disability, Development and Education*, *52*(3), 215–231. https://doi.org/10.1080/10349120500252858

Neufeld, P., & Fitzgerald, J. (2001). Early English reading development: Latino English learners in the "low" reading group. *Research in the Teaching of English*, *36*(1), 64–109.

Neuman, S. B., & McCormick, S. (Eds.). (1995). *Single-subject experimental research: Applications for literacy*. Newark, DE: International Reading Association.

Nilsson, N. L. (2008). A critical analysis of eight informal reading inventories. *The Reading Teacher*, *61*(7), 526–536. https://doi.org/10.1598/RT.61.7.2

Nolen, S. B. (2007). Young children's motivation to read and write: Development in social contexts. *Cognition and Instruction*, *25*(2-3), 219–270. https://doi.org/10.1080/07370000701301174

Noltemeyer, A., Joseph, L. M., & Watson, M. (2014). Improving reading prosody and oral retell fluency: A comparison of three intervention approaches. *Reading Improvement*, *51*(2), 221–232.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91–121. https://doi.org/10.1080/10888438.2010.529219

O'Connor, R. E. (2017). Reading fluency and students with reading disabilities: How fast is fast enough to promote reading comprehension? *Journal of Learning Disabilities*, 1–13. https://doi.org/10.1177/0022219417691835

O'Connor, R. E., Bell, K. M., Harty, K. R., Larkin, L. K., Sackor, S. M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology*, *94*(3), 474–485. https://doi.org/10.1037/0022-0663.94.3.474

O'Connor, R. E., Swanson, H. L., & Geraghty, C. (2010). Improvement in reading rate under independent and difficult text levels: Influences on word and comprehension skills. *Journal of Educational Psychology, 102*(1), 1–19. https://doi.org/10.1037/a0017488

Painter, C. (2013). *Reading visual narratives: Image analysis of children's picture books*. Sheffield: Equinox Publishing.

Palinscar, A., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction, 1*(2), 117–175.

Paris, A. G., & Paris, S. G. (2007). Teaching narrative comprehension strategies to first graders. *Cognition and Instruction, 25*(1), 1–44. https://doi.org/10.1080/07370000709336701

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202. https://doi.org/10.1598/RRQ.40.2.3

Paris, S. G., Carpenter, R. D., Paris, A. H., & Hamilton, E. E. (2005). Spurious and genuine correlates of children's reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 131–160). Mahwah, NJ: L Erlbaum Associates.

Parker, D. C., & Burns, M. K. (2014). Using the instructional level as a criterion to target reading interventions. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 30*(1), 79–94. https://doi.org/10.1080/10573569.2012.702047

Pearson, P. D. (1991). *Developing expertise in reading comprehension: What should be taught? How should it be taught?* (No. 512). Center for the Study of Reading.

Pearson, P. D. (2000). Reading in the twentieth century. In T. Good (Ed.), *American education: Yesterday, today, and tomorrow. Yearbook of the National Society for the Study of Education* (pp. 152–208). Chicago, IL: University of Chicago Press.

Pearson, P. D. (2013). Research foundations of the Common Core State Standards in English language arts. In S. B. Neuman & L. B. Gambrell (Eds.), *Quality reading instruction in the age of Common Core Standards* (pp. 237–261). Newark, DE: International Reading Association.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Perry, K. H. (2012). What is literacy? A critical overview of sociocultural perspectives. *Journal of Language and Literacy Education*, *8*(1), 50–71.

Peterson, B. (1991). Selecting books for beginning readers. In D. E. DeFord, C. A. Lyons, & G. S. Pinnell (Eds.), *Bridges to literacy: learning from reading recovery*. Portsmouth, NH: Heinemann.

Phillips, D. C. (2014). Research in the hard sciences, and in very hard "softer" domains. *Educational Researcher , 43*(1), 9–11. https://doi.org/10.3102/0013189X13520293

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*(6), 510–519. https://doi.org/10.1598/rt.58.6.2

Pinnell, G. S., & Fountas, I. C. (2010). *Benchmark assessment system 1* (2nd ed.). Portsmouth, NH: Heinemann.

Plavnick, J. B., & Ferreri, S. J. (2013). Single-case experimental designs in educational research: A methodology for causal analyses in teaching and learning. *Educational Psychology Review, 25*(4), 549–569. https://doi.org/10.1007/s10648-013-9230-6

Pressley, M., El-Dinary, P. B., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J., & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal*, *92*(5), 513–555.

Pressley, M., Wharton-McDonald, R., Allington, R. L., Block, C. C., Morrow, L., Tracey, D., … Woo, D. (2001). A study of effective first-grade literacy instruction. *Scientific Studies of Reading, 5*(1), 35–58. https://doi.org/10.1207/S1532799XSSR0501_2

Proctor, C. P., & Mo, E. (2009). The relationship between cognate awareness and English comprehension among Spanish-English bilingual fourth grade students. *TESOL Quarterly, 43*(1), 125–136.

Purcell-Gates, V., Duke, N. K., & Martineau, J. A. (2007). Learning to read and write genre-specific text: Roles of authentic experience and explicit teaching. *Reading Research Quarterly*, *42*(1), 8–45. https://doi.org/10.1598/RRQ.42.1.1

Purdy, J. (2008). Inviting conversation: Meaningful talk about texts for English language learners. *Literacy Discussion*, *42*(1), 44–51.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R & D program in reading comprehension* (No. MR-1465). Arlington, VA: RAND Publications.

Raphael, T. E., & McMahon, S. I. (1994). Book club: An alternative framework for reading instruction. *The Reading Teacher*, *48*(2), 102–116.

Rasinski, T., & Young, C. (2014). Assisted reading: A bridge from fluency to comprehension. *New England Reading Association Journal*, *50*(1), 1–4.

Reed, D. K., & Kershaw-Herrera, S. (2016). An examination of text complexity as characterized by readability and cohesion. *Journal of Experimental Education*, *84*(1), 75–97. https://doi.org/10.1080/00220973.2014.963214

Reed, D. K., & Vaughn, S. (2012). Retell as an indicator of reading comprehension. *Scientific Studies of Reading*, *16*(3), 187–217. https://doi.org/10.1080/10888438.2010.538780

Reinking, D., & Bradley, B. (2008). *Formative and design experiments: Approaches to language and literacy research*. New York: Teachers College Press.

Reis, S. M., Eckert, R. D., McCoach, D. B., Jacobs, J. K., & Coyne, M. (2008). Using enrichment reading practices to increase reading fluency, comprehension, and attitudes. *The Journal of Educational Research*, *101*(5), 299–314.

Renaissance Learning. (2015). *STAR Reading Assessment*.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*(6), 427–469. https://doi.org/10.1016/j.jsp.2009.07.001

Reutzel, D. R., & Smith, J. A. (2004). Accelerating struggling readers' progress: A comparative analysis of expert opinion and current research recommendations. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *20*(1), 63–89. https://doi.org/10.1080/10573560490242778

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children*, *82*(2), 144–169. https://doi.org/10.1177/0014402915585483

Rios-Aguilar, C., González-Canche, M., & Moll, L. c. (2010). *Implementing structured English Immersion in Arizona: Benefits, costs, challenges, & opportunities* (The Civil Rights Project: Proyecto Derechos Civiles). UCLA. Retrieved from http://civilrightsproject.ucla.edu/

Rodgers, E., D'Agostino, J. V., Kelly, R. H., & Mikita, C. (2018). Oral reading accuracy: Findings and implications from recent research. The Reading Teacher. https://doi.org/10.1002/trtr.1686

Rosebery, A., Warren, B., & Conant, F. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *The Journal of the Learning Sciences*, *2*, 61–94.

Rosenblatt, L. (2013). The transactional theory of reading and writing. In D. Alvermann, N. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 923–956). Newark, DE: International Reading Association. https://doi.org/10.1598/0710.35

Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research*, *95*(5), 259–272. https://doi.org/10.1080/00220670209596600

Ruddell, R. B., & Unrau, N. J. (1994). Reading as a meaning-construction process: The reader, the text, and the teacher. In R. B. Ruddell, M. R. Ruddell, & Singer (Eds.), *Theoretical models and processes of reading* (4th ed.). Newark, DE: International Reading Association.

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Los Angeles, CA: Sage.

Sanacore, J., & Palumbo, A. (2008). Understanding the fourth-grade slump: Our point of view. *The Educational Forum*, *73*(1), 67–74. https://doi.org/10.1080/00131720802539648

Santori, D. (2011). "Search for the answers" or "talk about the story"?: School-based literacy participation structures. *Language Arts*, *88*(3), 198–207.

Saul, E. W., & Dieckman, D. (2005). Choosing and using information trade books. *Reading Research Quarterly*, *40*(4), 502–513. https://doi.org/10.1598/RRQ.40.4.6

Saunders, W. M., & Goldenberg, C. (1999). Effects of instructional conversations and literature logs on limited- and fluent-English-proficient students' story comprehension and thematic understanding. *The Elementary School Journal*, *99*(4), 277–301.

Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4-12, 1914-2014. *Review of Educational Research*, *86*(3), 756–800. https://doi.org/10.3102/0034654316652942

Scarcella, R. C. (2003). *Accelerating academic English: A focus on the English learner*. Sacramento, CA: Regents of University of California.

Schisler, R., Joseph, L. M., Konrad, M., & Alber-Morgan, S. (2010). Comparison of the effectiveness efficiency of oral and written retelling and passage review as strategies for comprehending text. *Psychology in the Schools*, *47*, 135–152.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, *8*(2), 24–33. https://doi.org/10.1177/074193258700800206

Serafini, F. (2010). Reading multimodal texts: Perceptual, structural and ideological perspectives. *Children's Literature in Education*, *41*(2), 85–104. https://doi.org/10.1007/s10583-010-9100-5

Shanahan, T. (2011). Common Core Standards: Are we going to lower the fences or teach kids to climb? *Reading Today*, *29*(1), 20–21.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneier, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practice guide* (No. NCEE 2010-4038A). Washington, D.C: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=14

Sheehan, K. M., Kostin, I., Futagi, Y., & Flor, M. (2010). *Generating automatic text complexity classifications that are aligned with targeted text complexity standards* (ETS Research Report Series No. 2) (pp. i–44).

Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, *101*(2), 305–314. https://doi.org/10.1037/a0014217

Sipe, L. R. (1998). How picture books work: A semiotically framed theory of text-picture relationships. *Children's Literature in Education*, *29*(2).

Sipe, L. R. (2008). *Storytime: Young children's literary understanding in the classroom*. New York: Teachers College Press.

Snow, C. E., & Burns, S. M. (1998). *Preventing reading difficulties in young children*. Washington DC: National Academies Press.

Snow, C. E., & O'Connor, C. (2016). Close reading and far-reaching classroom discussion: Fostering a vital connection. *Journal of Education*, *196*(1), 1–8.

Snyder, E., Witmer, S. E., & Schmitt, H. (2016). English language learners and reading instruction: A review of the literature. *Preventing School Failure: Alternative Education for Children and Youth*, 1–10. https://doi.org/10.1080/1045988X.2016.1219301

Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal*, *53*(7), 410–413.

Sparks, R. L., Patton, J., & Murdoch, A. (1/2014). Early reading success and its relationship to reading achievement and reading volume: Replication of "10 years later." *Reading and Writing*, *27*(1), 189–211. https://doi.org/10.1007/s11145-013-9439-2

Spear-Swerling, L. (2015). A bridge too far?: Implications of the Common Core for students with different kinds of reading problems. *Perspectives on Language and Literacy*, *41*(2), 25–30.

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, *41*(1), 225–234. https://doi.org/10.1037/0012-1649.41.1.225

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72–110.

Stahl, S. A., & Heubach, K. M. (2005). Fluency-oriented reading instruction. *Journal of Literacy Research*, *37*(1), 25–60.

Stevens, R. J., Lu, X., Baker, D. P., Ray, M. N., Eckert, S. A., & Gamson, D. A. (2015). Assessing the cognitive demands of a century of reading curricula: An analysis of reading text and comprehension tasks from 1910 to 2000. *American Educational Research Journal*, *52*(3), 582–617. https://doi.org/10.3102/0002831215573531

Stillman, J., & Anderson, L. (2017). *Teaching for equity in complex times: Negotiating standards in a high-performing bilingual school*. New York, NY: Teachers College Press.

Taboada, A., Tonks, S. M., Wigfield, A., & Guthrie, J. T. (2009). Effects of motivational and cognitive variables on reading comprehension. *Reading and Writing*, *22*(1), 85–106. https://doi.org/10.1007/s11145-008-9133-y

Taylor, L. K., Alber, S. R., & Walker, D. W. (2002). The comparative effects of a modified self-questioning strategy and story mapping on the reading comprehension of elementary students with learning disabilities. *Journal of Behavioral Education*, *11*(2), 69–87.

Thomason, B., Brown, C. L., & Ward, N. (2016). Rigor or restriction: Examining close reading with high school English language learners. *TESOL Journal*. https://doi.org/10.1002/tesj.308

Topping, K. J., Samuels, J., & Paul, T. (2008). Independent reading: The relationship of challenge, non-fiction and gender to achievement. *British Educational Research Journal*, *34*(4), 505–524. https://doi.org/10.1080/01411920701609380

Treptow, M. A., Burns, M. K., & McComas, J. J. (2007). Reading at the frustration, instructional, and independent levels: The effects on students' reading comprehension and time on task. *School Psychology Review*, *36*(1), 159–166.

Turkan, S., & DaSilva Iddings, A. C. (2012). That child is a yellow: New immigrant children's conceptions of English language, literacy, and learners' identities in the NCLB era. *Theory into Practice*, *51*(4), 273–280. https://doi.org/10.1080/00405841.2012.726055

Umansky, I. M., & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal*, *51*(5), 879–912. https://doi.org/10.3102/0002831214545110

Unrau, N. J., & Quirk, M. (2014). Reading motivation and reading engagement: Clarifying commingled conceptions. *Reading Psychology*, *35*(3), 260–284. https://doi.org/10.1080/02702711.2012.684426

Valdés, G. (2001). *Learning and not learning English: Latino students in American schools*. New York: Teachers College Press.

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

Van den Branden, K. (2000). Does negotiation of meaning promote reading comprehension? A study of multilingual primary school classes. *Reading Research Quarterly*, *35*(3), 426–443.

VanDeWeghe, R. (2007). Research matters: What kinds of classroom discussion promote reading comprehension? *English Journal*, *96*(3), 86–91.

Vannest, K., Davis, J. L., & Parker, R. I. (2013). *Single case research in schools*. London: Routledge.

Vaughn, S., Mathes, P. G., Linan-Thompson, S., & Francis, D. J. (2005). Teaching English language learners at risk for reading disabilities to read: Putting research into practice. *Learning Disabilities Research & Practice*, *20*(1), 58–67.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walczyk, J. J., Wei, M., Grifith-Ross, D. A., Goubert, S. E., Cooper, A. L., & Zha, P. (2007). Development of the interplay between automatic processes and cognitive resources in reading. *Journal of Educational Psychology*, *99*, 867–887.

Wells, C. G. (1999). *Dialogic inquiry: towards a sociocultural practice and theory of education*. New York: Cambridge University Press.

Wertsch, J. V. (1998). *Mind as action*. New York, NY: Oxford University Press.

Wertsch, J. V., & Rogoff, B. (Eds.). (1984). *Children's learning in the "zone of proximal development."* San Francisco, CA: Jossey-Bass.

Wilhelm, J. (2015). Teaching texts to SOMEBODY! A case for interpretive complexity. *Voices from the Middle*, *22*(4), 44–46.

Williams, L. M., Hedrick, W. B., & Tuschlnski, L. (2008). Motivation: Going beyond testing to a lifetime of reading. *Childhood Education*, *84*(3), 135–141. https://doi.org/10.1080/00094056.2008.10522991

Williamson, G. L., Fitzgerald, J., & Stenner, A. J. (2013). The Common Core State Standards' quantitative text complexity trajectory: Figuring out how much complexity is enough. *Educational Researcher* , *42*(2), 59–69. https://doi.org/10.3102/0013189X12466695

Wong Fillmore, L. (2014). English Language Learners at the crossroads of educational reform. *TESOL Quarterly*, *48*(3), 624–632. https://doi.org/10.1002/tesq.174

Wong Fillmore, L., & Fillmore, C. (2012). What does text complexity mean for English learners and language minority students? Presented at the Understanding Language, Stanford, CA. Retrieved from http://ell.standford.edu

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *17*(2), 89–100.

Worsdell, A. S., Iwata, B. A., Dozier, C. L., Johnson, A. D., Neidert, P. L., & Thomason, J. L. (2005). Analysis of response repetition as an error-correction strategy during sight-word reading. *Journal of Applied Behavior Analysis*, *38*(4), 511–527. https://doi.org/10.1901/jaba.2005.115-04

Yesil-Dagli, U. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly*, *26*(1), 15–29. https://doi.org/10.1016/j.ecresq.2010.06.001

Yin, R. K. (2014). *Case study research: design and methods* (5th ed.). Los Angeles, CA: SAGE.

Zens, N. K., Gillon, G. T., & Moran, C. (2009). Effects of phonological awareness and semantic intervention on word-learning in children with SLI. *International Journal of Speech-Language Pathology*, *11*(6), 509–524. https://doi.org/10.3109/17549500902926881

Zhang, J., & Dougherty Stahl, K. A. (2011). Collaborative reasoning: Language-rich discussions for English learners. *The Reading Teacher*, *65*(4), 257–260. https://doi.org/10.1002/TRTR.01040

Zweirs, J., O'Hara, S., & Pritchard, R. (2014). Eight essential shifts for teaching new standards to academic English learners. Academic Language Development Network.

APPENDIX A

BOOKS USED WITH EACH GROUP

Book	Used for...
Arndt, I. (2014). *Best foot forward: Exploring feet, flippers, and claws*. New York City, NY: Holiday House.	Group 1, Matched Group 2, Difficult
Bishop, N. (2012). *Snakes*. New York City, NY: Scholastic Nonfiction.	Group 2, Difficult
Bishop, N. (2007). *Spiders*. New York City, NY: Scholastic Nonfiction.	Group 2, Difficult
Bradley, K. (2001). *Pop! A book about bubbles*. New York City, NY: HarperCollins.	Group 1, Matched
Chin, J. (2014). *Gravity*. Macmillan.	Group 2, Matched
Cowcher, H. (2009). *Antarctica*. New York City, NY: Square Fish.	Group 1, Matched
Cowley, J. (2006). *Red-eyed tree frog*. New York City, NY: Scholastic Paperbacks.	Group 2, Matched
de la Bedoyere, C. (2016). *Acorn to oak tree*. Irvine, CA: QEB Publishing.	Group 1, Matched
De la Bedoyere, C. (2012). *I am a frog*. Essex, United Kingdom: Miles Kelly Publishing Ltd.	Group 1, Difficult
Franco, B. (2002). *Amazing animals*. Chicago, IL: Children's Press.	Group 1, Matched
Fraser, M. A. (1998). *Where are the night animals?* New York City, NY: HarperCollins.	Group 1, Difficult
Jenkins, S. (2001). *What do you do when something wants to eat you?* Boston, MA: HMH Books for Young Readers.	Group 1, Difficult
Jenkins, S., & Page, R. (2008). *What do you do with a tail like this?* Boston, MA: HMH Books for Young Readers.	Group 1, Matched
Markle, S. (2014). *What if you had animal hair?* New York City, NY: Scholastic Paperbacks.	Group 1, Difficult
Newman, M. (2015). *Polar Bears*. New York City, NY: Square Fish.	Group 1, Difficult

Peterson, C., & Lundquist, D. R. (2012). *Seed, soil, sun: Earth's Recipe for Food*. Honesdale, PA: Boyds Mills Press.	Group 1, Difficult
Posada, M. (2000). *Dandelions: Stars in the grass*. Minneapolis, MN: Carolrhoda Books.	Group 2, Matched
Serafini, F. (2010). *Looking closely in the rain forest*. Toronto: Kids Can Press.	Group 1, Difficult
Sill, C. (2013). *About birds: A guide for children*. Atlanta, GA: Peachtree Publishers.	Group 2, Matched
Sill, C. (2017). *About fish: A guide for children*. Atlanta, GA: Peachtree Publishers.	Group 1, Matched
Simon, S. (2003). *Cool cars*. New York City, NY: Scholastic.	Group 1, Difficult
Squire, A. O. (2012). *Fossils*. Chicago, IL: Children's Press. (chapter 1 only)	Group 1, Difficult
Tatham, B. (2001). *Penguin chick*. New York City, NY: HarperCollins.	Group 1, Matched

APPENDIX B

METHODS SUMMARY

Qualitative data		
Measured outcome	*Definition (for purpose of this study)*	*Instruments & analysis*

Engagement	Students' social interaction and strategy use around texts in an effort to construct meaning and students' affective responses to text. (See Guthrie and Anderson 1999; Unrau and Quirk 2014).)	• A priori codes: strategy use, interaction for meaning, positive response, negative response
		• Additional codes developed in analysis:
		  • Engagement through talk (subcodes: about ideas, about pictures, about text)
		  • Subcodes developed under strategy use: background knowledge, cognates, connection to personal experience, connection to text, fix-up, genre, gesture, research, summarize
		  • Subcodes developed under interaction for meaning: agree, answer peer/self, answer teacher, ask, building on previous comment, disagree, introduce new topic, peer coaching
		• Students' ranking of book preferences
		• Count: code totals for each condition

Discussion participation	Students' verbal contributions to discussion	• Count: number of instances of verbal participation (Compare for each condition.) • A priori codes: interpretive responses, literal responses • Additional response sub-codes developed: incorrect

Quantitative data		
Comprehension	Understanding and constructing meaning from the ideas in text	• One minute timed retell: scored for number of words and number of correct idea propositions. (See Thomas 2012).
Fluency	Reading words accurately and automatically with good expression, phrasing, and prosody. (See Rasinski and Young 2014).	• One minute timed read: scored for words read correct per minute. (See Hasbrouck and Tindal 2006).

APPENDIX C

CODEBOOK

Code	Description
(RQ1, Engagement) Strategy use	
Strategy use: connection to text	Student connects text to other text (including multimodal texts such as television shows).
Strategy use: connection to personal experience	Student connects text to personal experience.
Strategy use: background knowledge	Student states relevant fact or previous learning connected to text.
Strategy use: monitor	Student says they understand or do not understand.
Strategy use: summary	Student summarizes text or portion of text.
Strategy use: fix-up (decoding)	Student troubleshoots or receives teacher support for an unknown word.
Strategy use: gesture	Student uses gesture to define/exemplify a term, estimate a measurement, show how something works, reenact text,
Strategy use: research	Student refers to peritextual or online material to get more information about a topic introduced in the main text.
Strategy use: cognates	With teacher prompting, student connects unknown English word to known Spanish word.
Strategy use: genre use	Student makes sense of text by reference to nonfiction/informational genre.
(RQ1, Engagement) Verbal interaction for meaning	
Interaction: ask	Student asks a question about text/topic/picture.
Interaction: answer peer/self	Student addresses a peer's question or addresses a question they raised themselves. Includes questions repeated or rephrased by teacher, but originally raised by student.
Interaction: answer teacher	Student answers the researcher's question.
Interaction: clarify	Student seeks clarification of what another participant has said.
Interaction: agree	Student agrees with another participant about discussion of text.

Interaction: disagree	Student disagrees with another participant about discussion of text.
Interaction: peer coaching	Student offers peer support such as coaching for unknown word or directing them to a page that addresses their question.
Interaction: building on previous comment	Student continues discussion by contributing an additional statement or elaborating a previous statement on the same topic.
Interaction: introduce new	Student continues discussion by contributing a statement on a new topic.
(RQ1, Engagement) Focus of Engagement	
Engaging with pictures through talk	Student talks or asks about pictures, with or without pointing at picture. Examples: commenting on what animals look like, expressing amazement at photo, asking about how something was drawn or photographed, asking what something in picture is
Engaging with text through talk	Student talks or asks about one or more exact words in the text. Student directly refers to text by saying, "it says" and then paraphrasing. Student references exact numbers from text. Student answers question with words from text. Examples: rereading, asking what a word means, quoting
Engaging with ideas through talk	Student talks or asks about idea introduced in text without specific reference to any particular text. Does not include one-word answers to teacher questions. Does not include introducing relevant background knowledge beyond ideas introduced in text. Examples: general questions, sharing connections
(RQ1, Engagement; RQ2 Discussion) Response	
Interpretive/inferential responses	Student makes comment that extends beyond information provided in text. Student asks inferential question, requiring thought beyond information provided in text. Student defines or clarifies a word. Does not include nonverbal responses.
Literal responses	Student restates idea or fact from text. Student asks "right there" question answered by text or other basic factual question even if the answer is not in the text. Does not

	include nonverbal responses. Does not include one-word answers to teacher questions.
Incorrect	Student makes literal or inferential statement/claim contradicted or not supported by book.

241

APPENDIX D

RESULTS BY BOOK

Best foot forward: Exploring feet, flippers, and claws (Level L)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	90	3
Jack	59	5
Rosa	66	2
Average	72	3.3
Penguin chick (Level L)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	88	4
Jack	62	6
Rosa	81	7
Average	77	5.7
Antarctica (Level L)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	81	10
Jack	50	4
Rosa	70	5
Average	67	6.3
Acorn to oak tree (Level L)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	106	4
Jack	66	9
Rosa	81	6
Average	84	6.3
What do you do with a tail like this? (Level L)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	110	6
Jack	81	5
Rosa	84	6
Average	92	5.7
Pop! A book about bubbles (Level K)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	102	6
Jack	61	5
Rosa	85	7
Average	83	6
Polar bears (Level N)		
	Fluency (wcpm)	Comprehension (ideas retold)

Alyssa	90	1
Jack	47	3
Rosa	67	4
Average	68	2.7
Fossils (Level N)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	76	3
Jack	49	4
Rosa	71	4
Average	65	3.7
What if you had animal hair? (Level N)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	93	5
Jack	58	6
Rosa	79	9
Average	77	6.7
Seed, soil, sun (Level N)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	90	4
Jack	42	2
Rosa	87	5
Average	73	3.7
Where are the night animals? (Level N)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	67	4
Jack	45	5
Rosa	44	7
Average	52	5.3
Looking closely in the rain forest (Level N)		
	Fluency (wcpm)	Comprehension (ideas retold)
Alyssa	101	7
Jack	71	7
Rosa	86	6
Average	86	6.7
Gravity		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	75	2
Gabriela	82	2
Elise	67	4

Average	75	2.7
About fish (Level I)		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	51	5
Gabriela	64	4
Elise	37	5
Average	51	4.7
Dandelions: Stars in the grass (Level K)		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	56	3
Gabriela	60	5
Elise	39	1
Average	52	3
About birds (Level I)		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	94	7
Gabriela	95	4
Elise	60	4
Average	83	5
Red-eyed tree frog (Level J)		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	55	4
Gabriela	88	5
Elise	59	1
Average	67	3.3
Amazing animals (Level F)		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	67	6
Gabriela	69	4
Elise	53	4
Average	63	4.7
Snakes (Level I)		
	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	44	3
Gabriela	45	5
Elise	31	5
Average	40	4.3
Best foot forward (Level L)		

	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	66	4
Gabriela	67	2
Elise	56	3
Average	63	3

What do you do when something wants to eat you? (Level K)

	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	43	3
Gabriela	52	3
Elise	26	4
Average	40	3.3

Cool cars (Level I)

	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	38	4
Gabriela	54	7
Elise	49	4
Average	47	5

I am a frog (Level H)

	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	51	7
Gabriela	63	3
Elise	55	3
Average	56	4.3

Spiders (Level I)

	Fluency (wcpm)	Comprehension (ideas retold)
Sarah	36	5
Gabriela	53	4
Elise	57	6
Average	49	5