Talking it Out: Scaffolding High Schoolers' Comprehension of Complex Texts

By

Dan Reynolds

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Learning, Teaching, and Diversity

October 31, 2017

Nashville, Tennessee

Approved:

Amanda Goodwin, Ph.D.

Deborah Rowe, Ph.D.

Rogers Hall, Ph.D.

Elfrieda Hiebert, Ph.D .

DEDICATION

To Laura, Mary, Bridget, and Jack

ACKNOWLEDGMENTS

This dissertation could not have been written without the patience, understanding, attention, and love of many people around me. First, the attention of my committee members—Dr. Rogers Hall, Dr. Debbie Rowe, and Dr. Freddy Hiebert—was invaluable, and each of them pushed me to strengthen different parts of this research. Most of all, the support of my advisor, Dr. Amanda Goodwin, provided me with the best combination of high expectations and high support over my four years at Vanderbilt. Thank you, Amanda.

The doctoral students, administrative staff, and professors in the Department of Teaching and Learning were wonderful supporters of my work, and I am a better educator for having been a part of DTL. Thanks are also due to the funders of the Bonsal Applied Education Research Award, without whose support I could never have dreamed of a project of this scope.

I would also like to thank the study's tutors as well as the teachers, and administration at the study high school, who helped organize the project. The tutoring team was a joy to meet every morning, and I am grateful for their work on behalf of this project.

I cannot forget the students of Redemptorist HS in Baton Rouge, LA and San Juan Diego Catholic HS in Austin, TX, where I taught before beginning doctoral work. They were my first teachers about reading, and in their spirit I continue to advocate for better, more responsive, and more human reading instruction for high school students.

Finally, I thank my family—my brothers, my parents, and my three children Mary, Bridget, and Jack. They are all a part of me. Of course, above all, I am ever-grateful for my wife, Laura, whose wisdom, expertise, and love steadied and guided me throug

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Educators have long grappled with the challenge of supporting students reading increasingly complex texts. A promising means of support is interactional scaffolding, defined as the responsive in-person support an expert provides to a novice reader to support the reader's comprehension (Athaneses & de Oliveira, 2014; Wood, Bruner, & Ross, 1976). Scaffolding for readers of complex texts is particularly relevant because of the widespread adoption of the Common Core State Standards (CCSS). Specifically, the CCSS argue that increasing the complexity of texts in reading instruction will better prepare K-12 students for the future demands of an increasingly information-driven society. Standard 10 of the CCSS outlines a three-part framework for assessing the complexity of texts, using qualitative, quantitative, and reader-task factors to establish a range of appropriate texts. Standard 10 states that students need to be able to read text in their grade band "proficiently, with scaffolding as needed at the high end of the range" (p. X; CCSS Appendix A, National Governors Association for Best Practices & Council of Chief State School Officers, 2010). Little, though, is included about what appropriate scaffolding should look like. This is likely because while research has examined how comprehension can be scaffolded, deeper understandings are needed about how to scaffold students' comprehension of complex texts—especially at the high school level. This dissertation addresses this need by analyzing an intervention in which small groups of high school students read complex texts with a tutor who scaffolded their comprehension. Results will show both the effectiveness of the scaffolding-based intervention at increasing student comprehension as well as links between types of scaffolding and student comprehension outcomes.

Understanding effective comprehension scaffolding is critical as many US high school students are reading below grade level and will need scaffolding to read texts at their grade level. In fact, on the most recent National Assessment of Educational Progress (NAEP) in reading, only 37% of high school seniors scored "proficient" or above—including only 7% of African-American and 12% of Hispanic seniors (National Center for Education Statistics, 2015). In addition, while fourth and eighth graders have made progress since NAEP reading tests began in 1992, twelfth graders' scores have decreased. Supporting high school readers should be an imminent concern for reading research.

Given these challenges, several researchers have voiced concerns about the implications of increases in text complexity without deeper understanding of effective scaffolding. For example, Hiebert and Mesmer (2013) contend that increasing text complexity without concomitant increases in teachers' ability to support readers could lead to frustrated and alienated readers—especially for the most vulnerable readers in US schools, such as English learners and struggling readers. Sanden (2014) echoes these concerns, asking, "In our zeal to create readers who *can* read, are we reducing the chances of creating students who *do* read?" (p. 8). Another concern involves the definition of complex texts: the language in Standard 10 devotes a great deal of attention and discussion to quantitative factors of text complexity such as word frequency and syntactic length, and far less attention to the host of contextual, developmental, and social variables that are important when considering how complex texts are to be read and scaffolded (Fisher & Frey, 2014; Goldman & Lee, 2014). In fact, Goldman and Lee (2014) argue that certain features of texts and tasks have strong impacts on how complex a text is for a particular reader. Therefore, that improving reading comprehension requires more than just using more challenging texts: the instruction must

respond to the contextual, developmental, and social factors that affect texts' complexity. Developing instructional approaches that account for these factors may help educators avoid the potential pitfalls and advance the intellectual promise of complex texts.

Contemporary Models of Reading and Literacy

Studies of scaffolding are particularly relevant now because of developments in contemporary models of literacy. Expanding on previous definitions from the *Becoming a Nation of Readers* report in 1985 and the RAND Reading Study Group in 2002, Frankel, Becker, Rowe and Pearson (2016) articulate key shifts in contemporary thinking about reading and literacy, and three are relevant to this study. First, the productive (writing, speaking) and receptive (reading, listening) dimensions of literacy are intertwined in literacy development. Consequently, instructional models must offer extensive opportunities for both. Second, contemporary models' increased emphasis on the social context of the reading requires an understanding of who students read with, why they read, and how others shape their emerging meaning-making. Third, research on the disciplinary nature of reading in the humanities, social sciences, natural sciences, and mathematics suggests that reading instruction must use texts from many disciplines and support students in tackling the different challenges posed by these texts.

A scaffolding-based intervention links well with these new models of reading. In such an intervention, students must constantly be engaged in productive processes to allow scaffolders to respond to their thinking. In this study, students constantly talked out the text (i.e. paraphrased it) as they read with their tutors, and their constant productive talk enabled tutors to support their emerging comprehension. In addition, focusing on interactional scaffolding attends to the social nature of students' emerging construction—in this study's

small groups, students' comprehension was not just the product of their individual reading but was shaped by their interactions with their tutors and peers. Finally, interactional scaffolding's responsive in-person support offers flexibility to support students across texts from many disciplines. Thus, this study's instructional model is aligned with contemporary models of literacy across developmental levels.

The Lack of Research about High School Readers

Though much research exists on student literacy and scaffolding, developing a scaffolding-based intervention for high school readers is particularly challenging because of the dearth of research about them. Seven years ago, Snow and Moje (2010) highlighted the contrast between high schoolers' NAEP static reading scores and the increasing scores of their elementary and middle school counterparts, and noted the "massive investment in primary grades literacy instruction while neglecting later literacy development" (p. 66). They critiqued the widespread belief that an early dose of protective reading instruction in elementary school can protect a student from future reading difficulties, naming this belief the "inoculation fallacy". They suggested that the literacy challenges of increasingly complex texts, disciplinary literacy, and academic language represents a unique set of challenges for middle and high school students, and therefore recommended that reading research tackle those challenges directly.

Despite this recommendation seven years ago, the lack of rigorous experimental research on interventions focused on improving high school reading is striking. Paul and Clarke (2016) conducted a literature review of randomized controlled trials aimed at improving middle and high school students' reading comprehension, checking over 10,000 initial search findings, but they found only eight such studies conducted since 1992. Slavin,

Cheung, Groff, and Lake (2008) reviewed over 300 studies of middle and high school reading in their search for rigorous studies of middle and high school reading programs and found only 33, only four of which used random assignment. The currently published research available on the What Works Clearinghouse database (2017) lists dozens of interventions with strong evidence of effectiveness for the preschool, kindergarten, and elementary level, but only four that extend to the high school level. Additionally, those interventions are broad interventions designed for grades K-10, 5-9, 4-10, and 5-12, but almost none of the interventions were tested in high school students and, in fact, the 11th and 12th grades are completely unrepresented in the WWC database. It appears, then, that there are few empirically-tested interventions specifically designed for typical high school students' reading comprehension[1]. Even more noteworthy is that the few studies defined as "high school" in all of the reviews above are concentrated in early high school (ninth and tenth grade): no studies in the Paul and Clarke (2016) and Slavin and colleagues (2008) reviews and in the What Works Clearinghouse database (2017) worked with students older than tenth grade. While strong non-experimental work has built a good foundation establishing principles for intervention design, reading research is devoid of rigorous experimental evidence for how to improve late high schoolers' reading comprehension.

The Current Study

Within the broader literature on reading research, research on interactional scaffolding also suffers from the inoculation fallacy pattern. A recent systematic review of research on interactional scaffolding and reading comprehension noted that of the 57 studies included,

[1] Three reviews of reading interventions for high school students with learning or reading disabilities (Edmonds et al., 2009; Wanzek et al., 2013; Reed, Sorrells, Cole, & Takakawa, 2012) noted studies designed for high school readers, but all of these studies were specifically designed for students with such disabilities.

only nine focused on high school readers (Reynolds, 2017), justifying a focus on high school readers. That review also coded the studies' experimental designs, finding that while 11 experimental studies of scaffolding verified the superiority of more-scaffolded instruction to less-scaffolded control groups, only one of those studies was conducted at the high school level (Lee, 1995), and none of those studies specifically conceptualized their texts as complex. Consequently, an intervention investigating whether scaffolding can be effective in supporting high school students reading complex text is needed. Finally, while the 11 experimental studies investigated *whether* scaffolding is effective, and 38 observational studies show *how* it can work, only five studies investigated whether certain *kinds* of scaffolding are associated with improved student comprehension—and none of those were conducted with students older than sixth grade. Because of this lack of evidence on scaffolding in high school coupled with the lack of experimental evidence on high school reading interventions, a study unraveling links between kinds of scaffolding and high school students' comprehension growth would be a strong addition to the literature. Therefore, to address these gaps in the research literature, this study was designed as both a rigorous experimental investigation of whether a scaffolding-based curriculum supports student comprehension and a correlational investigation to identify the kinds of scaffolding that are linked to student comprehension outcomes.

CHAPTER 2

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

This chapter specifies a theoretical framework that links three bodies of research: scaffolding, reading comprehension, and text complexity. Following that framework, the literature review examines the findings of prior research with the goal of informing the design of a scaffolding intervention. The chapter concludes with the study's research questions.

Theoretical Framework

Origins and theoretical roots of interactional scaffolding. Interactional scaffolding traces its roots to the term *scaffolding*, coined by Wood, Bruner, and Ross (1976) as the domain-general "process that enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts" (p. 90). Since then, the term has become lexically broad, and over the last four decades, education research has used many terms for similar forms of scaffolding across domains and theories of learning, which are summarized in Table 1. These terms fall into two categories: *interactional*, which is the responsive in-person support an expert provides to a novice, and *planned*, which refers to scaffolding determined before the student begins learning, such as lesson plans, computer programs, or curricular tools. Putanbekar and Hubscher (2005) argue that interactional scaffolding hews closer to Wood and colleagues' (1976) original definition of scaffolding because it focuses more directly on the responsivity of the scaffolder to the scaffoldee. Similarly, Reynolds and Daniel (2017) have called for research to focus on interactional scaffolding because the assumptions built into pre-planned scaffolding can inadvertently encourage scaffolders to overlook emergent student comprehension. Research, however, has also shown that responsive interactional scaffolding can be woven out of planned scaffolding

(Many, 2002) and thus research on interactional scaffolding must be carefully designed to separate it from planned scaffolding (Reynolds & Daniel, 2017). Though the intervention described in this study incorporated planned scaffolding to give the tutors a common set of texts and lesson plans from which to work, the intervention design and analytic focus is on the interactional scaffolding.

Even within studies focusing on interactional scaffolding, lexical breadth abounds. Table 1 summarizes the many different terms used for the same scaffolding constructs, each term emphasizing slightly different dimensions of the scaffolding. For example, *soft* (Brush & Saye, 2002) and *adaptive* (Azevedo & Hadwin, 2005) attend to scaffolding's malleability, *social* (Pea, 2004) and *direct interaction* (Applebee & Langer, 1983) emphasize scaffolding's human interactions, and *moment-to-moment* (Clark & Graves, 2005) and *temporary* (Roehler & Cantlon, 1997) focus on its temporal brevity. Though each of these terms carry certain connotations, *interactional* (Athaneses & de Oliveira, 2014) is used in this study because it concisely captures all of them, accounting for the malleable, social, temporal nature of the support.

***Theoretical grounding for scaffolding*.** The concept of interactional scaffolding is rooted in the work of Vygotsky (1978). Though he never used the term *scaffolding* in his writing, scholars have linked the term to his theories of learning (Belland, 2014; Pea, 2004; Putanbekar & Hubscher, 2005; Stone, 1993; Van de Pol, Volman & Beishuzen, 2010)[2]. Three key elements of a Vygotskyan view of learning are important to an understanding of how interactional scaffolding can help students read complex texts.

[2] In fact, even the authors of CCSS Standard 10 note that scaffolding is rooted in Vygotskyan conception of learning, although their description of scaffolding is limited to merely asserting that students already receive "considerable" amounts of it.

Table 1
Terms for Similar Forms of Scaffolding

Author	Field	Interactional term	Planned term
Pea (2004)	Science	Social	Technological
Azevedo (2005)	Science	Adaptive	Fixed
Stone (1998)	Special education	The dynamics of instruction	Preplanned activities or instructional materials
Brush and Saye (2002)	Social studies	Soft	Hard
Putanbekar and Hubscher (2005)	Science	Original notion of scaffolding	Evolved notion of scaffolding (cf. 2005)
Clark and Graves (2005)	Literacy	Moment-to-moment	Instructional frameworks
Applebee and Langer (1983)	Literacy	Direct interaction	Group-oriented instruction
Roehler and Cantlon (1997)	Literacy	Temporary	Permanent
Athaneses and de Oliveira (2014) [from Hammond & Gibbons, 2005]	Literacy	Interactional	Planned

First, Vygotsky's (1978) specification of the zone of proximal development (ZPD) is closely aligned with complex text instruction. He argued that "the only 'good learning' is that which is in advance of development" (p. 89). This suggests that increasingly challenging texts as a foundation for instruction could be catalysts for students' reading development.

Second, Vygotsky claims that learning in the ZPD is activated "only when a child is interacting with people in his environment" (p. 90). In this case, interactional scaffolding focuses squarely on the interaction between the expert and novice readers as an agent for the novices' development. This interactive, social process privileges the expert's ability to respond to the emerging needs of the novice as he or she progresses—and this responsiveness can modulate the texts' complexity to ensure that instruction remains within students' ZPDs. In their conceptual review of scaffolding, Van de Pol and colleagues (2010) identified the

construct of *contingency*: that is, the scaffolder's responsiveness to the scaffoldee's knowledge. Contingency is crucial, they argue, because without it, instruction is less able to ensure that learning is optimally within the students' ZPD. Others have argued that interactional scaffolding's in-the-moment responsiveness is more contingent than planned scaffolding, because planned scaffolding is designed for certain learners but not responsive to them in the moment (Athaneses & de Oliveira, 2014; Reynolds & Daniel, 2017).

Finally, Vygotsky also specifies that after learning interactions take place in the ZPD, the mental processes become internalized and become part of the child's independent developmental achievement and transfers to new tasks and settings. The evidence of the child's developmental achievement, however, can appear in multiple ways. For example, interactional scaffolding might support students' improvement in independently developing coherent interpretations of texts without assistance, which might be detected in traditional comprehension assessments. Alternatively, interactional scaffolding might support a student's improvement in using the language and procedures of strategic readers, which might then be seen in surveys or think-alouds used on independent texts. In addition, scaffolding might also support a student's ability to defend his or her interpretations against others', evidence of which might be visible in discourse analyses. Thus, studies of interactional scaffolding can use similar Vygotskyan frameworks for comprehension development but offer different kinds of evidence to support claims about the success of comprehension scaffolding.

Wertsch's extension of Vygotsky's ZPD. Though Vygotsky's work laid the foundations for understanding scaffolding, Wertsch (1984) argued that Vygotsky left underspecified three key elements of interactional scaffolding learning in the ZPD: situation definitions, semiotic mediation, and intersubjectivity. Lee's (1995) study of teachers

scaffolding underachieving African American high school seniors learning literary interpretation skills illustrates Wertsch's argument in the context of a comprehension scaffolding study[3].

First, the *situation definition* is the disparate initial representation of the learning tasks. For example, Lee's students, long schooled in traditional methods of didactic reading instruction, viewed their everyday talk as fundamentally distinct from literary interpretation. In contrast, Lee recognized that the rich discursive resources her students brought to the classroom, especially the practice of signifying—which makes use of literary tropes such as irony, double entendre, satire, and metaphor—were much like those necessary for literary interpretation. Resolving these contrasting situation definitions is the task of scaffolding.

These disparate representations are gradually reconciled through *semiotic mediation,* which is the process of experts and novices using tools and signs to mediate their learning interactions. Lee used her students' existing understandings of their own discourse as a bridge toward literary interpretation by converting the students' implicit strategies for understanding discourse into explicit strategies for comprehending texts. She began instruction by undertaking a collaborative study not of literature, but of signifying—a characteristic of African-American oral discourse. This decision to temporarily abandon the situation definition of traditional academic literary interpretation helped the students recognize the richness embedded in their everyday language. Then, after an in-depth analysis of everyday language, she then scaffolded the students toward sharing her representation of the task of literary interpretation by showing them how similar habits of interpretation could be applied

[3] Lee (2000) explicitly connected her instructional approach to Wertsch's constructs, and much of this section relies on those connections.

not just to oral discourse but to literary texts.

Finally, *intersubjectivity* is the end goal of scaffolding where the novice shares the expert's representation of the task. For instance, Lee's students eventually attended to imagery and figurative language in text just as they would in oral language, suggesting the students moved from their initially disparate situation definition to share Lee's view. Additionally, Lee's quantiative analysis of comprehension outcomes showed that her students' successful reading extended beyond classroom interactions and discussion to independent reading tasks. Thus, Lee's study demonstrates how interactional scaffolding aligned with Wertsch's conceptual framework can lead to improved comprehension that is visible on both independent outcome measures and in more sophisticated classroom discussions. Seeing Lee's work through the lens of Wertsch's frame emphasizes the importance of eliciting and responding to students' initial situation definitions, using mediational means that are closely aligned with students' existing knowledge, maintaining the teacher's rigorously defined situation (i.e. Lee's high expectations of literary analysis) and ultimately achieving an intersubjective vision of the task of reading comprehension and the mediational means for achieving it.

Connecting interactional scaffolding to reading comprehension pedagogies. Both the theoretical specification of Wertstch (1984) and an empirical review of the scaffolding research by Van de Pol and colleagues (2010) highlight the importance of understanding the intersubjective end goals and purposes of scaffolding. These authors point out the importance of the student and teacher defining *what* is being scaffolded, not just describing how the teacher does it. Similarly, an early article by Searle (1984), titled "Scaffolding: Who's Building Whose Building?" pointed out that the metaphor of scaffolding presupposes some

kind of building—but that it is not clear who is in control of the interaction, and whose ultimate goal structures the interaction. Searle presents an example of a first-grader who has brought a walkie-talkie in for show and tell; the transcript shows the teacher modelling how to ask questions, present details, and talk about the item, but it is not clear that the child was ever given any agency in speaking. Searle points out that instead of supporting the child in achieving his goal—telling his classmates what he sees is important about his walkie-talkie—the teacher instead changed the situation "into one in which the child was left to figure out the teacher's understanding and intentions" (1984, p. 482). The very definition, then, of scaffolding includes the potential for tension when the scaffolder and the scaffoldee have different visions of the end of the interaction.

This discussion of intersubjective ends of scaffolding by Wertsch (1984) and Van de Pol and colleagues (2010), however, remained domain-general. How, then, could this theoretical perspective integrate more specifically with theories of reading comprehension? A framework devised by Aukerman (2013) helps forge this missing link. This framework classifies reading comprehension pedagogies into three approaches: comprehension-as-outcome, comprehension-as-procedure, and comprehension-as-sensemaking. Understanding the different potential end goals of comprehension scaffolding can better support design considerations for how to investigate it.

***Comprehension-as-outcome*.** Aukerman (2013) defines comprehension-as-outcome pedagogies as those which strive to "produce readers who generate 'correct' reading of texts" (p. A3). At the outset of the learning interaction, the disparate situation definitions are clear: the student's interpretation of the meaning of the text does not match the expert's. Over the course of the interaction, the expert provides discrete pieces of knowledge, such as the

definitions of vocabulary words or background information, to guide the learner towards the scaffolder's accepted interpretation of the text. In addition to providing knowledge, teachers may focus on repairing student misconceptions by assessing and correcting nonstandard interpretations, often in an initiate-response-evaluate discursive structure in which the teacher checks to see if a student has made progress toward the correct reading by asking questions to evaluate that progress, and, if progress has not met his or her standard for matching the correct reading, steps in to supply any missing pieces of knowledge to move the child forward (Mehan, 1979). Because knowledge is envisioned in discrete units and with a proper way to be assembled, the learning in this instructional approach can be quickly and efficiently assessed through comprehension outcome assessments—that is, assessments that determine how well the student's interpretation of the text matches the generally accepted interpretation. Aukerman (2015), however, cautions educators that such an approach could over-focus on assessing students' reproduction of standard interpretations and potentially encourage teachers to overlook emerging comprehension in students' nonstandard interpretations that could contain meaningful insight into their thinking and learning process.

An example of this comprehension pedagogy in a study of scaffolding can be found in Crowe's (2003) study of oral reading feedback (i.e., scaffolding). This study defines the task of a reader as "reconstruct[ing] the author's meaning" (p. 18), and the task of the scaffolder as supporting the reader in reconstructing that meaning. In this case, the text has a fixed meaning that the author has embedded in the text, and it is the job of the reader to comprehend that meaning. For Crowe (2003), the scaffolder "assists in establishing the content of the author's message before reading, setting the scene, simplifying complex sentences, establishing relationships between and discussing or expanding upon unfamiliar vocabulary or

concepts in context" (p. 18). Here, the scaffolder assists the student towards the correct interpretation of the author's message in whatever way possible, and good scaffolders must be prepared with background knowledge and vocabulary to provide to students when it is needed. In sum, such a view of comprehension scaffolding requires the scaffolder to structure the scaffolding and orient his or her contingent support toward the outcome of the reading process: the scaffolder-sanctioned correct reading.

Comprehension-as-procedure. While the end goal of comprehension-as-outcome pedagogies is to produce a correct interpretation, the goal of comprehension-as-procedure pedagogies is producing a correct reading method (Aukerman, 2013). In Wertsch's terms, the reader's initial situation definition means that he or she does not use the reading strategies and methods used by the expert and other good readers. During the semiotic mediation of scaffolding, the expert models, demonstrates, and gradually transfers the strategies to the learner. Intersubjectivity, in this pedagogy, is achieved when the learner then uses those strategies independently. This paradigm is often expressed as cognitive strategy instruction. Harris and Pressley (1991) describe this pedagogy as "inform[ing] all students about cognitive tools often invented by or informally taught to outstanding readers" (p. 395). Thus, good comprehension scaffolding succeeds to the extent that the scaffolder's semiotic mediation enables the transfer of the strategies to the learner.

An example of this comprehension pedagogy in a study of scaffolding is Palincsar and Brown's (1984) Reciprocal Teaching (RT) intervention. The core work of RT involves the scaffolder's modeling, scaffolding, and transferring the practice of four key comprehension strategies used by good readers: summarizing, questioning, clarifying, and predicting. In fact, Palincsar and Brown defined "comprehension [as] the product of three main factors: (1)

considerate texts, (2) the compatibility of the reader's knowledge and the text content, and (3) the active strategies the reader employs to enhance understanding and retention, and to circumvent comprehension failures" (1984, p.18). Here, as in comprehension-as-outcome pedagogies, Palincsar and Brown fix the meaning to be comprehended as embedded within the text, as well as claim that correct background information is necessary for comprehension. Yet Palincsar and Brown go beyond comprehension-as-outcome approaches and include the active strategies component of comprehension. Consequently, scaffolding in RT prescribes "guided learning interactions in which the teacher could both model appropriate comprehension-fostering activities and at the same time guide the child to participate in an ever-increasing level of competence" (1984, p. 124). The scaffolding during Reciprocal Teaching lessons begins with the scaffolder modeling how to do the appropriate strategy, doing the strategy with the learners, and gradually transferring responsibility to them. In such comprehension-as-process pedagogies, the focus on the transfer of comprehension processes means that scaffolders should explicitly name the strategy and practice it with the students. Assessment, in this paradigm, focuses not just on whether the students' interpretation matches standard interpretations, but also whether they can use the appropriate strategies independently at the end of the scaffolding interaction. Consequently, studies employing such a comprehension pedagogy often measure both comprehension outcomes and increased strategy use.

Comprehension-as-sensemaking. Finally, a third comprehension pedagogy avoids reproducing specific readings or transferring specific comprehension strategies and instead views the ends of comprehension scaffolding as supporting emergent student sensemaking. Aukerman (2013) clarified that this approach can take two forms: a reader-response form

(Rosenblatt, 1982) in which readers form their own relationships with the text and offer their personal feelings about its meaning, or a dialogic form where the seeds of a reader's comprehension are sown in his or her personal interaction with the text but are refracted through dialogic interaction with others' interpretations (Bakhtin, 1994). The first form (i.e. self-contained sensemaking) focuses on the individual's own reactions to the text[4], but the second form (i.e. dialogic sensemaking) focuses on how readers' interpretations are transformed, enriched, and scaffolded by others. In fact, a recent study by Daniel and colleagues (2015) connected the dialogic form of comprehension-as-sensemaking pedagogy with a focus on contingent scaffolding. Both Aukerman (2013) and Daniel and colleagues (2015) argue that a sensemaking pedagogy honors students' own strategies for understanding discourse, challenges scaffolders to attend to emergent comprehension even in nonstandard interpretations, and helps scaffolders avoid deficit perspectives.

The Lee (1995) study discussed above also serves as a strong example of this pedagogy. Rather than specifying particular comprehension strategies, Lee's instruction focused on eliciting students' existing interpretive strategies for making sense of their own discourse. Then, her dialogic responses to students' visible thinking invited them to extend and refine their interpretations of their own discourse. Then, when the instruction shifted towards interpreting literary texts, the preexisting discourse forms were taken up as she used dialogic scaffolding of students' sensemaking to refine their interpretations of literature. As evidence of how such scaffolding worked to support student comprehension, Lee presents discourse analyses that show how the teacher's prompts such as "How do you know that?",

[4] Studies such as Anderson (1984) are exemplars—they note how, in isolation, readers' own schemas and histories can lead them to quite divergent readings of text based on their initial textual transactions.

"Why did you underline that?" invited students to make their existing sensemaking more visible to others in the classroom community, and students responded by defending their decisions using textual evidence (1995, p. 622-3). Lee speculated that "the attention to imagery and figurative language in which the students attend in their participating in signifying talk in their community experiences has now expanded to include attention to literary images in fiction" (p. 626). Thus, responses are evidence of how the sensemaking-oriented instruction encouraged students to use their own comprehension resources in order to generate more coherent and nuanced interpretations.

Implications of comprehension pedagogies for intervention design. Table 2 summarizes these three comprehension pedagogies and traces how they are operationalized in Wertsch's (1984) terms. For instructional purposes, distinguishing between these pedagogies is not meant to mark distinct boundaries between them—in fact, Aukerman (2013) notes that instruction can incorporate multiple pedagogies, and research has documented ways in which teachers negotiate the different comprehension pedagogies encouraged by high-stakes standardized tests, summative assessments of classroom units, and formative assessments used to guide instruction (Handsfield, Crumpler, & Dean 2010; Segal, Snell, & Lefstein, 2016). In theory, a truly contingent scaffolder might switch between comprehension pedagogies to respond to students' situation definitions. For example, students deeply schooled in comprehension-as-outcome pedagogies and used to the teacher providing background knowledge and vocabulary might require some transition before they can participate in a comprehension-as-sensemaking pedagogy.

Table 2
Linking Scaffolding to Pedagogies

	Location of meaning	**Theoretical tradition**	**Learner's initial situation definition**	**Semiotic mediation**	**Intersub-jectivity**
Outcome	In text, waiting to be understood by the reader	Instructivism and Core knowledge (Hirsch, 2003)	Interpretation does not match expert's	Expert provides and activates prior knowledge and vocab	Learner shares expert's interpretation.
Procedure	In ideal processes of good readers	Cognitive strategy instruction (Harris & Pressley, 1991; National Reading Panel, 2000)	Does not use strategies used by the expert/good readers	Through gradual release, expert demonstrates and transfers strategy	Learner uses expert's strategies
Sense-making	In learners' own interpretive processes	Dialogic pedagogies (Aukerman, 2013; Bakhtin, 1994)	Sensemaking does not produce internally coherent interpretation	Through dialogic discussion, experts push for coherence	Learner's repertoire expands to produce more coherent interpretation

One implication of this pedagogical framework, however, is that researchers working within each pedagogy have traditionally used different kinds of evidence to support claims about students' comprehension. For example, Reynolds (2017) found that studies using comprehension-as-sensemaking pedagogies were more likely to use observational methods and qualitative data analysis of teacher-student discourse to show comprehension as built between scaffolders and students, while studies using comprehension-as-outcome pedagogies were more likely to use experimental methods and quantitative data analyses of students' independent test scores. Because of the lack of rigorous experimental evidence for high school

students (Paul & Clarke, 2016; What Works Clearinghouse, 2017), this study focused on outcome assessment data, but future studies from this intervention will examine video of the instruction to see how students' discourse and language use may also demonstrate improvement in comprehension.

Connecting scaffolding and comprehension to text complexity. A discussion of how to scaffold readers' comprehension of complex texts would not be complete without considering factors that make texts complex for readers. Because Wertsch (1984) emphasized the importance of the mediational means which can enable or constrain learning, and because interactional scaffolding is affected by the planned scaffolding within which it occurs, scaffolding studies must consider how certain features of text can affect the scaffolding interactions, though scaffolding research has not generally done so (Reynolds, 2017).

Text complexity is a challenging field, and many frameworks exist to assess just how complex a text truly is. For example, the framework employed in the Common Core State Standards uses a three-part system incorporating qualitative factors determined by attentive human readers (e.g. levels of meaning, language conventionality), quantitative dimensions determined by software (e.g. word frequency or sentence length), and reader and task considerations determined by teachers and instructional designers (National Governors Association and Council of Chief State School Officers, Appendix A, 2010). The CCSS discussion of quantitative factors proposes several algorithms that assign values to a text's complexity, such as the Lexile system. This study's focus on contingent scaffolding (and not just reader-text matching), however, needed an approach to assessing text complexity that prioritized instructional value over text leveling. For a solution, Pearson and Hiebert (2014) argued that the purpose of qualitative features systems using rubrics and exemplars "is to

involve teachers in identifying text features that can promote (or impede) students' capacities to read a text, rather than assigning a specific level to a text" (p. 163). These systems place greater focus on the instructional implications of text complexity with less attention to quantifying linguistic features of text, and thus is consistent with the goal of this intervention. Similarly, Pearson and Hiebert (2014) point out that such systems "indicate the scaffolds and supports a teacher might need to provide in a given classroom to help a range of students" (p. 163). Identifying such scaffolds—and ascertaining how well they support comprehension outcomes—is one goal of this research.

The RSVP framework. Existing empirical studies of scaffolding offer little guidance about how scaffolding works with increasingly complex texts (Reynolds, 2017). Fortunately, Pearson and Hiebert's review of qualitative text systems for identifying instructional features of text identifies the RSVP system (ACT Inc., 2006) as the system that "shows the most potential to provide direction for teachers on how to scaffold texts that challenge students" (2014, p. 169). Adapting a scaffolding approach based on RSVP could support students in reading complex texts.

The ACT Inc. report (2006) argued that high school students' ability to read increasingly complex texts—and not their ability to handle increasingly difficult questions or answer questions about different elements of texts—showed the strongest relationship with overall reading scores. Consequently, the report specified six dimensions of text complexity and coined the acronym RSVP (technically, RRSSVP) to describe their text complexity framework:

- "Relationships (interactions among ideas or characters)
- Richness (amount and sophistication of information conveyed through data or literary

devices)

- Structure (how the texts is organized and how it progresses)
- Style (author's tone and use of language)
- Vocabulary (author's word choice)
- Purpose (author's intent in writing the text" (ACT Inc., 2006, p. 15).

While this six-part framework offers a strong foundation toward understanding scaffolding, and the report offers some exemplars of ACT texts annotated with particular elements of that complexity rubric, the report does not offer much detail or explanation about how each dimension was operationalized and measured. Consequently, it is difficult to see how instruction might align with these textual features. Fortunately, recent research on academic language by Uccelli, Phillips Galloway, Barr, Meneses, and Dobbs (2015) links to elements of the RSVP framework.

Using CALS to extend the RSVP approach. Though Uccelli and colleagues' (2015) work did not explicitly connect to the RSVP framework, significant correspondence exists between the two frameworks' conceptualization of how text complexity is operationalized in academic language. Uccelli and colleagues refer to the construct as Cognitive Academic Language Skills (CALS). Table 3 shows some of the overlap between the two systems. To illustrate one example of how CALS elaborates the simpler conceptions of RSVP, consider the RSVP category "Vocabulary" (the fourth row in Table 3). The ACT report goes no further than suggesting that vocabulary is a dimension of text complexity, but the CALS study conceptualizes vocabulary dimensions of academic language as both how students can unpack morphologically complex words as well as recognizing that certain vocabulary words (both morphologically simple and morphologically complex) can take on particular definitions,

meanings, and connotations in academic texts that are rarely found in colloquial language. Students instructionally supported in recognizing the polysemy of words used in both academic and everyday language would, theoretically, be better prepared to comprehend complex academic texts. For another example, the RSVP construct of relationships focuses on the "interactions among ideas or characters in the text" (ACT, Inc., 2006; p. 7). Similarly, the CALS construct about connecting ideas logically through attending to discourse markers (e.g. *although*) refers to students' ability to see how ideas interact within the author's argument. For another example, the CALS construct about tracking participants refers to students' skills in resolving conceptual anaphoric and cataphoric referents—that is, helping them remember who is doing what, and how the relationships between agents and actions plays out over the course of a text. This aligns closely with the RSVP dimension of "Relationships". Similar overlap with CALS constructs exists across the other four constructs of the RSVP. Viewed together, these frameworks offer guidance for how scaffolding can attend to complex features and academic language commonly found in school texts.

Table 3

Aligning the RSVP and CALS frameworks

RSVP (ACT 2006)	**Core Academic Language Skills (Uccelli et al., 2015)[5]**
Relationships	Connecting ideas logically Tracking participants and themes
Richness	Unpacking complex words Unpacking complex sentences
Style	Recognizing academic register
Vocabulary	Unpacking complex words Recognizing academic register
Purpose Structure	Organizing argumentative texts Connecting ideas logically

[5] See Appendix A of Uccelli et al (2015) for more elaborate description of these constructs.

Literature Review

Designing an intervention requires establishing a coherent theoretical framework as well as synthesizing findings from prior published studies about interactional scaffolding, reading comprehension, and text complexity. The following section outlines the strengths and gaps in the research designs that have been used in the literature and specifies key recurring themes that are relevant to designing a scaffolding-based intervention (Reynolds, 2017).

Existing research designs and gaps in the research. While existing research offers a strong foundation for understanding scaffolding, several important gaps exist in the research (Reynolds, 2017). Observational studies constitute the majority of studies of interactional scaffolding (38 of 57). Many of these studies offer in-depth examinations of expert teachers' scaffolding in their regular teaching practice, and they show what interactional scaffolding looks like across different developmental levels and in classroom settings (e.g. Aukerman, 2007; Christoph & Nystrand, 2001; Lee, 1995; Many, 2002). While these studies describe how teachers enact contingent scaffolding and draw on positive classroom cultures, they often have little to say about comprehension outcomes. In addition to observational studies, somewhat less common are the 11 experimental studies. These studies find nearly universal positive evidence of comprehension growth, both on researcher-designed (e.g. Alfassi, 1998) and standardized measures (Crowe, 2003; Lysynchuk, Pressley & Vye, 1990). Though the evidence is strongly positive from these existing studies, their relatively small sample sizes (only Brabham & Lynch-Brown, 2002 included a treatment group of 60 or more students) means that scaffolding has not been examined across larger groups of students. In addition, the experimental comparisons all compare more-scaffolded treatment groups to less-scaffolded control groups; these designs show that scaffolding can improve comprehension,

but do not offer insight into what kinds of scaffolding work best.

Correlational designs may be able to offer some insight into the kinds of scaffolding that are most effective, but few studies exist: three studies combined correlation with observational methods and two were stand-alone correlational studies. The mixed- methods studies analyze large corpuses of talk to show the importance of teachers' talk moves—showing that teachers' scaffolding talk predicted the quality of students' textual discussions even when the teacher was absent (Jadallah et al., 2011; Lin et al., 2015). The stand-alone studies also focus on the importance of teacher talk, showing that teachers' scaffolding talk that reduces cognitive challenge was linked to decreased standardized reading comprehension scores (McElhone, 2012), and that tutors' motivational scaffolding predicted growth on a standardized comprehension assessment (Reynolds & Goodwin, 2016b). Overall, correlational designs can examine the effectiveness of different kinds of scaffolding across larger groups of teachers and students (e.g. McElhone, 2012; Reynolds & Goodwin, 2016b) and across large samples of talk (e.g. Jadallah et al., 2011., and Lin et al., 2015). These designs suggest the importance of teacher talk for good scaffolding, and that scaffolders should maintain high engagement and high cognitive challenge during the interaction. More research employing correlational designs could confirm these findings and investigate other potential pathways for scaffolding's effectiveness.

Beyond research design, Reynolds (2017) also documented gaps in the research regarding text complexity. Only four of the 57 studies incorporated texts that could be considered complex (i.e. above students' reading levels), and only one of these was conducted at the high school level (Moss, Lapp & O'Shea, 2011). In addition, none of the four studies devoted extensive written space to conceptualizing the complexity of the texts. Because the

various dimensions of complex texts can present unique challenges to readers, research should address this gap.

Themes in existing research that inform intervention design. Current research indicates interactional scaffolding supports student comprehension, although questions remain when considering how to best scaffold successful high school reading of complex texts. Three key themes informing intervention design are found from the literature: scaffolding should be contingent on student thinking, scaffolding depends on contextual factors, and scaffolding should avoid specific types of pitfalls (Reynolds, 2017).

The importance of contingency. In a review of the literature on scaffolding across different domains of learning, Van de Pol and colleagues (2010) identified contingency on student thinking as a key feature of good scaffolding in general, and Reynolds (2017) found this was true of the interactional scaffolding of reading comprehension as well. For example, studies show that students bring rich interpersonal and background knowledge with them as they read, and that expert scaffolders can elicit that knowledge and use contingent scaffolding productively in the service of reading comprehension (Lee, 1995, 2001; Moll, Neff, Amanti, & Gonzalez, 1992). For example, Athaneses and de Oliveira's (2014) comparative case study at a high school serving Latino students demonstrated how a teacher successfully drew on students' cultural and community knowledge in her scaffolding. Using this knowledge enriched her students' comprehension more than the routinized scaffolding approach of a comparison teacher. In this case, the more successful teacher was more contingent—she responded to student knowledge rather than relying on routines. Similarly, Wortham's (1995) study of teachers in an urban high school teaching a Great Books program showed how teachers used analogies to everyday life and participant examples to build bridges between

students' everyday knowledge and the literary worlds of the texts. The teachers capitalized on what students already knew about relationships and human motivations to contingently transform the classical Great Books texts from alien to more accessible and relatable—without dumbing them down. These examples suggest that contingent scaffolding can productively leverage students' existing knowledge toward reading comprehension.

Research also shows that effective, contingent scaffolders build on students' responses, using careful discourse moves to encourage students to elaborate and extend their responses, thus enriching comprehension. Three qualitative analyses of expert scaffolders document how teachers probe for student knowledge with careful questioning, and then build on student responses to extend meaningful discussion (Fisher & Frey, 2010; Gaskins, Rauch, Gensemer, & Cunicelli, 1997; Palincsar, 1986). In addition, three quantitative discourse analyses show similar findings. Mariage (1995) and Boyd and Rubin (2006) found that increased comprehension was linked to scaffolders' contingently building on students' contributions. Similarly, McElhone (2012) studied discourse and reading comprehension in 21 classrooms, finding that teachers' reduction of conceptual press (i.e. accepting students' initial answers without asking them to extend them) negatively predicted reading comprehension. Taken together, research suggests that good scaffolding both draws on students' existing knowledge and extends students' spoken responses to scaffold toward deep comprehension.

Contextual influences on interactional scaffolding. Research also shows that scaffolding is shaped by the contextual factors of classroom cultures, material tools, and disciplinary goals. In fact, nearly all the studies of interactional scaffolding conducted in classrooms emphasize the importance of classroom culture to successful interactional

scaffolding (e.g. Aukerman, 2007; Celani, McIntyre, & Rightmyer, 2006; Christoph & Nystrand, 2001; Many, 2002; Maloch, 2004). These studies show how teachers must build cultures of safety and belonging so that students are willing to make their emergent thinking public in front of their teacher and peers. Several studies suggest that teachers who avoid immediately evaluating student responses can help build positive classroom culture (Aukerman, 2007; Celani, McIntyre, & Rightmyer, 2006; Christoph & Nystrand, 2001; Many, 2002; Maloch, 2004). For example, Aukerman (2007) closely analyzed an episode where a classroom teacher consciously avoided evaluating a student's response, and how students in that class slowly began to rely more on their own reasoning and evidence rather than looking to the teacher to approve their interpretations. From this research, it appears that scaffolders who build quality relationships and whose classroom cultures privilege students' thinking with less attention to evaluation are creating a social and epistemic capital that can be used towards productive scaffolding.

Beyond the cultural and interpersonal dimensions, scaffolding is also shaped by the material tools and disciplinary goals of the instruction. For example, Meskill (2005) documented how, in a classroom with a curriculum organized around computer-based activities, interactional scaffolding built on students' computer work, extended their thinking, and enriched their overall comprehension. Similarly, Daniel, Martin-Beltran, Peercy and Silverman used a design-research intervention study (2015) to showed how, in an early iteration of a reading curriculum, the question cards used to spark discussion inadvertently constrained scaffolding and limited students to yes-or-no answers, but once the question cards were revised in a later iteration of the intervention, the materials enabled scaffolding of rich discussion of textual meanings and deep comprehension. The material tools used during

scaffolding instruction are often aligned with the disciplinary goals of the instructional setting, and research shows that the disciplinary course context also shapes interactional scaffolding, especially in literature (Lee, 1995, 2001; Moss, Lapp, & O'Shea, 2011; Wortham, 1995) and social studies (Athaneses & de Oliveira, 2014; Reisman, 2015). For example, studies in literature suggest that asking students to enter the unfamiliar worlds of classic works of literature supports the disciplinary goal of literary interpretation (Moss, Lapp, & O'Shea, 2011; Wortham, 1995), while the teachers' authoritative reviewing of historical content and facts scaffolded towards comprehension and critical understanding of historical texts (Reisman, 2015). These studies show how good scaffolding works with material tools and disciplinary goals to build toward student comprehension.

Potential pitfalls. While research provides portraits of exemplary scaffolding, it also documents potential pitfalls of scaffolding when scaffolders do not maintain rigorous comprehension expectations or struggle to organize classroom discourse. For example, several studies have documented that teachers, even those who aim for high comprehension expectations, can inadvertently enact comprehension as basic literacy word-reading tasks or simple recall questions instead of deeper understanding (McElhone, 2012; Poole, 2008; Mertzmann, 2007). This may be because research also shows that even experienced teachers with the support of professional development initiatives may not be able to implement the challenging discursive practices that exemplify good interactional scaffolding. For example, Hacker and Tenant (2002) studied a three-year implementation of Reciprocal Teaching and found that teachers struggled to release the discussion to their students, instead retaining authoritative control over the discussions and avoiding contingent scaffolding or extending student talk. Several case studies of individual teachers also document the challenge of

teachers transitioning from a classroom discourse that evaluates student comprehension toward one that scaffolds it (Billings & Fitzgerald, 2002; Christoph & Nystrand, 2001; Maloch, 2002). Though other studies document how good classroom culture supports scaffolding (e.g. Maloch, 2004), these case studies of teachers suggest that even well-intentioned teachers must be willing to rethink their classroom discourse if they hope to achieve the promise of scaffolding toward comprehension. Overall, the existing literature's findings about the importance of scaffolders' contingently responding to student thinking, attending to contextual factors, and avoiding particular pitfalls form a foundation for good scaffolding design in an intervention.

Research Questions

While substantial research exists on scaffolding, reading comprehension, and text complexity, no existing framework has integrated these three conceptual bodies of research. Integrating them here deepens understanding of how interactional scaffolding can support comprehension of complex texts. The current study uses that framework to investigate how interactional scaffolding can support high school readers in reading complex text. The research questions are:

- Does a scaffolding-based intervention using complex texts improve high schoolers' comprehension additionally when compared to a comparison group? Does the effect of the intervention vary among subgroups of students? [RQ1]
- When adolescent readers read complex texts in small groups with scaffolders' support, what kinds of scaffolding predict growth in student reading comprehension? [RQ2]

CHAPTER 3

METHODOLOGY

Research Design, Site Selection, and Counterfactual

Translating the theoretical framework into instructional design requires careful attention to implementation. The initial stages of the research design process focused on the four elements of the theoretical framework displayed in the left column of Table 4: contingency, planned scaffolding, mediational means, and situation definitions. The right column shows how the research design accounts for each of these theoretical considerations.

Table 4

Linking scaffolding theory to intervention design

Theoretical scaffolding feature	Consequent design choice
Interactional scaffolding contingently responds to emergent comprehension.	Scaffolds identified based on theory, but applied (or not) as tutors see fit
Interactional scaffolding is affected by planned scaffolding	Holding curriculum constant allows comparisons across groups
Interactional scaffolding depends on the mediational means used	Text selection is crucial—and holding texts and other mediational means constant allows comparisons across groups and texts
Students' situation definitions have to be immediately and consistently visible to permit contingency	Curriculum must prioritize making student comprehension visible

The next step of the research design was site selection. Given the gaps in the research literature about typical high school reading, a large public high school in the southeastern US was selected for its diversity of race, socioeconomic status, and reading ability. The school had several academic tracks, including both regular academics and an International Baccalaureate track with many high-achieving students on track for admission to selective universities, but it also had only a 32% proficiency rate on the state's standardized test for

English III in 2015[6]. In addition, an initial power analysis suggested seeking out a sample size of approximately 150 students (75 treatment, 75 control), and the school's size was sufficient to permit such a study. Therefore, site selection emphasized diversity in student race, socioeconomic status, and reading achievement scores in a typical public high school setting large enough to detect a meaningful effect on student reading and conduct a larger-scale experiment than many existing studies of scaffolding.

An additional design concern supported the site selection—its willingness to participate in a regular-practice study design that would not rely on student volunteers. Because sixteen- and seventeen-year-old high school students are generally less likely than elementary school students to return consent forms sent home to their parents, it can be harder to conduct studies of high school reading[7]. Consequently, permission was sought and received from the school district's research office as well as the Institutional Review Board to conduct the study using de-identified data from students participating in the study instruction. As the study's instruction was structurally similar to what the students would normally receive during the school day, parental consent would not be required. Not requiring parental consent enabled both a larger sample size than a consenting study would have and guarded against selection bias, as there would be reason to think that students returning parental consent forms would be systematically different than those who did not return forms.

To answer RQ1 and address the lack of rigorous experimental evidence of reading interventions for late high school readers (Paul & Clarke, 2016; What Works Clearinghouse,

[6] This 32% figure is could be misleading as students in the IB Programme do not take English III and thus do not take the English III end-of-course exam. Thus, the 32% passing rate represents not the whole school, but the students not on the IB track.

[7] Beyond the inoculation fallacy, this may be another factor contributing to the imbalance of research between the elementary and secondary levels.

2017), the study employed a pretest-posttest randomized controlled trial design. To provide a valid experimental comparison group, the study was conducted during the school's daily 37-minute advisory period. During the advisory period while the treatment group students were receiving the small-group tutoring instruction, the comparison group students were engaged in computer-based self-paced ACT preparation materials from the website ePrep.com. This activity consisted of students taking practice tests, watching videos targeted by the software's algorithms to the students' learning needs, and doing additional practice exercises. The program bills its instructional approach to students as:

> "an online program that allows you to take full length exams with instant feedback. After you have completed the tests online or offline, ePrep instantly grades your tests and provides video lessons from an expert tutor for every single test question. If you want more, we provide more video lessons that drill deeper into the core concepts covered on the exams" (ePrep.com, 2017).

The individualized and student-directed nature of the computerized instruction meant that each student may have covered different subject matter during comparison instruction. Interestingly, the ePrep program appears to strive for contingency—showing video content tailored to the questions students were missing on the practice test. This contingency, however, used the student's multiple-choice answers to standardized test data to individualize their instruction, while the treatment group tutors' contingent instruction relied on students' direct paraphrases of the text to support their scaffolding. While students were working on the ePrep software, the advisory teachers were present and coordinating their activities, but these teachers did not deliver any of the instruction, and thus there was no human-delivered interactional scaffolding in the comparison condition. Therefore, this experiment's causal

counterfactual argument compares small-group interactional scaffolding while reading complex texts to what would have happened in its absence—individualized computer-based test-prep instruction. Teacher recruitment was conducted by the author and the school's principal sent out emails and visited advisory teachers' classrooms inviting them to participate in the study. Eight of the fifteen eleventh-grade advisory teachers agreed to participate.

Participants

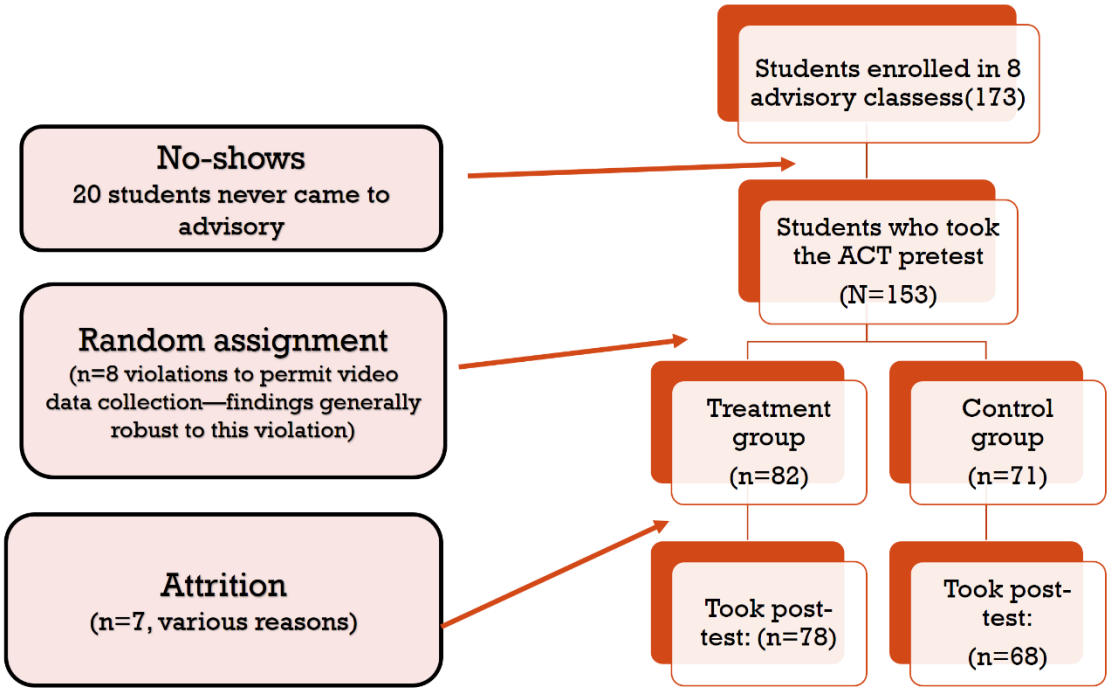
Students. Student participants were 153 eleventh-graders drawn from eight advisory teachers' classes. The students were 44% White, 50% Black, 2% Hispanic, and 4% Asian-American; 29% were eligible for free or reduced-price lunch, 5% spoke another language at home, fewer than 1% were receiving special services as ELLs, and 5% were classified as having some form of learning disability. These data confirm that the site selection succeeded in selecting a general population of students that had racial and socioeconomic diversity, though with less linguistic variability.

The details of experimental design, assignment, and attrition are summarized in Figure 1. The eight advisories officially enrolled 173 students, but only 153 came to the advisory during the study, and so the sample was limited to those 153 students who took the ACT pretest. After pretesting and as a safeguard toward equalizing the conditions' reading abilities across the intervention and comparison groups, students were blocked into pairs of roughly equal pretest scores and then one student from each block was randomly assigned to treatment and the other to the comparison condition. The treatment group was slightly larger than the comparison group due to eight violations of random assignment to permit robust video data collection. Since video data collection required parental consent and only 40 of the 153 students returned their consent forms, eight students (5.2% of the sample) who returned forms

were assigned to treatment without being subject to random assignment. Sensitivity analyses will examine results both with and without these students to investigate whether the violation of random assignment affected the results (Murnane & Willett, 2011). Attrition in both experimental and comparison groups was low and equally distributed across the two groups, with the reasons for the seven students' attrition ranging from moving out of the city, switching schools, being under out-of-school suspension, and, in one case, refusal to participate in posttests. Students in both conditions were compensated $10 for participation.

Figure 1

Experimental design, sampling, assignment, and attrition



Tutors. Nine tutors were selected to provide a range of experience levels typical of the ranges seen in US schools. Tutors' education levels, teaching experience, and number of study groups and students taught are summarized in Table 5. Three of the tutors (tutors 1-3 in Table

5) were highly experienced high school teachers, each with a master's degree and seven years of high school classroom teaching experience and all with at least four additional years of education experience in tutoring, curriculum design, or university teaching. Two of the three experienced tutors also had experience in high school administration. Four of the tutors (tutors 4-7 in Table 5) were education students with some practicum training and relatively little classroom experience (two or fewer years of experience as a teacher of record). Finally, two tutors (tutors 8-9) were well-educated medical researchers completing MD/PhD degrees but with no classroom experience. In addition, two additional coders and substitute tutors had middle and high school teaching experience and were pursuing PhDs in literacy education. Students were spread across tutors, with no tutor teaching over 23% of the students.

Table 5

Tutors' experience levels and study teaching loads

Tutor	Degree	Teaching years	Other education experience	Groups taught	Students taught
1	PhD Reading Education*	7	4 university + tutoring	4	17
2	M. Ed.	7	4 tutoring	1	4
3	M. Ed.	7	4+ Curriculum design	1	4
4	M. Ed.*	1	1 practicum	3	10
5	M. Ed.*	0	1 practicum	2	9
6	M. Ed.*	2	1 practicum	1	5
7	B.S. Education*	0	1 practicum	1	4
8	MD/PhD*	0	0	2	9
9	MD/PhD*	0	2 tutoring	4	18

*=degree in progress

Procedure

Instructional Design.

Groups. To facilitate close-knit interactional scaffolding between tutors and students and permit a high degree of contingency, but to also conduct the experiment at sufficient scale per the power analysis, students were assigned to small groups of 2-5 students (average size 3.1). The groups were assigned to produce heterogeneous groups by ordering the students within their advisories from highest to lowest ACT pretest scores, then assigning the highest scorer to one group, the second highest to the next group, and continuing until each group had one high-scoring student. Then, the next highest scoring student was assigned to the first group, and so on. This produced groups that had heterogeneous ability but prevented extreme heterogeneity (i.e. no group had both the highest- and lowest-achieving students). Once formed, groups were then randomly assigned to tutors. Assignment was generally done within advisory groups. The students received an eight-session intervention curriculum, and treatment group attendance ranged from 0-8 sessions, with an average of 5.7 sessions attended for students assigned to the tutoring groups.

Lesson plans. The focal activity of the lesson plans was paraphrasing the text—putting the text into their own words aloud—which was chosen for several reasons. First, a scaffolding-based intervention requires instructional activities that elicit students' situation definitions so scaffolders can respond (Wertsch, 1984). Second, paraphrasing the text balanced students' productive speaking activity (i.e. generating responses via talk) with their receptive reading activity (i.e. taking in information about the text; Frankel et al., 2016). Third, putting the text in their own words explicitly values the linguistic and cultural resources that students bring to the reading interaction (Lee, 1995). Additionally, since few students

were ELLs, linguistic proficiency was not a concern and it was assumed that all students had the requisite linguistic skill to participate in the instruction.

The empirical literature justifying paraphrasing as a focal activity is modest but positive. Research has documented positive correlations between paraphrasing skill and reading comprehension (Glover et al. 1981; Haynes & Fillmer, 1984; Crain-Thoreson, Lippman, & McClendon-Maguson, 1997). Single-group interventions with elementary (Hagaman & Reid, 2008; Hagaman, Casey & Reid, 2012), high school (Lauterbach & Bender, 1995), and postsecondary students (Hua et al., 2014) also show improvement when students are trained to paraphrase. One experiment also shows the superiority of a paraphrasing training intervention to business as usual reading instruction, reporting an statistically significant adjusted effect size of 0.48[8] on a researcher-designed 10-question multiple choice comprehension assessment (Katims & Harris, 1997). Thus, sufficient empirical evidence exists to support the selection of paraphrasing as a focal instructional activity.

The lesson plans were designed to consider the key elements of the theoretical framework and the findings from prior research. The lessons included attention to conversation norms and interpersonal relationships to build a positive small-group culture that would facilitate good scaffolding (Reynolds, 2017). The lessons also incorporated the transfer of responsibility, another key element of scaffolding, by having the tutor read the passage aloud to the students sentence-by-sentence in the first session, but then slowly transferring the responsibility to students reading to each other in chunks with lengths of their own choice[9].

[8] The study did not report an effect size, but did report means and standard deviations, so Cohen's *d* was calculated using that information. The adjusted effect size refers to the slight initial differences between the treatment and comparison conditions.

[9] Due to the need for students to make their situation definitions (i.e. their emerging comprehension) visible, they did not read silently or to themselves during the intervention.

Still, the tutors could maintain contingency and if students were struggling to read the text aloud (which was the case for one particularly complex text in lesson #7), they could step in and support the students as necessary by reading aloud while still encouraging for full transfer of responsibility to students. In addition, tutors were required to use a planned scaffold of introducing the text's "big question" (e.g. "What landforms are on the bottom of the ocean?") in the first four lessons, but were instructed not to use that planned scaffold in the final four sessions unless their students needed additional contingent support.

All tutors worked from the same lesson plans and were required to implement the same framework of planned scaffolding. Because of the importance of classroom culture to scaffolding (Reynolds, 2017), each day's lesson began with a norms-building conversation that encouraged tutors to build friendly relationships with their students and specified norms for turn-taking during the scaffolded paraphrasing. Three conversation norms were specified:

- Everyone takes turns talking out text
- Taking guesses and making mistakes is part of learning
- No talking down others' response: build on them.

Following these norms helped the talk-intensive instruction go smoothly. After building norms and relationships, tutors would show a sheet printing with the "Big Question" of the day that prepared students for the textual content, and discuss what students already knew about the topic. To limit the amount of pre-reading and maximize reading time, the norms-building and the Big Question took about the first 5-7 minutes of each session (decreasing as the intervention went on), and the reading took about 20-25 minutes. Then, tutors passed out a copy of the day's text to each student; these copies were the only materials used by the students during the intervention.

After building group norms and preparing to read, the groups then participated in the focal activity: scaffolded paraphrasing. At first in sessions 1-3, tutors briefly modeled how to do the paraphrasing strategy, gradually releasing so that sessions 4-8 contained only a cursory review of how to do so. After this review, the next activity was the scaffolded paraphrasing of the complex texts, in which one student would read a section of text (a sentence, sometimes two or three) and another student would paraphrase it, with tutors providing contingent scaffolding as needed. Each lesson then closed with a discussion that summarized what was learned during reading. Keeping the curricular framework fixed across tutors and groups permitted the direct investigation of the interactional scaffolding. An example lesson plan, from the third of the eight lessons, is included in Appendix A. Bold text was the suggested script—while tutors were encouraged to use consistent language, the tutors were not required to read from the script so they could talk naturally to the students. Unbolded text were instructions on how to implement that component of the lesson.

To assess whether the planned scaffolding curriculum was implemented with fidelity, each of the daily lesson plan components (building norms, introducing text, reviewing paraphrasing, scaffolded paraphrasing, and concluding discussion) had a check-box on the lesson plan. At the end of the lesson, tutors checked the boxes according to whether they had completed each of the components of the lesson, recorded how many lines of text they had read with the students, and noted the contents of the closing discussion. Tutor-observers, coding for fidelity of implementation, recorded the same information. This was done to verify whether the tutors' self-reports of implementation fidelity were reliable.

Texts. The nature of the intervention required short, complex, self-contained passages. Short texts were needed because of the limited instructional time in each session, and self-

contained passages were preferable as students who missed a session would not have to catch up on earlier content. In addition, because the RSVP and CALS frameworks suggested that addressing text complexity via academic language scaffolding could work across multiple academic disciplines, a diverse diet of student reading material was needed. The solution was to use texts from prior years' ACT exams, which are multidisciplinary self-contained texts of 700-800 words (ACT, Inc. 2010). Additionally, these texts offered the advantages of being aligned to the RSVP text complexity framework, and these texts also offered some motivation for the eleventh graders as they could see that these were the types of texts they would be reading on college entrance examinations and potentially in future college classes, and that gave many students an additional purpose for reading beyond learning content[10]. However, the instruction was *not* test preparation: the tutors never discussed any test preparation strategies, nor were any of the test questions ever shown to the students during instruction. Also, the dialogue of the tutors was focused on how to improve students' reading of complex texts, not how to beat a test.

Identifying complex texts was important both to provide a foundation for the instruction that would be challenging (i.e. in the ZPD for most students), and to see whether scaffolding could support students of widely varying reading abilities in reading complex texts. If contingent scaffolding could, as theorized, modulate the relative complexity of the texts, such texts—with scaffolding support—could be a powerful way for students to tackle challenging reading. Published reports show that ACT passages vary in complexity over three categories ("Uncomplicated", "More Challenging", and "Complex"), but the ACT Inc.,

[10] Though the texts may have had the opposite effect on students who were convinced that they would not take the ACT or attend college. Ultimately, the scale of these motivational effects are not clear.

ratings of complexity are not published. Therefore, to select complex passages, 32 passages were reviewed by the author per the RSVP guidelines to determine which would be considered complex (ACT, 2010). Each passage was rated from 1-3 (1=Uncomplicated, 2=more challenging, 3=complex) on each of the six categories. The passages with the highest averages were considered the most complex.

After reviewing the texts, selecting which would be used in instruction required consideration of the instructional sequence. Passages in each of the four disciplines used on the ACT that met the RSVP's criteria as "complex" were identified: prose fiction, social science, humanities, and natural science. At the recommendation of one of study teachers (i.e. a teacher who knew her advisory students well) and to provide accessible entry points for students while they learned to paraphrase and established group discussion norms, "more challenging" passages (highlighted in green in Table 6) was used in each of the first two sessions. The ensuing six sessions, constituting the bulk of the intervention reading, all used texts rated complex (highlighted in red in Table 6).

Table 6

RSVP Complexity Rubric (ACT, 2010) and Ratings for the 32 ACT Passages Reviewed and Eight Passages Selected for Instruction

	Degree of Text Complexity		
Aspect of Text	**Uncomplicated**	**More Challenging**	**Complex**
Relationships	Basic, straightforward	Sometimes implicit	Subtle, involved, deeply embedded
Richness	Minimal/limited	Moderate/more detailed	Sizable/highly sophisticated
Structure	Simple, conventional	More involved	Elaborate, sometimes unconventional
Style	Plain, accessible	Richer, less plain	Often intricate
Vocabulary	Familiar	Some difficult, context-dependent words	Demanding, highly context dependent
Purpose	Clear	Conveyed with some subtlety	Implicit, sometimes ambiguous

Test Form	Genre	Informal Passage Title	Relationships	Richness	Structure	Style	Vocabulary	Purpose	Average
72C	Prose Fiction	Rushdie	1	2	2	3	3	1	2.00
72C	Social Science	Atlantic	3	2	3	3	3	3	**2.83**
72C	Humanities	Bradbury	1	3	2	2	1	2	1.83
72C	Natural Science	Trapjaw Ants	1	1	1	2	2	1	1.33
72F	Prose Fiction	Stones for Ibarra	3	3	2	2	2	1	2.17
72F	Social Science	Taxonomy	2	1	2	3	2	2	2.00
72F	Humanities	Salinger	1	1	1	1	1	2	1.17
72F	Natural Science	Ants nesting	2	1	1	1	1	2	1.33
1MC	Prose Fiction	Linda Rose	3	3	1	2	1	3	2.17
1MC	Social Science	Government	2	1	3	1	3	1	1.83
1MC	Humanities	Indian-American	3	3	3	3	3	3	**3.00**
1MC	Natural Science	Dumb dinos	3	2	1	3	2	1	**2.00**
2MC	Prose Fiction	Saigon to VA	1	1	2	1	2	1	1.33
2MC	Social Science	Jefferson	3	2	2	3	3	3	**2.67**
2MC	Humanities	Star Trek	1	1	1	2	2	1	1.33
2MC	Natural Science	Color theory	3	1	2	3	3	3	**2.50**
3MC	Prose Fiction	Ted	2	3	2	2	1	2	2.00

3MC	Social Science	No waste economy	3	1	2	2	3	1	2.00
3MC	Humanities	Shakespearean CDs	1	1	1	1	2	1	1.17
3MC	Natural Science	Martian Water	3	1	3	2	2	1	2.00
64E	Prose Fiction	Dreaming woman	3	3	2	2	1	3	**2.33**
64E	Social Science	The little ice age	3	3	1	2	3	2	2.33
64E	Humanities	Louis Armstrong	2	3	1	2	2	1	1.83
64E	Natural Science	Acupuncture	2	1	1	1	2	2	1.50
66F	Prose Fiction	Jazz show	2	3	1	2	3	1	**2.00**
66F	Social Science	Sprawl	2	2	1	1	1	1	1.33
66F	Humanities	Field guide	2	2	1	1	2	2	1.67
66F	Natural Science	Snowflakes	2	2	1	1	2	2	1.67
67A	Prose Fiction	Winter wheat	2	3	2	1	1	3	2.00
67A	Social Science	Colosseum	3	1	3	1	2	3	2.14
67A	Humanities	South Asian music	3	3	3	2	3	2	**2.71**
67A	Natural Science	Squid	3	3	2	2	3	3	2.71

**Note: The two green passages were the "more challenging" passages used in sessions 1 and 2, while the red-highlighted passages were the "complex" passages used in sessions 3-8.

To triangulate these selections and confirm their complexity, the Qualitative Assessment of Text Difficulty was consulted (QATD; Chall, Bissex, Conrad, & Harris-Sharples, 1996). The QATD goes beyond the RSVP's domain-general complexity framework and incorporates discipline-specific elements of complexity. Specifically, the QATD pays special attention to how issues of narration (e.g. unreliable narrators and multiple perspectives) enhance text complexity in humanities and prose fiction texts, and how methodological issues (e.g. the limits and ambiguities of scientific and historical knowledge, the historical evolution of researchers' methods) enhance text complexity in the social and natural sciences. For example, the passage presented in Appendix B, is an account of historical oceanography (i.e. how humans have investigated the topography on the ocean floor),

Not only does the passage describe current oceanography and undersea geography, it also includes a historical discussion of how ordinary sailors and undersea-cable-layers have aided the emerging process of *how* the science of undersea topography has evolved. In this way,

complex texts like these refer to not only complex content, but also methodological issues in science over time. The passage's complex purpose is not just about history and science, but how historical and scientific knowledge is made.

Similarly, the QATD confirmed that complex content in social science and history emphasizes methodological dimensions. The "Jefferson" passage used in lesson #7, a discussion of whether Jefferson plagiarized any or all of the Declaration of Independence, certainly meets the QATD's complexity requirements for discipline-specific vocabulary and sentence structure. Like the discussion of the evolution of undersea topography, the passage wrestles with how historical views about Jefferson evolved over time, and how contemporary historians have moved beyond idealized portraits of a Founding Father and reconstructed an increasingly more historically accurate Jeffersonian narrative through examining his other writings, his public statements about the Declaration, and contemporary events. Ultimately, the passage comes to an ambiguous conclusion about the potential plagiarism. This conclusion acknowledges both the complexity of its central question and because the limitations of historical study and the evidence base. As in science passages, the complexity of the passage does not merely recount what happened, but also explains historians' process of knowing what happened.

The "Atlantic" and "Jefferson" passages illustrate how the QATD rubric leveled each of the RSVP-selected texts as an "13th grade/early college" level, which confirmed the texts as appropriately complex for the goal of the intervention in finding out whether reading complex texts with intensive support would be effective reading instruction. The titles of the works from which the texts were excerpted and one sample text ("Atlantic") used in the third lesson are presented in Appendix B.

The ten classes of scaffolds. Based on the theoretical framework and prior research, a list of scaffolds was developed from which tutors could select to contingently support students' thinking. The scaffolds were grouped into ten classes aligned with their theoretical links: extending talk, background knowledge, morphology, academic register, syntax, connections, tracking participants, structure, motivation, mediators, and rereading. Examples are summarized in Table 7, and a full list of scaffolds is presented in Appendix C.

Table 7

Classes and examples of scaffolds

Research Links	**Scaffolds**	**#**	**Example**
Vygotsky (1978)	Extending talk	6	"Can you extend your thinking?" "Where is the evidence for your paraphrase?"
Hirsch (2003)	Background knowledge	5	Tutor provides definition of vocabulary word. Tutor provides his or her interpretation of a sentence.
RSVP/CALS (ACT 2010; Uccelli et al., 2015)	Morphology	4	"What do the parts of this word tell you?"
	Academic register	5	"What does this word mean in *this* context?"
	Syntax	3	"How do the different parts of the sentence work together?"
	Connections	3	"How has (character/idea) changed so far?"
	Tracking participants	3	"Who did that action?" "Who does this pronoun/nominalization refer to?"
	Structure	4	"What is the structure of this text?"
Reynolds & Goodwin (2016b)	Motivation	7	"I like how you did ___" "You can do this on your own—you don't need me"
Wertsch (1984)	Mediators	7	Tutor acts texts out with hands or draws a sketch to demonstrate.
	Rereading	2	Tutor re-reads sentence or asks student to re-read.

The six extending talk scaffolds were broad prompts where tutors asked students to elaborate on their thinking. Grounded in the interactive view of scaffolding, they are also supported by research that shows positive benefits of teachers asking students to extend their comprehension talk through encouraging students to elaborate their ideas, provide evidence for their claims, or clarify their perspectives (Boyd & Rubin, 2006; Jadallah et al., 2011). McElhone (2012) called this phenomenon "conceptual press" and noted that teachers who did not engage in it during textual discussions were associated with lower student comprehension scores.

The five scaffolds classified as background knowledge scaffolds offered tutors the opportunity to provide definitions of vocabulary words or interpretations of sentences, paragraphs, or the entire passage. These scaffolds are linked to an outcome view of scaffolding that sees content knowledge as the most important link to comprehension (Hirsch, 2003), and thus were not aligned with the sensemaking nature of the intervention. Still, these scaffolds were included as students may have brought their expectations of the tutors as knowledge providers, and additionally they could help tutors could control students' frustration (Wood, Bruner, & Ross, 1976). Sometimes, students just wanted tutors to tell them what a word meant. Additionally, tutors could have used these if they felt that providing authoritative knowledge would stabilize the content under discussion and therefore facilitate additional sensemaking (Reisman, 2015).

Morphology scaffolds were supports in which tutors could prompt students to attend to the morphological nature of words in the complex text. Morphological interventions have been shown to have positive effects on literacy (Goodwin & Ahn, 2010, 2013). In addition, breaking down complex words is a dimension of the CALS academic language framework

(Uccelli et al., 2015). For these scaffolds, tutors might support students paraphrasing a sentence with the relatively rare word *unroughened* by asking students to identify the root word *rough*.

Another CALS dimension inspired the next class of scaffolds: recognizing the academic register. These scaffolds allowed tutors to highlight the ways in which everyday words can take on different meanings in academic registers, thus encouraging students to take flexible approaches to interpreting words based on context clues. For example, as a student paraphrased the phrase "the grand and sweeping energies hidden below [the surface of the ocean]", a tutor might support by highlighting the academic register in which the word *sweeping* is used here and asking what it might mean in this context. Similarly, tutors might point out the ways in which authors use punctuation such as dashes, quotation marks, or italics to add emphasis or connotations to ordinary words, shifting their contextual meanings. In this way, the scaffolds go beyond merely recognizing the academic register (the CALS assessment skill) and support students in clarifying the potential meanings of academic register discourse in context.

A third group of scaffolds, labeled as syntax scaffolds, align with the CALS skill of unpacking complex sentences. Research has shown that syntactic comprehension makes a unique contribution to overall passage comprehension (Poulsen & Gravgaard, 2016; Scott, 2009). In these scaffolds, tutors could prompt students to recognize how different clauses work together, or how interruptions work within a sentence. Tutors could also highlight transition words that clarify syntactic relationships, such as conjunctions. For example, students faced with paraphrasing the 39-word sentence: "*People assumed that this "Middle Ground," "Telegraph Plateau," or "Dolphin Rise," as it was variously called, was an*

ancient and drowned land bridge, or a lost continent, but sailors repairing transatlantic telegraph cable unknowingly produced evidence to prove otherwise"[11] might ask students to first paraphrase the first part of the sentence (before the *but*), and then the second part (after the *but)* as a way to support students in unpacking complex sentences. These scaffolds are reading-specific adaptations of what Wood and colleagues (1976) called scaffolds which reduce in degrees of freedom, as they show how tutors could support struggling students by breaking a problem (in this case, a sentence) down into smaller parts.

The CALS framework also specifies making logical connections across ideas as an important dimension of academic language of complex texts, a finding echoed by other recent research (Crosson & Lesaux, 2013; Welie, Schoonen, Kuiken, & Van den Bergh, 2016). In these scaffolds, tutors asked questions like "How has this character/idea changed so far?" or "Is this similar to or different from what we saw earlier in this text?" A specific example from the text in Appendix B could be the word "Today" at the beginning of the fourth paragraph, which marks a transition in the chronological organization of the paragraph, showing how historical methods for describing the topography on the ocean floor have culminated in modern sonar. Overall, this group of scaffolds was theorized to support students in building coherent interpretations that account for the ways ideas and characters changed over the course of the text.

The fifth dimension of the CALS construct includes tracking participants within texts. Since complex texts often include tricky referents such as nominalizations that obscure actions and pronouns that refer to several sentences earlier in the text (Snow, 2010), helping students unravel these referents could support their comprehension (Sanchez & Garcia, 2009).

[11] Ssee Appendix B for the passage from which this sentence was excerpted

In addition, since academic language often contains abstract entities acting as agents, tracking participants might help students clarify just who is performing a particular action (Snow & Uccelli, 2009). At times, this might even point out textual ambiguities. For example, in the phrase "highly sophisticated sound waves bring the hazy images of those early soundings into sharp focus" (Appendix B, paragraph 4), the grammatical subject of the sentence is *sound waves*, an abstract entity. Scaffolders using the tracking participants scaffolds could have chosen to point out that it is unclear just who is using the sound waves to create the images. Together, the scaffolds were created to support students in determining, when possible, the actions taken by persons in the text as it unfolded.

The sixth and final CALS category of scaffolds is labeled structure, and it links to the CALS dimension of organizing argumentative texts and to research suggesting the links between understanding text structure and reading comprehension (Hebert, Bohaty, Nelson & Brown, 2016; Meyer, Brandt, & Bluth, 1980). When using these scaffolds, tutors would highlight structural features of the text or ask students to consider how the text was organized. This could include asking broad questions about the whole text like "What is the structure of this text?" or narrower questions like "How does this sentence fit in the structure?" Thus, a tutor might have encouraged students to consider how the text in Appendix B was structured both chronologically, describing the evolution of methods of determining ocean-floor topography, and comparatively, contrasting ocean-floor topography with land topography.

Since scaffolding includes both cognitive and motivational dimensions (Belland, Hannafin, & Kim, 2013, Wood, Brunger & Ross, 1976), a group of motivational scaffolds was included, patterned after research that showed their effectiveness (Reynolds & Goodwin, 2016b). These scaffolds offered tutors several ways to encourage students' participation.

Examples include general praise, specific praise for student thinking moves, or refocusing students' attention.

Two final groups of scaffolds—labeled mediators and rereading—are theoretically based on Wertsch (1984). Because the semiotic mediation of talk may not always be sufficient to convey particular meanings, the mediator scaffolds offered tutors the chance to use hand gestures, drawings, concept maps, or external media to convey meaning. This was included as activating other channels of communicating might facilitate reaching intersubjective understandings of text support students' conceptual understandings (Glenberg, 2011; Wertsch, 1984). The rereading scaffolds were just that: at times, tutors could simply re-read a portion of the text or ask students to do so. Re-reading the text could help mediate the interaction by establishing just which piece of text was being paraphrased or scaffolded, as well as by clarifying word pronunciations or syntactic prosody (Paige, Rasinksi, & Magpuri-Lavell, 2012; Rasinski, Rikli, & Johnston, 2009).

Tutor training. After designing the intervention, selecting the texts, and developing the scaffolds, the tutors had to be trained to implement the intervention. Tutors completed three two-hour sessions of training before teaching[12]. The training used the intervention texts and scaffolds described above as well as video from a pilot test of the materials done by the author with a small group of high school students. The packet of materials given to the tutors during training is reproduced in Appendix D.

Training day 1: Understanding paraphrasing and contingent scaffolding. The first day of training focused on the core instructional activity: contingently responding to students'

[12] Four late-joining tutors got one-on-one training from the researcher and were able to complete the assignments and the training in three hours.

paraphrases. First, tutors participated in teambuilding exercises and were given a brief overview of the theory, terms, and goals of the intervention, with a focus on the nature of contingency and its importance to the intervention. Then, they were given a sample text from the intervention and the list of scaffolds and asked to identify potential scaffolding opportunities linked to the demands of the complex text. In discussion, tutors found key transition words, words used in academic registers, structural changes in the author's argument, and connections across paragraphs—all scaffolding opportunities linked to the theoretical framework. Next, the tutors watched a video of a tutor delivering the instruction with a pilot group of students, and, using a transcript of the interaction during the pilot instruction, they discussed whether the tutor had responded contingently to the students' emerging comprehension. For example, the pilot video included students paraphrasing the following sentence from the text in Appendix B: "*Today, sophisticated sound waves bring those early sound waves into sharp focus, revealing that one of the largest and most salient geographic features on the planet lies on the floor of the ocean."*. The a selection of the transcript from Appendix B reports about 20 seconds of the interaction (starting at the 16:17 mark) a tutor scaffolds the paraphrasing of students A, L, and K :

> [16:17] A: Means a large something under the ocean--
> [16:21] Tutor: There's a large somethin'...
> [16:24] Tutor: .. good.... what about that "Today"? What about the sound waves?
> [16:29] L: They were highly sophisticated
> [16:30] Tutor: Yeah... what does that mean? Let's explain this to a six-year-old. A six-year old might not know what "sophisticated" means
> [16:37] K: Interesting, smart, educated
> [16:39] Tutor: Yeah, really complicated, alright... so let's put it all together, we got some sound waves, we got the hazy images of those early soundings into sharp focus and we revealed there's somethin' ...

When the team of tutor trainees discussed the transcript, they noted that the tutor focused on

the connective word “Today” as a way to support students in recognizing the chronological structure of the text, but also noted that the “Today” did not respond directly to A’s comment about “large something under the ocean”. This discussion revealed to the tutors that contingency could take many forms—including both asking students to elaborate on information they provide in the initial paraphrases, but also challenging them to incorporate additional textual information (e.g. “Today”) not visible in their first attempt. Either of these approaches would be considered enhancing student sensemaking. Thus, the team defined practicing contingency not as striving for a single right scaffold applied to a student response, but a way of teaching that strove to respond directly to student knowledge in some form that was appropriate in the moment.

The team’s discussion also addressed the tutor’s move at 16:39. After K paraphrased “sophisticated” as “interesting, smart, educated”, the tutor responded “Yeah, really complicated”. The team of trainees discussed whether the tutor’s scaffold was contingent on K’s response, or whether his response merely provided a definition of the word in context, without responding to K’s paraphrase (i.e. was not contingent), or whether his response helped clarify K’s conception of the word so she could then respond to his scaffold in the next part of the sentence at 16:39 about putting it all together to paraphrase the whole sentence. Again, the team discussed how contingency in comprehension scaffolding might look different for different students based on the tutor’s responsive in-the-moment assessment of the student’s comprehension needs. In a case like this, where the student’s initial response (i.e. “a large something under the ocean”) includes few of the details from the text (i.e. omitted “sound waves”, “hazy focus”, and “today”), the tutor may have needed to provide extensive support to enable the students to produce a coherent interpretation. That is,

operationalizing contingency suggests that the type and degree of support vary according to students' responses.

After discussing how contingent scaffolding could be enacted using transcripts from the pilot video, tutors engaged in practice teaching with one another using intervention texts and a partner, and after teach partner practiced the teaching, they discussed which scaffolds they had selected and how they had tried to respond contingently. At the end of the first day's session, the tutors' homework was to make a short video of themselves practicing a short sequence of scaffolding using the one of the intervention texts with a friend or family member pretending to be a high school student. The videos were reviewed and oral feedback was provided by the author during the time before the second session, Feedback included commentary on correct implementation of the planned scaffolding as well as comments on successful contingent scaffolding and opportunities for additional contingency.

Training day 2: Delivering the intervention. The second day of training focused on implementing the structure of the intervention with fidelity. After a brief teambuilding exercise, the team collaboratively reviewed two videos from the homework assignment, comparing how different tutors had responded contingently to their students' paraphrases, and seeing how contingency was operationalized differently. For example, one tutor, when her student paraphrased only the first half of a long, complicated sentence, asked him to incorporate the second half of the sentence into his paraphrase, and the team agreed that that was contingent scaffolding (Scaffold #6 in Appendix D, in the extending talk group). Another tutor, when her student produced a relatively complete paraphrase of a sentence, contingently scaffolded by pushing her student to consider why the author had included that detail at that point in the text (Scaffold #36 in Appendix D, in the Structure group). The team discussed

how a tutor responded to a different kind of student knowledge with a different and contingent scaffold.

After discussing the homework videos, the tutors were introduced to the eight-lesson structure of the intervention, reviewed a sample lesson plan and the texts, and discussed the nature of the transfer of responsibility that was built into the planned scaffolding dimension of the intervention. Tutors were introduced to the way experimental fidelity would be maintained, and how the scaffolding would be self-reported with observers' reports measuring the self-reports' reliability. Then, tutors engaged in another practice teaching with partners, this time using the lesson plans and practicing self-reporting the scaffolds used in each session. Their homework was to record a full practice 37-minute lesson with a friend or family member and to share the video with the author to determine if the planned scaffolding framework was being correctly implemented. The author provided similar feedback for the second video as for the first.

Training day 3: Practicing reliable coding. On the third day of training, tutors practiced reliable coding. Informed by the practice sessions and homework teaching, team discussions were held to come to consensus about how the potential different ways each particular scaffold could be operationalized, which strengthened the team's shared definition of each scaffold. The team watched another video of the author scaffolding and identified scaffolds in it.

Finally, tutors assumed the roles of coders and watched a full 37-minute lesson of the author teaching the intervention, coding the implementation of each segment of the lesson plan (i.e. the planned scaffolding) as well as identifying the scaffolds used in that session (i.e. the interactional scaffolding).

Tutor supervision. Ongoing supervision during the intervention supplemented the initial training. When tutors were observed by another member of the research team, the tutor and observer would each submit independent reports of that session's scaffolding, but would then conference to see if they had agreed, and the ensuing discussion strengthened their sense of agreement about the components of the instruction and the definitions of the scaffolds. In addition, after each day's instruction, the team gathered to discuss issues that had come up during instruction—a post-teaching conference that helped ensure consistent implementation. As one more fidelity safeguard, for five tutors whose groups were being videotaped, the researcher reviewed one video in the middle of the intervention (after lessons 3 or 4) and provided feedback to the tutor. Feedback included pointing out instances of clearly contingent scaffolding and praising positive rapport between tutor and students as well as noting instances when the tutors may have failed to give the students space to develop their own paraphrases or instances where tutors were over-scaffolding by insisting that students replicate tutors' interpretations.

Sample scaffolding sequence. To illustrate what this contingent scaffolding looked like across, Table 8 presents transcripts of enacted scaffolding by two different tutors at the same point in the same text (excerpted from the second paragraph of the sample text in Appendix B). The top row presents the common text, and the two columns display the tutor-student interactional scaffolding. The students were reading to answer the big question, "What landforms are on the bottom of the ocean?". In these cases, the tutors were not required to scaffold at this particular point, but given the initial responses provided by students in turns 1 and 6, tutors elected to step in and scaffold.

Table 8.

Sample scaffolding sequence showing two tutors' contingent scaffolding

Common Text			
Only one thousand miles offshore, the *Cramer* has already sailed through some of Atlantic's deepest waters. Contrary to what one might guess, Atlantic's deepest waters, like those in other oceans, are along her edges. **As we continue east toward the middle of the sea, the bottom rises. The unmarked plains of the abyss here, flattened by layers of sediment, give way to rising foothills, then to mountains."**			
Turn	Scaffolding in Group 1	Turn	Scaffolding in Group 2
1	Student 1: As they were moving the water rises.	6	Student 3: As they continue to travel there's stuff going on under the sea.
2	Tutor 1: Not necessarily the water—what rises? (Scaffold: Tracking participants)	7	Tutor 2: What do you mean by stuff? What is the stuff that's going on? (Scaffold: Extending talk)
3	S1: The bottom	8	S3: It's calm on the top but it's not calm under.
4	T1: The bottom, good. And then in the next sentence, we see, what do we see about the bottom? (Scaffold: Connections)	9	T2: What about the ... there are mountains down there. What's going on with the mountains? […] (Scaffold: Extending talk)
5	S2: Describes them as mountains and foothills.	10	S3: I guess maybe just mountains are forming under?

In turns 1 and 6, students present two different emerging comprehensions as their paraphrase. As their paraphrases do not account for all the elements in the bolded text, the scaffolders respond contingently in turns 2 and 7. In Group 1, tutor 1 uses a "tracking participants" scaffold to challenge the student to account for who was doing the action. On the other hand, Tutor 2 responds asking student 3 to extend her talk, pushing her to more coherently account for the "stuff" she had paraphrased. A similar pattern of tutor contingency in turns 4 and 9 show Tutor 1 again using an academic-language scaffold asking her student to make logical connections across sentences, and Tutor 2 using an extending-talk scaffold to encourage Student 3 to incorporate more of the text into her response.

Consistent with the sensemaking model, the students' final responses in turns 5 and 10, when combined with their initial responses, are more coherent interpretations than their initial responses alone in turns 1 and 6, because they account for more features of the text. In Group 1, Student 1's initial paraphrase in turn 1 is supplemented in turn 5 by Student 2's incorporation of the landforms on the bottom of the ocean, producing an overall more coherent interpretation. In Group 2, Student 3's initial paraphrase with the vague "stuff going on" develops into the more coherent account of "mountains are forming". While no student's final paraphrase accounts for every connotation or feature of the original text—for example, neither group mentions potential implications of the text mentioning "layers of sediment"—the tutors were practicing sensemaking orientation by pushing students' coherence without insisting on the students replicating a standard reading.

Measures

Comprehension. Research has shown comprehension-as-sensemaking instruction can lead to detectable changes in outcome measures (Aukerman, Martin, Gargani, & McCallum, 2016; Kong & Pearson, 2003; Lee, 1995; McElhone, 2012). This may be because outcome measures attempt to assess a student's ability to produce a coherent interpretation (albeit a standard one), just as sensemaking pedagogies do. With this in mind, two outcome measures of comprehension were selected—one tapping passage-level reading comprehension, and one tapping sentence-level reading efficiency.

ACT reading section. For a measure of passage-level comprehension, the 35-minute ACT reading test was selected (ACT Inc., 2014), featuring four passages and 40 questions, both literal and inferential. The passages were the same length and addressed the same four disciplinary content areas. Unlike the instructional texts, which were only composed of texts

rated “complex”, the pre- and post-tests used only intact ACTs that included uncomplicated, more challenging, and complex passages; this was done to ensure the validity of the scale score interpretations. This measure was selected for several reasons. First, it is an easily interpretable measure relevant to students, teachers, and educators. Second, it is linked to the study’s text complexity framework and includes four disciplinary content areas but, was designed, as college entrance examinations are, with relatable but unfamiliar content for typical eleventh graders in order to minimize the effect of students’ prior knowledge. Third, it offered students an additional purpose for reading even beyond learning about the topics: to understand the types of texts that appear on entrance examinations as well as those they might see in college.

Two different forms were selected from older versions of the ACT that students would be unlikely to have seen before—form 61C for the pre-test and 71G for the post-test. The versions administered were older versions of the ACT reading test, which included both complex and less complex passages. The ACT technical manual specifies that the median for the reliability of the scale score across the six national ACT administrations in 2011-12, was 0.88, with the range extending from 0.86 to 0.90 across the six forms (ACT Inc., 2014). Data are not available for the specific forms of the ACT reading used for this study. The assessments were administered by the research team with assistance from advisory teachers during regular advisory time in large groups including both the intervention and comparison groups. Scoring of the multiple-choice questions was done by the author using automated optical scanning software. As each version of the ACT had accompanying scale score guidelines so that scores on the different versions could be compared, scale scores from 1-36 were used for interpretation.

TOSREC. For a complementary measure that tapped a different dimension of reading by measuring sentence-level reading efficiency, the 10th-12th grade Test of Silent Reading Efficiency and Comprehension (TOSREC) was used (Wagner, Torgesen, Rashotte, & Pearson, 2010). This measure helped investigate whether scaffolding would support reading efficiency as well as comprehension. The TOSREC is a three-minute test where students read a series of sentences of increasing length, but not increasing complexity and verify whether each sentence is true or false. For example, a sentence early in the assessment was "A companion might go on a trip with you". Then, sentences in the middle of the assessment were longer, such as "You should terminate a transaction with an individual who seems to be dishonest", and the final sentences are the generally the longest, such as, "Daydreaming, turning in homework late, and not studying for exams can jeopardize chances of getting good grades". While the designers of the test have not released information about the relative complexity of these sentences, the sentences appear to grow longer and include increasingly-rare vocabulary. However, the sentence-level nature of the test means that it cannot incorporate multi-sentence, paragraph-level, or discourse-level dimensions of complexity, such as the relationships, structure, or purpose elements of the RSVP framework. In addition, the three-minute nature of this test was far shorter than the ACT. Thus, when compared to the dimensions of reading tapped by the ACT, the TOSREC focuses on reading efficiency with texts of limited complexity.

Because the ACT pretest took 35 minutes of the 37-minute advisory period, both pretests could not be conducted on the same day. Thus, for intervention students, the TOSREC pretest was administered by tutors right before the first instructional lesson in their tutoring small groups while the comparison students completed it in a medium-sized group (8-

15 students) in their advisory classroom administered by a research team member. For the post-test, a similar procedure was followed at the end of the 8th instructional session. Per the test manufacturers, test-retest reliability exceeds 0.85 for all forms (Torgesen et al., 2010). Index scores were used for interpretation; the normed average score is 100 with an SD of 15.

Scaffolding. Researchers have noted the challenge of measuring scaffolding (Van de Pol et al., 2010; Rodgers, D'Agostino, Harmey, Kelly, & Brownfield, 2016). This may be because many studies do not adequately separate planned from interactional scaffolding, which this study does by controlling the texts and lesson plans across groups in a measurement strategy similar to prior research (Reynolds & Goodwin 2016a, 2016b). To measure scaffolding, tutors were given the list of potential scaffolds in Appendix C, and at the end of each lesson, they marked the checklist if they used that scaffold in that session. Then, a sum score was created that totaled the number of scaffolds each student received across the eight sessions.

To ensure the reliability of the tutors' self-reports, 23% of the sessions were observed by another member of the research team. Based on prior research and to account for the challenge of reliably identifying both high- and low-frequency scaffolds, a reliability threshold of 80% agreement or a Cohen's $\kappa>0.5$ had to be achieved (Landis & Koch, 1977; McHugh, 2012). Consequently, the extending talk scaffolding group were dropped because it was not reliable ($\kappa=0.44$, 72% agreement), and one tutor's scaffolding data had to be dropped because it was not reliable even after the training and supervision ($\kappa=0.34$, 67% agreement). Reliabilities for the remaining scaffold classes ranged between 79-91%, with $0.46<\kappa<0.64$. The overall reliability for the entire measure was 84% agreement, $\kappa=0.55$, suggesting the trustworthiness of the tutors' self-report data.

Demographics. Because of the necessity to protect student privacy in the no-consent design, district personnel provided demographic data linked to the de-identified student IDs, which were used to link the pre- and post-assessments and the demographics. The demographic data included categorical data about student race, dichotomous data about eligibility for free or reduced price lunch, dichotomous data for special education services eligibility with accompanying descriptions of disability, lists of languages spoken at home, and categorical data about eligibility for ELL services (four categories: receiving ELL services, exited ELL services, refused ELL services, or never classified as ELL). As the percentage of Hispanic and Asian-American students was very low, the analyses combined the categorical race data with the Black students into one dichotomous category here labeled "student of color".

Data Analysis

Because of the nested nature of the data where students were within advisory classes and tutoring groups, multilevel regression was used with Stata 14 (StataCorp, 2015). First, intraclass correlations (ICCs) were explored at three potential levels to address potential nesting effects: advisory teacher, intervention tutor, or tutoring group. The 0.1 guideline suggested by Snijders and Bosker (2012) was used. The ICCs for the advisory teacher level were 0.20 for the ACT outcome and 0.22 for the TOSREC, which suggested including it as a level. The ICCs for intervention groups approached the inclusion criterion (0.08 for the ACT, 0.09 for the TOSREC), but since assignment was largely done within advisory, this was likely an effect of the advisory clustering. For the tutor level, ICCs were negligible (0.025 for the ACT and 0.005 for the TOSREC). Thus, for both outcomes, two-level models of students nested within advisory teachers were used.

For RQ1, the experimental comparison between the treatment and comparison group, models investigated students in both intent-to-treat group (ITT) consisting of the 82 students who were assigned to treatment and treatment-on-treated group (TOT) consisting of the 48 students who fully participated in the intervention, defined as attending 75% or more of the instructional sessions[13]. The multilevel model for the ACT measure included 145 final observations representing the initial sample of 153 students without the seven attrited students and one student whose posttest went missing. Due to mis-administered assessments, 17 students' TOSREC pretests had to be discarded, and thus the sample sizes are lower for that outcome.

Models also included students' demographic data as reported by the school district, including their race, their free and reduced-price lunch status, and their special education status. As the sample included only one ELL, that variable was not included in the final model. While instruction was administered in four different cycles with differing intensities—some students received daily instruction and others received twice- or thrice-weekly instruction—no differences were detected according to these intensities (p=0.22), and it was dropped from the final model. The final multilevel regression for RQ1 was:

$$POST_{ijk} = \gamma_{000k} + \gamma_{10jk}PRE_{ijk} + DEMOGRAPHICS_{ijk} + INTERVENTION_{ijk} + u_{0k} + u_{0jk} + e_{ijk}$$

where DEMOGRAPHICS was a vector of dichotomous dummy variable indicators for students' race, lunch status, or if the student had a learning disability. The INTERVENTION coefficient of interest is a dummy variable indicating either the offer of the intervention (ITT models) or participation in the intervention (TOT models).

[13] As state data reports suggested that average attendance for the school in the preceding year was 92%. Because the intervention was eight sessions long, students missing one session would have been near the typical attendance rate, so one other absence for a total of two was considered appropriate for the attendance threshold.

Additionally, interaction effects were investigated to see whether intervention effectiveness interacted with students' demographic characteristics or pretest scores. No significant interactions were found for pretest scores (p=0.84), student of color status (p=0.21), eligibility for free or reduced-price lunch (p=0.70), or classification with a learning disability (p=0.17) and thus no interactions were included in the final models.

Estimating the overall explanatory power of the intervention or the scaffolds on the outcome measures required some way to estimate the amount of student-level outcome variance explained, which can be tricky in multilevel models. LaHuis, Hartman, Hakoyama and Clark's (2014) Monte Carlo study of different measures of explained variance found that calculating R^2 values in an OLS manner results in unbiased estimates and permits intuitive interpretation of the level-1 outcome. Because student-level outcome variance—and not cross- or multi-level variance—is of interest in this study, that OLS method was used to offer an intuitive interpretation.

As an additional safeguard (beyond the randomization) against the possibility that students in the treatment group might have differentially received external ACT tutoring or test prep that might have biased the results, the students were surveyed at pretest and posttest to see if they had taken extra ACT reading practice tests or received ACT tutoring either before or during the intervention. Chi-square tests detected no significant differences between the intervention groups and the comparison groups in experience on the ACT reading test (p=0.64), or receiving ACT tutoring (p=0.39). Consequently, these variables were not included in the final model.

For RQ2, the regression analyses associating the classes of scaffolds to the comprehension outcomes, the ITT models investigated only the students who received at least

one session of intervention treatment. This was chosen because, unlike the intervention evaluation in RQ1, the offer of scaffolding (i.e. including students who were offered intervention instruction but did not attend) is incompatible with the theory of contingency that required the scaffolding to be tailored to each student. Thus, the ITT models included 73 students: the 82 students assigned to treatment minus four attrited students, four students who were discarded for this analysis because of their tutor's unreliable data, and one missing posttest. The TOT models included only 48 students, representing only the students from the TOT group who attended 75% of the sessions.

The final multilevel regression for RQ2 was identical to RQ1, except with replacing the intervention coefficient with a vector of the 10 groups of scaffolds:

$$POST_{ijk} = \gamma_{000k} + \gamma_{10jk}PRE_i + DEMOGRAPHICS_{ijk} + SCAFFOLDS_{ijk} + u_{0k} + u_{0jk} + e_{ijk}$$

Effect sizes were calculated for interpreting the importance of the findings. For RQ1, because the assignment variable was dichotomous, Cohen's *d* was calculated. Though the treatment and control groups' pretest scores were not significantly different, the intervention group had slightly higher pretest scores. Therefore, effect sizes were adjusted to account for the small pretest differences. For RQ2, because the scaffold variables were continuous (and not dichotomous, as required for Cohen's *d* and as used in RQ1), Cohen's f^2 was used to estimate an local effect size for the coefficients of interest per the recommendation and procedures in Selya, Rose, Dierker, Hedeker, & Mermelstein (2012) and performed in Stata 14 (StataCorp, 2015) as such:

$$f^2 = \frac{(R^2_{full\ model} - R^2_{model\ without\ variables\ of\ interest})}{(1 - R^2_{full\ model})}$$

Because it is rarely used in education research and comparison guidelines are not widely

available (as they are for Cohen's *d*), interpretation guidelines for the f^2 statistic were drawn from Cohen (1988), who specified that f^2 values of 0.02 should be considered small, 0.15 considered medium, and 0.35 considered large. Given the comparatively lower effect sizes for yearly high school comprehension growth when compared to general effect size guidelines (Cohen, 1988; Hill et al., 2009), significant effect sizes lower than these general prescriptions would be notable findings for this study.

CHAPTER 4

RESULTS

Preliminaries

Equivalence and fidelity. Equivalence checks found that the randomization was successful, as no significant differences were found between the treatment and comparison groups on a t-test for the pretest measure (p=0.64), or on chi-squared tests of the demographic variables of race (p=0.98), eligibility for free or reduced-price lunch (p=0.88), or classification with a learning disability (p=0.27) . These equivalence checks bolster the inference that the randomization successfully distributed unobserved variables equally between the intervention and control groups.

Checks for fidelity of intervention implementation examined how frequently the tutors self-reported completing the components of the lesson plans. That tutors self-reported completing 92% of lesson plan components (i.e. checking the boxes on each part of the lesson plan, see sample lesson plan in Appendix A), and observers (present at 23% of session) reported tutors completing 95% of lesson plan components. Thus, the planned scaffolding curriculum was implemented in close accordance with the lesson plans. Overall, both the planned scaffolding components of the intervention were delivered as intended.

Descriptive Statistics.

Comprehension. Descriptive statistics for the comprehension measures across all students, the ITT group, and the TOT group are reported in Table 9. The sample pretest average ACT score of 17.7 is below both the state average of 20.3 and the national average of 21.3. As a validity check, pretest data were compared to official ACT reading scores for 27 students for whom data were available and predated the intervention, finding that the study's

pretest scores were slightly lower (on average by 0.7 pts) but not significantly different (p=0.30). The sample pretest average TOSREC score of 105 is slightly above the national average of 100, with the sample SD of 22.0 exceeding the national SD of 15.

Table 9

Descriptive statistics for comprehension measures

Group	Measure	N	Mean	SD	Min	Max
All students (N=153)	Pre-ACT	153	17.7	5.5	9	36
	Post-ACT	140	18.0	6.1	9	36
	Pre-TOSREC	121	104.7	22.0	55	156
	Post-TOSREC	144	105.2	19.8	50	145
Students offered treatment (ITT; n=82)	Pre-ACT	82	17.9	5.5	11	36
	Post-ACT	75	18.5	6.2	9	36
	Pre-TOSREC	70	105.2	22.1	58	156
	Post-TOSREC	77	105.2	21.4	50	145
Students not offered treatment (n=71)	Pre-ACT	70	17.5	5.6	9	33
	Post-ACT	68	17.2	6.0	9	30
	Pre-TOSREC	51	104.0	22.2	55	151
	Post-TOSREC	68	104.8	18.1	72	145
Students who participated (TOT; n=52)	Pre-ACT	52	17.8	5.0	11	30
	Post-ACT	48	18.6	6.1	10	33
	Pre-TOSREC	63	106.7	22.3	58	156
	Post-TOSREC	67	106.5	22.3	50	145
Students who did not participate (n=101)	Pre-ACT	100	17.7	5.8	9	36
	Post-ACT	96	17.6	6.2	9	36
	Pre-TOSREC	74	102.8	22.4	55	156
	Post-TOSREC	94	103.9	19.0	50	145

Scaffolding. Table 10 presents descriptive information about the scaffolding measure. The "# of scaffolds column" corresponds to the number of different scaffolds in each scaffold group (corresponding to the list in Appendix C), and the "mean" column corresponds to the sum score for all the scaffolds that student received across the eight sessions of the intervention.

Notable is the variability in all classes of scaffoldings, offering opportunities for differential predictive power. The background knowledge, academic register, and motivation

scaffolds were used most often, averaging almost ten or more scaffolds for each student over the course of the intervention. The structure, rereading, and mediators were used least often.

Table 10

Descriptive statistics for scaffolding measure

Scaffold group	# of scaffolds	Mean [over 8 sessions]	SD	Min	Max
Background knowledge	5	9.95	5.97	0	27
Morphology	4	5.15	3.28	0	14
Academic register	5	11.46	6.27	0	28
Syntax	3	3.72	3.08	0	15
Connections	3	4.38	2.53	0	12
Tracking participants	3	6.17	3.54	0	14
Structure	4	2.63	1.61	0	8
Motivation	7	13.26	7.72	1	34
Mediators	7	4.42	2.64	0	11
Reread	2	2.56	1.89	0	7

RQ1: Intervention Effectiveness

Passage-based comprehension (ACT outcome). Table 11 presents the results of the analysis of the intervention's effectiveness on passage-based comprehension (i.e., the ACT outcome). Column 1 presents the model with just the pretest, column 2 adds the demographic predictors, column 3 adds the intervention status for the ITT group, and column 4 switches the intervention coefficient to the TOT group. The overall variance explained (i.e. R^2 by just the pretest) was 51%, increasing to 54% after adding demographic controls, and then increasing to 56% where the intervention coefficients were added. This suggests that the intervention variables (in the ITT and TOT analysis separately) explained about 2% of the variance in the outcome across the 153 students.

Table 11

Intervention effectives for passage-based comprehension

	(1)	(2)	(3) ITT	(4): TOT
Constant	3.48**	7.14***	6.54***	6.78***
Pretest	0.80***	0.69***	0.68***	0.68***
Student of color		-1.51*	-1.49*	-1.72*
Lunch status		-2.01**	-2.04**	-1.98**
Learning disability		-1.59	-1.73	-1.60
Intervention status			1.20*	1.41*
Residual σ^2 teacher level	0.43	0.30	0.30	0.32
Residual σ^2 student level	1.34***	1.30***	1.29***	1.28***
Observations	145	145	145	145
-2 log likelihood	-412.07	-404.11	-402.43	-402.39
DF model	1	4	5	5
R^2	0.51	0.54	0.56	0.56

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The average effect of the offer of the intervention over the 82 students significantly predicted a 1.20-point increase in ACT posttest scores (p=0.048, 95% CI .01 2.39) when controlling for pretest scores and student demographic characteristics. When the treatment group was limited to the 52 participating students and the same variables were controlled for, participation predicted a 1.41-point increase in ACT posttest scores (p=0.028, 95% CI .15 2.67). Effect sizes for each group were Cohen's d=0.12 for the ITT group and d=0.17 for the TOT group, meaning the intervention effect was stronger for those students who attended at least six of the sessions.

To see if the eight students (5% of the total sample) who self-selected into the intervention biased the results, a robustness check was conducted by excluding those students from the analyses and re-running the models. While the ITT model coefficient decreased slightly in magnitude and significance (β=1.05, p=0.09), the TOT model coefficient increased in both magnitude and significance (β=1.49, p=0.025). Given these results, it appears that the

effect of the self-selection was minimal.

Sentence-level reading efficiency (TOSREC outcome). Table 12 presents the results for the TOSREC outcome in a similar manner. For this outcome, the offer of intervention (ITT) predicted a tiny increase of 0.91 on the post-TOSREC when controlling for demographic variables, but that increase was not significant (p=0.68). Participation in the intervention (TOT) predicted a slightly larger but still small post-test increase of 1.70 with similar control variables, but that was also not significant (p=0.31). Effect sizes for each group were d=0.04 for the ITT group and d=0.08 for the TOT group, meaning the effect of the intervention was stronger for those who participated, but also that this insignificant effect was about a third the size of the significant effect for the ACT measure.

Table 12

Intervention effectiveness for sentence-level reading efficiency

	(1)	(2)	(3): ITT	(4): TOT
Constant	21.53***	27.50***	27.08***	27.69***
Pretest	0.81***	0.77***	0.77***	0.76***
Race		-1.55	-1.63	-1.94
Lunch		-3.78*	-3.74	-3.73
LD		-7.61	-7.54	-7.48
Intervention			0.91	1.70
Residual σ^2 teacher level	0.95*	0.75	0.75	0.77
Residual σ^2 student level	2.15***	2.13***	2.13***	2.12***
Observations	117	117	117	117
-2 log likelihood	-423.14	-418.81	-418.65	-418.37
DF model	1	4	5	5
R^2	0.76	0.77	0.77	0.77

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

RQ2: Scaffolding Effectiveness

Passage-based comprehension (ACT outcome). Table 13 reports the results for the effectiveness of the scaffolding groups on the ACT outcome in a similar manner to RQ1.

Notably, the R^2 level was 0.50 in model 2 meaning pre-test scores and demographics explained about half the variance in outcome scores, but the R^2 increased to 0.61 for the ITT group (model 3) and 0.63 for the TOT group (model 4). It appears that the scaffolds are explaining about 12% of the variance in the ACT outcomes—in fact, that R^2 change was larger than the 0.07 increase when the demographics were added to model 1. In other words, the scaffolds explained more variance than the demographics. This suggests that the scaffolds themselves were contributing significant variation to outcomes—which could not be ascertained from the analysis in RQ1. These results suggests that it is not just the intervention that influenced outcomes, but the scaffolding within the intervention.

In the ITT model, one scaffold groups significantly predicted lower ACT posttest scores when controlling for the effects of all the other scaffold groups: rereading (β=-0.87, p=0.01, f^2=0.02). Conversely, two scaffold groups significantly predicted higher ACT posttest scores when controlling for all the others: syntax (β=0.54, p=0.04, f^2=0.05), and structure (β=1.18, p<0.01, f^2=0.18). Per Cohen's (1988) interpretive guidelines, there is a small negative effect for rereading scaffolds, and a small positive effect for syntax scaffolds and a medium-sized positive effect for structure scaffolds. Thus, the positive scaffolds appeared to have a somewhat larger effect than the negative scaffolds. Note that the effects of each scaffold group are independent—the modeling approach of adding all the scaffold groups as a block effectively controls for all of the others at once, indicating that these scaffold group's effects contribute uniquely to the outcome.

The TOT analysis produced similar results for all coefficients except syntax, academic register, and structure. The β-coefficient for syntax, controlling for the same variables as the ITT analysis, decreased markedly, as its significance decreased drastically to p=0.58, and its

effect size was negligible (f^2<0.01). The estimate for the academic register scaffolds, which had been marginally significant (p=0.09) in the ITT model, became significant in the TOT model (p=0.046), and increased slightly in magnitude from -0.25 to -0.35, with an effect size of f^2=0.02. For the structure scaffolds, the β-coefficient increased drastically to 1.80 in the TOT model and remained significant, and its effect size increased to f^2=0.35, a large effect size per Cohen (1988). This suggests that an increased dosage of structure scaffolds was associated with increased effectiveness.

Table 13

Scaffold effectiveness for passage-based reading efficiency

	(1)	(2)	(3) ITT	(4) TOT
Constant	4.05*	8.81***	10.22***	13.85*
Pretest	0.79***	0.65***	0.67***	0.56***
Student of color		-1.93^{+}	-2.12^{+}	-1.69
Lunch status		-2.33^{+}	-2.15^{+}	-3.03^{+}
Learning disability		-3.84^{+}	-3.31	1.43
Background			-0.31^{+}	-0.29
Reread			**-0.87****	**-0.88***
Mediators			0.03	0.08
Motivation			0.22	0.24
Morphology			0.10	0.08
Academic register			**-0.25^{+}**	**-0.35***
Connections			-0.63^{+}	-0.88^{+}
Syntax			**0.54***	**0.20**
Tracking participants			0.35^{+}	0.35^{+}
Structure			**1.19*****	**1.80*****
Residual σ^2 teacher level	0.22	-0.30	-17.20	-14.84*
Residual σ^2 student level	1.48***	1.41***	1.28***	1.26***
Observations	73	73	73	48
-2 log likelihood	-214.35	-207.91	-197.10	-128.41
DF model	1	4	14	14
R^2	0.43	0.50	0.61	0.63

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Sentence-level reading efficiency (TOSREC outcome). The results for the sentence-level reading efficiency outcome are in Table 14. The high R^2 of 0.78 in the pretest-only

model indicates that the pretest alone explained 78% of the variance in the post-test. In comparison, the pretest-only model of passage-based comprehension (Table 13) explained just 43% of the outcome variance, suggesting that sentence-level reading efficiency is much less malleable to instruction than passage-level reading comprehension. Adding the scaffolds to model 2 had very little overall effect on the explained variance, accounting for only 1% of the variance in the outcome, and the effect grew no stronger with the increased dosages in the ITT model. Therefore, interpretation of the significance of the TOSREC coefficients is not meaningful.

Table 14

Scaffold effectiveness for sentence-level reading efficiency

	(1)	(2)	(3): ITT	(4): TOT
Constant	20.98***	32.49***	27.28**	12.31
Pretest	0.82***	0.78***	0.78***	0.90***
Student of color		-4.57	-1.98	-1.87
Lunch		0.34	-2.58	-2.20
Learning disability		-12.61	-13.95	-9.64
Background knowledge			0.86+	1.43*
Reread			0.93	0.84
Mediators			-0.70	-0.26
Motivation			-0.82+	-1.17+
Morphology			0.02	0.32
Academic register			0.39	0.64
Syntax			-1.22	-1.64
Connections			1.88+	2.02
Tracking participants			-0.31	-0.60
Structure			-0.26	-0.97
Residual σ^2 teacher level	1.21*	1.10*	-17.68	-21.49**
Residual σ^2 student level	2.23***	2.19***	2.17***	2.17***
Observations	63	63	63	44
-2 log likelihood	-233.52	-230.96	-226.24	-157.98
DF model	1	4	14	14
R^2	0.78	0.79	0.80	0.80

CHAPTER 5

DISCUSSION

The goal of this study was study to investigate a scaffolding-based intervention in which tutors read complex texts with high school students and scaffold their comprehension. The analysis of RQ1, using a randomized controlled trial design, found that the intervention caused significant growth on a standardized measure of passage-level reading comprehension. The analysis of RQ2, using a multilevel regression approach, found that the scaffolding in the intervention also explained meaningful variance in standardized passage-level reading comprehension, and that specific types of scaffolds had significant effects. The discussion below contextualizes the findings in the research literature, assesses the strengths and limitations of the research design, and offers implications for future research.

Understanding the Causal Effects of the Intervention

Positive effects on passage-level comprehension. The main finding of this study is that an intervention focused on scaffolding, paraphrasing, and complex texts caused significant growth in 11th graders' reading comprehension. As few interventions with rigorous randomized designs have been conducted in high school—and none on late high school students in 11th or 12th grade—this study offers important evidence to inform future research and practice at the high school level. This evidence is particularly relevant because of the strength of the research design: the randomization of students to condition, the no-parental-consent sampling procedure including students with a wide range of reading abilities, the regular fidelity observations, robustness checks, and validity checks are all layers of the research design that strengthen the causal claims and overall significance of the study. Therefore, it is reasonable to claim that small group instruction that scaffolds complex texts

within a paraphrasing intervention can support high school students' reading abilities, even transferring to performance on standardized reading assessments.

Understanding the meaning of the effect size also requires understanding the treatment difference between the treatment group and the counterfactual. Students in the counterfactual were doing computer-based instruction, but it is not clear how many passages of complex text they may have read during the time of the intervention—perhaps, working alone, they read more than one passage per session (the amount read by the treatment group), or perhaps they focused on other ACT sections and read less than the treatment group. Such data were not available, and thus the experimental design was unable to control for reading volume, which may have contributed to the differential effect size. In addition, while the tutors were generally able to keep the students engaged in the lesson and participating in the instruction (as verified by the tutor-observers), no comparable fidelity data are available for the computerized instruction. Thus, the experimental design was unable to control for instructional engagement, which may have differed between the two conditions. In this case, though, the treatment difference is deliberate—the treatment approach grounded in interactional scaffolding was designed to provide an accessible entry point for students to engage with the text (i.e. paraphrasing), and both the planned scaffolding (i.e. the norms of all students participating) and the interactional scaffolding were designed to maintain high student interaction with text. Though the computer program presented students with individualized curricula, the program appears to rely on students' individual motivations to stay engaged in their learning. In addition, it is important to note that the high school purchased the ePrep program for each of its eleventh-grade students as a way to improve students' ACT scores, and though improvement on other sections (i.e. math, science, or

English) may have occurred due to software use, the results of this study show that the treatment group improved significantly more than the comparison group in their reading scores.

Contextualizing the statistically significant and practically meaningful effect sizes (d=0.12 for the ITT group and d=0.17 for the TOT group) requires comparing them to other similar interventions. While general effect size guidelines such as those originally prescribed by Cohen (1988) for psychological research studies might consider these effect sizes small, those general guidelines do not consider educational guidelines for interpretation. Extensive existing research in reading comprehension can add a more nuanced understanding of the meaning of these effect sizes. Perhaps most helpful in understanding the effect sizes for 11th graders' standardized reading comprehension comes from a meta-analysis by Hill, Bloom, Black, and Lipsey (2008). Their study estimated the average yearly effect size growth (in the absence of special interventions) across each K-12 grade and across seven different standardized measures of reading comprehension. They found that the yearly growth in the early grades was large (d=1.52 between grades K-1, d=0.97 between grades 1-2), but declined drastically with age, culminating in an average effect size growth for 11th-12th graders of 0.06, with a margin of error as ±0.11. This shows that late high school students' natural yearly growth on standardized comprehension measures is quite small. Therefore, it appears that this study's effect sizes represent more than a full year's growth for these students—powerful evidence of improvement for a group of students whose reading is largely unexamined by rigorous experimental interventions. This is particularly noteworthy because the intervention took only eight 37 minute experimental sessions, much less than the dosage implied in the yearly growth estimates provided by Hill and colleagues (2008). These findings suggest that

intensive small-group scaffolding is a strong way to improve student comprehension.

Contextualizing this study's significant effect sizes of 0.12 (ITT) and 0.17 (TOT) to other comparable interventions is difficult because of the paucity of such studies published in peer-reviewed journals. Paul and Clarke's (2016) review of RCTs conducted for secondary students revealed only one RCT conducted since 1992 on typically developing high school students; that study reported that a yearlong intervention for students in the bottom third of reading achievement in their schools receiving a Learning Strategies Curriculum—a comprehension-strategies intervention—had no significant effects on 9th grade students' reading comprehension or strategy use (Cantrell, Almasi, Carter, Rintamaa, & Madden 2010). One other intervention—Lee's (1995) six-week quasi-experiment with 52 treatment students and 25 control students—provided enough information to calculate a substantial effect size of d=0.75 on a 10-question short-answer researcher-designed measure of comprehension (as opposed to the standardized measures used in this study). As Lee (1995) also explicitly designed interactional scaffolding and supporting students' existing linguistic resources as part of the intervention, this study corroborates the justification for including them as part of reading instruction for high school students.

Because of the limited evidence for high school reading interventions in peer-reviewed research, contextualizing the findings from this intervention requires investigating technical reports as well. A synthesis of evaluation studies funded by the federal Striving Readers grant incorporated findings from evaluations meeting What Works Clearinghouse evidence screens that were published in technical reports but not in peer-reviewed journals (Boulay, Goodson, Frye, Blocklin, & Price, 2015). This report revealed more evidence including several additional RCTs targeting high school students' comprehension, but with the same patterns

seen in peer-reviewed li

terature: working with whole classes of 15-25 students, working almost exclusively with 9th grade students (and none with 11th or 12th grade students) and largely targeting students already struggling in reading. In addition, these studies evaluated yearlong curricular programs with students receiving substantially larger dosage of intervention instruction than the students in this study did. These programs include READ 180 (Sprague, Zaller, Kite, & Hussar, 2012; Swanlund et al., 2012), Xtreme Reading (Faddis et al., 2011), Voyager Passport Reading Journeys (Dimitrov, Frye, Juyrich, Saychko, & Lammert,, 2014), and the Fusion Reading Program (Schiller et al., 2012), and the Kentucky Cognitive Literacy Model (Cantrell, Almasi, Carter, Rintamaa, & Madden, 2010), and Strategies for Literacy Independence across the Curriculum (Hofstetter et al, 2011). Many of these studies reported null effects; the only studies reporting significant positive results were READ 180, with effect sizes ranging between 0.14 and 0.18 for studies conducted in high schools, and Xtreme reading with effect sizes between 0.09 and 0.21. In this context, the intervention involving paraphrasing and scaffolding small groups of students reading of complex text appears to be producing a similar effect size in substantially less time—about four hours of instruction versus the yearlong programs presented in Boulay and colleagues' synthesis (2015).

Beyond high school interventions, this study's findings are consistent with those found in the 11 experiments evaluating scaffolding described in Reynolds (2017): scaffolded reading instruction is superior to less-scaffolded reading instruction across developmental levels. This study offers only the second (Lee, 1995) experimental study of scaffolding at the high school level, so these findings confirm that scaffolding is effective for elementary, middle, and high school students. A meta-analysis of Reciprocal Teaching studies—similar to this study in that

they are all of small-group scaffolded reading with a focus on comprehension—found that the median effect size for 16 Reciprocal Teaching interventions was 0.32 (Rosenshine & Meister, 1994). None of the studies, however, evaluated interventions at the typical high school level[14], so the larger effect size of those interventions on mostly younger students fits with the trend identified by Hill and colleagues (2008) of lower effect sizes with increasing student grade. Overall, this study's findings fit with both Reynolds (2017) and Rosenshine and Meister's (1994) reviews: scaffolding is effective in supporting student comprehension. In addition, the study goes beyond prior studies to show that scaffolding is specifically effective in supporting high school students reading complex texts.

The study's lack of interaction effects according to students' pretest ability, race, socioeconomic status, or special education classification suggests that the intervention effect was equally consistent across these kinds of students—especially noteworthy because of the diversity of reading abilities of the students in the study. The pretest scores on the ACT ranged from 9 (equivalent to random guessing) to 36 (a perfect score). The consistent effectiveness is particularly important given the troubling gaps in reading achievement in US high schools for struggling readers, students of color, and students of lower socioeconomic strata (NCES, 2015). Reading research should be constantly striving for instruction that achieves equitable outcomes, and this study's evidence suggests that a scaffolding-based vision of reading instruction may support that goal.

In addition, as many of the null findings for existing interventions described by Boulay and colleagues (2015) and Paul and Clarke (2016) prioritized strategy-based curricula (i.e.

[14] Thirteen of the studies were conducted in elementary and middle school levels, while one was conducted at the "vocational" grade level, one with college students, and one with adults.

planned scaffolding), this study presents evidence for a different route to improving high school students' comprehension. Perhaps reading instruction that places interactional scaffolding at its center (and not planned scaffolding curricula) may be an ideal way to prioritize extending the existing linguistic and intellectual resources of each student instead of teaching students the strategies used by others.

One reason this study may have proved effective in improving the generally less-malleable domain of reading comprehension is because it drew on the fundamentals of scaffolding research (Reynolds, 2017). The study's focus on delivering the scaffolding contingently and responding to student knowledge (Van de Pol et al., 2010) likely supported students in developing more sophisticated paraphrasing of the complex texts. In fact, contingency research has shown that even less-than-perfect contingency in reading scaffolding may still be adequate for significant student growth, provided the scaffolders strive for increasing contingency (Rodgers et al., 2016). Therefore, it seems likely that the focus on contingency was a positive contributor to the intervention success. In addition, attending to the small groups' culture by explicitly building relationships with students and establishing conversation norms (as suggested by Christoph & Nystrand, 2001; Many, 2002; Maloch, 2004), even though it may have decreased the number of minutes spent focusing on text, likely provided the necessary relational groundwork for talking out the challenging texts.

In addition, the study's success might also be in part because students' own linguistic resources were explicitly valued—both in tutors' framing of the instruction and in the planned scaffolding structure of paraphrasing activity. Transcripts from the intervention show that students, encouraged to talk, put themselves in the positions of historical oceanographers plumbing the depths of the Atlantic and compared Jefferson's potential plagiarism to a high

schooler writing a paper for class. Like in Lee (1995), it seems that building an instructional model that explicitly valued and drew on students' linguistic resources, but anchored those resources in textual interaction, may have been a positive factor in the success of the intervention.

Finally, choosing the complex texts as the texts of the intervention likely encouraged tutors to have rigorous situation definitions (Wertsch, 1984), and discouraged them from reducing the challenge of the task in the interaction (McElhone, 2012; Poole, 2008; Mertzmann, 2007). This is the hope of text-complexity researchers—that such texts will raise the expectations for comprehension and allow students to learn about complex content. Overall, while the research design does not identify which of these dimension of the scaffolding contributed to the outcome, the theoretically-driven components of scaffolding research probably supported the students' overall comprehension growth.

Beyond scaffolding, this study adds to the existing research about interventions including paraphrasing as a focal activity. Fundamentally, paraphrasing strategies seek to improve comprehension by leveraging students' linguistic resources to put academic texts into their own words—and they prioritize students' productive speaking abilities as vectors of improvement (Frankel et al., 2016). In a way, the paraphrasing strategy insisted that the students not just make their comprehension visible to the scaffolder, but also read and construct meaning actively, avoiding a passive or receptive stance toward the text. As the paraphrasing was done sentence-by-sentence (or at times every two or three sentences), students and tutors were collectively monitoring the students' comprehension of the text quite closely, which may have supported the students' comprehension-monitoring skills.

To situate the findings of this study in the research literature on paraphrasing, single-

group studies (e.g. Hagaman et al, 2012; Hua et al., 2014) have shown the effectiveness of paraphrasing as a strategy, but the lack of comparison groups in those studies prevents meaningful understanding of effect sizes. One study (Katims & Harris, 1997) reported an effect size of 0.48 for an experimental comparison, though study used researcher-designed measures of comprehension, which meta-analyses have found to be associated with higher effect sizes than standardized measures of comprehension (Edmonds et al., 2009; Rosenshine & Meister, 1994; Scammacca et al., 2007). Still, this is emerging positive evidence for paraphrasing. Therefore, as the only experiments experimentally evaluating the paraphrasing strategy, this study and Katims and Harris (1997) offer a trend of evidence that paraphrasing may be a helpful approach to comprehension for middle and high school students.

This study also adds to research about scaffolding and text complexity. The literature review of interactional scaffolding and comprehension by Reynolds (2017) found that only four of the 57 studies surveyed explicitly conceptualized their texts as complex. Three of the studies were observational, showing how teachers' scaffolding supported students' comprehension, but without comparing it to a less-scaffolded instructional condition (Palincsar, 1986; Maloch, 2008, Moss, Lapp, & O'Shea, 2011). Reynolds and Goodwin's study (2016a, 2016b) of complex text scaffolding with fifth and sixth graders found that, unlike this study, motivational scaffolding but not any other kind of scaffolding predicted increased sentence-level reading efficiency (i.e. TOSREC gains). However, that study did not include a control group that did not receive complex text instruction, so the comparative benefits of the complex text instruction could not be ascertained. In addition, none of the previous eleven experimental studies of scaffolding used complex texts (Reynolds, 2017). Thus, this study's use of a randomized experimental design represents the first evidence that

suggests that scaffolding students reading complex texts is an instructional approach powerful enough to increase independent comprehension.

Null effects on sentence-level reading efficiency. When compared with the significant findings for passage-based comprehension, the positive but insignificant findings on the sentence-level reading efficiency outcome might seem puzzling. However, these two measures tap different dimensions of reading, and the intervention's focus on sensemaking, discussion, and complexity were not closely aligned with the TOSREC's focus on efficiency and relative lack of complexity. While students did practice reading aloud with complex texts, tutors often scaffolded their comprehension by asking them to refine their paraphrases over multiple turns of talk—not the sort of approach that would be beneficial on a test of efficiency. Similarly, some of the academic-language scaffolding would have supported students' making logical connections and tracking participants across multiple sentences—or even across paragraphs—of the complex passages, a skill that would not have been helpful on the TOSREC because its sentences were all independent and not linked by any content.

Understanding the Effects of Scaffolding

Scaffolding and passage-based comprehension. Examining the RQ2 results for the passage-based comprehension outcome (ACT) offers additional insights into understanding the effects of the intervention. The 12% increase in explained variance when the scaffolds were added to the model (models 3 and 4) suggests that the scaffolding measure—even though it was tutors' self-reports—tracked instructionally relevant discourse that had discernible effects on student reading. In other words, while the experimental comparison could not parse out whether scaffolding, paraphrasing, or text complexity were explaining student comprehension growth, the regression analyses of scaffolds suggests that scaffolding

is making a meaningful contribution. This supports the assertion that interactional scaffolding is not just the incidental adaptations that teachers make as they deliver a planned scaffolding curriculum, but is an instructionally relevant domain in its own right, and a worthy subject for future implementation research.

The finding that three scaffold groups were significant predictors of passage-based comprehension is also noteworthy—rereading and academic register scaffolds negatively so, and structure positively so. Based on the modeling procedure, each of these estimate for the effect of a scaffold group was effectively controlling for the others, and thus each scaffold group's relation to the outcome was independent, and above and beyond the contributions made by the scaffolds in general, which is visible in the amount of additional variance explained in models 3 and 4. Given the lack of evidence for late high school readers about what kinds of scaffolding are more effective, these results show that the types of interactional, contingent scaffolding mattered, and that certain scaffolds were particularly effective. These findings also show that planned scaffolding is only part of the picture of planning for high school reading—teachers should be listening closely to emerging student comprehension and working to develop their skills in contingently responding to students' knowledge prioritizing certain scaffolds that seem more supportive than others. These findings offer some guidelines for researchers and educators to refine knowledge of the kinds of scaffolding.

Syntax and structure. These academic-language scaffolds were positive predictors of passage-based comprehension, with small effect size for the ITT model. The oral nature of the intervention meant that it was easy for tutors and students to note long sentences—transcripts of the intervention note students saying "that was a lot" "that was a mouthful" after reading and before paraphrasing a long sentence. Syntax scaffolds may have helped students break

down lengthy sentences to get at the gist of their meanings. Yet syntactic complexity refers not just to sentence length but also to grammatical and stylistic features. This dimension of sentence complexity is represented in the RSVP category of Richness, and is increasingly recognized as a key index of overall text complexity (Starr, Frantz, & Bailey, 2015). For example, Uccelli and colleagues (2015) noted that "denser syntactic structures, such as center-embedded-clauses, are widely used in academic texts" (p. 342). Similarly, Snow (2010) pointed out that academic texts are more likely to use grammatical embeddings to convey complex relationships when compared to everyday language, which is more likely to present simpler sequential information. Scott (2009), when reviewing research on syntax and comprehension, noted that research has documented the links between syntactic processing and overall comprehension, suggesting that techniques for supporting students in tackling complex syntax could be positive routes to supporting overall comprehension. Therefore, it may be that the syntax scaffolds—such as #26, which prompted students to consider how different parts of the sentence work together—assisted students in unpacking the complex texts' grammatical complexity, and consequently supporting their overall comprehension.

Curiously, the syntax scaffolds were statistically significant in the ITT model, but not in the TOT model. This suggests that they may be effective in lower dosages---the ITT students received, on average, 2.8 syntax scaffolds, while the TOT students received 5.8. Perhaps there is a threshold effect where a few syntax scaffolds can help students unpack complex sentences, but repeatedly scaffolding in this way may not be any more effective. Still, as syntactic complexity has been established as a key dimension of text complexity (Frantz, Starr, & Bailey, 2015) and that individual differences in unpacking complex sentences are linked to reading comprehension (Poulsen & Gravgaard, 2016; Uccelli et el.,

2015), more research would be needed to establish a firm relationship between the degree of syntactic scaffolding and comprehension.

The most powerful scaffolds of all were the structure scaffolds, which link to the CALS skill of organizing argumentative texts. Already the largest effect size of any scaffold group when estimated with the ITT model, the effect size nearly doubled when applied to the TOT group, so an increased dosage appears to be even more effective; students in the TOT group received on average 3.3 structure scaffolds, while the ITT group received only 2.6. The positive finding is consistent with studies which have established that teaching text structures is positively linked to comprehension growth (Cain & Oakhill, 2006; Duke & Pearson, 2008; Pearson & Gallagher, 1983). Interestingly, this study differs from most of that research because instead of explicitly teaching text structure analysis, the scaffolds were only applied as tutors thought they would be effective. It may be that attending to the author's overall purpose—and asking about how specific details fit within that structural purpose as scaffold #36 did—helps readers pay attention to both content and craft, contributing to overall comprehension. These scaffolds, more so than the word-centered morphology scaffolds or the sentence-focused syntax scaffolds, also function integratively. That is, they ask students to combine insights about small details with the bigger picture of the structure, potentially increasing the overall coherence of their interpretation.

In addition, scaffolding students' understanding of text structures may also implicitly link to understanding the author's purposes. Going beyond simply establishing a topic (e.g. undersea landforms), or a main idea (e.g. the author's discussion of the uniqueness of undersea landform), understanding the author's purpose and how he or she structures the argument to achieve that piece may involve a deeper and more meaningful comprehension

(e.g. The author attempts to persuade the reader that unknown landforms on the bottom of the sea are actually more physically impressive and geologically noteworthy than their counterparts on land).

When examining the descriptive statistics, the structure scaffolds group was the second-least used group of scaffolds—on average, students received only 2.6 structure scaffolds over the course of the eight lessons[15]. This may be because, while some scaffold types were useful at any point in the text, structure scaffolds were less likely to be used at the beginning of the text, before the author's structural choices become apparent, and so were more useful toward the end of a text, as the author's structure fully materialized. In addition, the tutors were instructed only to respond to contingent student knowledge and not pre-teach anything, so they may have been waiting till later in the text to use structure scaffolds. Another possible explanation for the low dosage of structure scaffolds is that the sentence-by-sentence nature of the paraphrasing, while facilitating comprehension monitoring at that level, kept the tutors' focus on supporting students in comprehending smaller units of meaning and not larger structural components of the text. While the exact reason for the low dosage of structure scaffolds remains elusive, the large coefficients and effect sizes as well as the consistency with other empirical and theoretical reading research suggests that tutors and teachers should keep these scaffolds as part of their comprehension-scaffolding repertoire.

Rereading. The negative association of rereading scaffolds with passage-based comprehension show that not all scaffolding is helpful. The magnitude of the effect was consistent for both the ITT and TOT groups, suggesting that the effect of rereading is

[15] As there were four types of structure scaffolds (See Appendix C), the potential maximum student score would have been 32.

consistent no matter how many sessions the students attended. The effect size of 0.02 was relatively small by Cohen's (1988) guidelines and when compared to the other scaffolds. This scaffold class included two simple scaffolds: tutors re-reading portions of the text, or tutors asking students to re-read text. Three potential explanations for the negative associations come to mind. First, in situations where tutors asked a student to re-read the sentence because he or she pronounced it incorrectly, tutors may have been inadvertently reinforcing a performance-based view of reading which requires one correct pronunciation—which may be at odds with constructing comprehension via sensemaking stance in which students are encouraged to make optimal use of their existing comprehension strategies. Research on younger students has shown that students who are seen as weak readers because of poor pronunciation can receive more basic literacy instruction rather than higher-level thinking questions, even when placed in heterogeneous groups with the goal of having high expectations for all students (Poole, 2008). In sum, these scaffolds may have reinforced a performance- or pronunciation-oriented view of reading that was not aligned with the instruction or the comprehension measures.

A second explanation could be that the rereading scaffolds may have been less contingent than the other kinds of scaffolds. For example, if a student's initial paraphrase did not coherently account for textual features, but the tutor's scaffolding choice was to simply reread the entire sentence rather than guiding the students to address the words or propositions elided by the student's paraphrase, rereading scaffolds could have been less contingent than academic-language or background knowledge scaffolds. It may be that less contingent responses may have been less effective in supporting students' emergent sensemaking. A third potential explanation suggests that, when rereading parts of the passage to students, tutors

were not transferring the responsibility to the students—a key dimension of scaffolding (Van de Pol et al., 2010). In short, tutors who used these scaffolds frequently may have been doing too much of the reading work for their students. While occasionally rereading scaffolds may have been necessary to come to intersubjective agreement about word identification, it appears that overall they may not be effective scaffolds for supporting students' comprehension.

Academic register. The significant negative associations of academic register scaffolds with passage-based comprehension for both the TOT and ITT groups was an unexpected finding, as the academic-language dimensions of these scaffolds was thought to align with the academic discourse of the complex texts and would thus help students unravel how the authors talked about the complex ideas in those texts. This effect was relatively small (f^2=0.02), similar in size to the negative effect of rereading scaffolds.

In retrospect, examining an instance of academic register scaffolding might help understand how they were enacted. In this sequence, the tutor is working with a group of three students (two of whom speak in this transcript) scaffolding the complex sentence "*Today, highly sophisticated sound waves bring the hazy images of those early soundings into sharp focus, revealing that one of the largest and most salient geographic features on the planet lies on the floor of the ocean*" (from the text in Appendix C, also the sentence discussed earlier, albeit with a different scaffolding group, in the Tutor Training Day 1 section). The student's paraphrase did not pick up the running thread through the text that the author compares underwater landforms to those on land. Ostensibly to scaffold understanding, the tutor selected the word *salient* to scaffold (Scaffold #18 in Appendix D) in this sequence:

Tutor: If I were to ask you what *salient* means what do y'all think? Now knowing more about this mountain range ... what could we infer *salient* means?
S1: Oceanic ...
Tutor: Oceanic ... Okay, what do you think, [Student 2]?
S2: I don't know.
Tutor: We know it's the backbone of the sea. We know it's extensive. We know it's almost as large as all of the land continents.
S1: So we're trying to figure out what *salient* means.
Tutor: Yeah.
S1: Ancient ...
Tutor: Ancient, okay.
A1: Impressive.
Tutor: Impressive. Good. Okay I like these, I like these guesses. Okay, good any other thoughts about what salient could mean? Nope, okay. So, sometimes just know that when you're reading complex text you can read further and use the context to try and help you figure out the little details

In this case, it is clear that S1 and S2 did not already know what the word *salient* means, and S1's responses of "oceanic", "ancient", "impressive" make somewhat reasonable guesses based on the immediate context of the sentence, but they also seem somewhat random, as though S1 were simply trying to match what he thinks the tutor wants him to say. On the other hand, S2 simply chooses not to hazard any attempts. At this point, after the tutor has attempted scaffolding through pointing out a key word, asking students to infer its meaning, and trying to supply what she thought might be relevant context ("We know it's…. We know it's… We know it's…"), the tutor is still not satisfied ("Any other thoughts about what salient could mean?", and the sequence does not achieve either students recognizing either the definition of *salient* or recognizing the larger thematic connection of the word *salient* (i.e. the author's comparison of undersea and land topography). This sequence illustrates a potential explanation for the negative effects of academic register scaffolds: several of the academic register scaffolds focus on asking students for their vocabulary knowledge. Scaffold 18 asks

students to consider the meaning of a rare word, #20 asks students if they know the meaning of the word in an academic context, and #21 asks students to consider alternate word meanings. While these were intended as academic-register scaffolds, they could easily have been enacted as outcome-model scaffolds in which the tutors quizzed students for vocabulary knowledge (as it appeared in the sequence with S1 and S2) rather than challenging their sensemaking to become more coherent overall. Therefore, if these scaffolds were enacted in an outcome-model way (i.e. requesting students match tutor-accepted definitions for the vocabulary), it may have reinforced the idea that students had to have extensive background vocabulary to make sense of the text. If students approached the independent assessment with the belief that their vocabulary was inadequate, they may not have been positioned to succeed.

Scaffolding and sentence-based reading efficiency. As in the analysis of the overall effectiveness of the intervention, the effects of scaffolding on passage-based reading comprehension were much larger than the effects on sentence-based reading efficiency. Adding the scaffolds to the model only increased the overall variance explained in sentence-level reading efficiency by 1%, and the scaffolds' effect sizes were negligible. Therefore, this evidence suggests that academic-language scaffolding does not have a notable effect on sentence-level reading efficiency with high school students. This is notable because it contrasts with the positive relationship between motivational scaffolding and sentence-level reading efficiency reported in an intervention with small groups receiving scaffolding while reading complex texts (Reynolds and Goodwin, 2016b). Because the motivational scaffolds this study were conceptualized on and similar to those in that 2016 study, but that study was with 5th and 6th graders, it appears there may be developmental differences in the relationship between motivational scaffolding and comprehension. Perhaps younger students' engagement

and motivation can be scaffolded in a way that increases their engagement with the text and thus their comprehension skills, while that does not work with high school students. Put simply, it may be easier to get younger kids excited about the text and thus help them learn, while high schoolers' motivation to engage with complex texts may be more complicated and less affected by simple scaffolds.

Strengths & Limitations of Design and Implementation

Strengths. The findings of this study are buttressed by the strengths of its design. The random assignment offers a strong causal argument that the intervention was the sole cause of the increase in comprehension between the treatment and control groups. The lack of interaction effects suggests that the intervention was equally effective for students of different backgrounds and abilities—a useful finding for schools serving diverse populations of readers.

The implementation details also constitute strengths that offer possibilities for future work to build on. Keeping planned scaffolding consistent across groups allowed the intended focus of the study to be on the interactional scaffolding in which tutors responded to their students' emerging comprehension—a design similar to Reynolds and Goodwin (2016a, 2016b) and Rodgers and colleagues (2016) and recommended by Reynolds and Daniel (2017) as a model for future studies of interactional scaffolding. In addition, the layers of validity in the study—checking the validity of students' pretest scores, surveying to rule out possible extraneous causes for comprehension growth, observers' documenting the trustworthiness of tutors' self-reports, robustness checks against self-selection—strengthen the causal arguments for the intervention and the integrity of the correlational findings.

Finally, the setting and the tutor populations are also strengths of the study. The

student population—especially as sampled through the no-parental-consent procedure—is similar to other US high schools, and conducting it within the setting of a high school suggests that other studies might be able to do the same. In addition, the effects were achieved using tutors that did not have extraordinary expertise—while three of the nine tutors had extensive high school experience, four were relatively inexperienced in the field of education, and two had little education experience. Since the ICCs did not detect significant variability at the tutor level, it appears that the tutors were relatively equally effective. This suggests that this intervention could be delivered without an enormous investment in human capital.

Limitations and practical concerns. While the strengths of this study offer strong reasons to support the findings, several limitations and practical methodological concerns exist as well. First, the control group offered a meaningful comparison because students were working on ACT preparation through test-prep software, but since it was self-paced, it is not clear which ACT subject they were working on at any given moment. Future interventions could be conducted during English classes to provide a clearer comparison to business-as-usual reading instruction.

Another practical concern of the study was that it was conducted during the late spring, close to the times in which students took their battery of state end-of-course tests over a two-week period. Students did complain of testing fatigue, especially on the post-tests; while there is no evidence to suggest the testing fatigue affected the control group differently than the intervention group, it may have exerted a conservative bias on overall student growth.

Finally, measurement limitations also existed. Because of the difficulty of identifying scaffolding intensity within sessions, the measure is an imperfect estimate of the intensity of the scaffolding each student received. Future studies will have to examine how to measure the

within-session intensity of scaffolding in the context of implementation. Finally, measuring extending-talk scaffolds could not be done reliably, likely because of their general nature (e.g. scaffold #1 asking students to extend their paraphrases) made them difficult to identify precisely. While extending students' talk has been shown to support student comprehension (Boyd & Rubin, 2006; McElhone, 2012), those studies used video data to code their teachers' scaffolding moves, and thus more fine-grained coding systems could be established that were not possible with the live-observation measurement methodology. Future studies considering measures of scaffolding should investigate whether such scaffolds can be reliably identified in real time by tutors and observers, or whether retrospective video analysis would be required.

The small-group scale of the intervention could be considered a strength or a limitation. Some scaffolding-intensive programs show significant improvement with one-on-one literacy tutoring; Juel (1996) pointed out that that the individualized nature of tutoring—which, in scaffolding terms, is ideal for contingent response—is nearly always effective for students, but noted that its high cost makes it prohibitive to implement at large scale across schools and districts. In this view, then, the groups of 2-5 students in this study could represent amore scaled-up efficient approach than individually tutoring each student. Still, the 2-5 student groups are far smaller than typical high school English classes, and it would be difficult for a teacher maintain the same level of contingent scaffolding in a far larger group. Research on how to extend the benefits of scaffolding to a whole class has begun in the domain of middle school mathematics (e.g. Smit, van Eerde, & Bakker, 2013), and extensive research exists on instructional discourse and comprehension in English classes (e.g. Applebee, Langer, Nystand & Gamoran, 2003; Murphy, Wilkinson, Soter, Hennessey & Alexander, 2009; Nystrand, 2006). Future research might consider manipulating group sizes

as a way to consider whether contingent scaffolding can be scaled up to whole-class settings or whether smaller groups are necessary for contingency.

Implications for Researchers and Educators

These findings support the argument that, despite the relatively small yearly progress on standardized reading for 11th and 12th grade students (Hill et al., 2008) and relative lack of existing intervention models, interventions can make measurable progress in supporting those students' reading comprehension. While asserting that small-group reading instruction can help students may seem obvious, the research base for interventions with older adolescents has documented very few successful interventions. The contrasting findings of the intervention's significant effect on passage-based comprehension but insignificant effects on sentence-based reading efficiency also suggests that future interventions should consider whether comprehension or efficiency is the more desirable outcome.

While the evidence suggests that the intervention as a whole was effective, and scaffolding contributed to a meaningful portion of outcome variance, it challenging to specify the degree to which each of the elements of the intervention (complex texts, paraphrasing, or scaffolding) accounted for the students' improvement. The mixed evidence about the effectiveness of different kinds of scaffolding hints that it may not have been the sole contributor to student growth. Future research might consider systematically varying the complexities of texts used or the focal instructional activities in order to parse out the contributions of text and activity. However, as the scaffolding in this study was designed to account for text complexity and the paraphrasing activity selected to prioritize eliciting students' situation definitions and prioritize their existing language resources, future studies should consider how the text features and the instructional activities align with their approach

to scaffolding.

These findings also show that scaffolding can support complex text reading even without highly experienced teachers delivering the intervention. Though research has shown that developing contingent scaffolding skill is challenging for teacher educators to develop in novice teachers (Hedin & Gaffney, 2013, Rodgers et al., 2016), research has also shown that than instruction need not be perfectly contingent to support student learning (Rodgers et al., 2015; Van de Pol & Elbers, 2013, Wood & Wood, 1996). While the data from this intervention do not allow for claims about the degree of contingency enacted in the instruction, it may be that the training provided tutors with enough contingent skills to push their students toward increasingly coherent comprehension. As research on contingency in comprehension scaffolding is still in its infancy, and scaffolding measurement has been identified as a challenge to researchers (Van de Pol et al., 2010), future research might investigate routes to better measure and operationalize teachers' contingency.

In addition, many of the empirical tests of interventions for adolescent readers in the last decade have focused on programs combining various amounts of explicit instruction in multiple comprehension strategies, explicit vocabulary instruction, and independent student reading practice (Boulay et al., 2015; Paul & Clarke, 2016), all of which were not included in this intervention. Perhaps a scaffolding-based intervention is better suited to build on students' existing strategic resources and respond to students' emerging comprehension, rather than prescribing a set of predetermined planned scaffold strategies. The success of this intervention suggests that researchers seeking instructional models for comprehension might consider a scaffolding-oriented model instead of a strategies-oriented model.

Findings from the correlational analyses suggest, also, that the kinds of scaffolds

matter. In striving to be contingent on student responses, tutors had an array of scaffolds to choose from, and the array of scaffolds designed to address the comprehension of complex texts is a contribution to teaching practice literature. The list of scaffolds presented in Appendix C are written to be grounded in theories of comprehension and academic language, broadly applicable across texts in different domains, usable in different schools and districts, and helpful for different kinds of students. Teachers might use the scaffold list to develop and refine their existing practices of interactional scaffolding. Teachers could self-report the scaffolds to track their usage, or consider using colleagues to observe and track their scaffolds. Or, teachers might videotape their instruction and bring the list of scaffolds to a video club or lesson study group, using the scaffold list as a way to uncover glimmers of emergent comprehension in their students' classroom talk and to identify particular kinds of instructional moves that may be helpful for particular groups of students. In applying the scaffolds, teachers are encouraged to use them contingently as possible, knowing that students have valuable linguistic resources from which to elicit emerging comprehension.

Findings indicate, however, that not all scaffolds were equally effective. The strength and significance of the association between structure scaffolding and passage-based comprehension outcomes could be the most practical finding for teachers. This finding is consistent with a meta-analysis indicating that interventions teaching text structures support student comprehension (e.g. Hebert, Bohaty, Nelson & Brown, 2016). This study's approach, however, did not explicitly teach students about kinds of text structure or skills in ascertaining texts' structure—instead, tutors applied these scaffolds contingently. Teachers might be able to derive similar benefits in comprehension by contingently responding to students' knowledge during reading and during class discussion without spending precious class time

on explicit instruction.

On a cautionary note, when developing their practices of scaffolding in their classrooms or during small-group instruction, teachers should consider limiting their reliance on rereading texts to students as a scaffold. Teachers might consider ways in which they can respond to students' emerging comprehension in a nuanced and contingent way, rather than merely asking for rereading. To develop this skill, teachers might watch video of textual discussion and practice identifying students' emergent (but not completely coherent) comprehensions and consider different discursive routes to build on that comprehension. In addition, teachers should be careful in implementing their academic-language scaffolding: academic-register scaffolding, though hypothesized to be a helpful set of scaffolds in this study, actually predicted lower comprehension scores, possibly because they may have been enacted in a way that focused on ascertaining definitions of words outside students' lexicon. While it may be that optimally enacted academic-register scaffolding could help students recognize the flexible registers of academic vocabulary words and therefore support their comprehension, teachers should exercise caution with these scaffolds based on this evidence.

Directions for Future Research

Extending the current study. The next phase of analysis in this study involves both cross-sectional and longitudinal analyses of the video data collected during instruction. Table 9 demonstrates how, given the same texts and the same activity frame, two students paraphrased a complex sentence in different ways. The cross-sectional analyses examining the six groups participating in the same lesson, enabled by the consistent curriculum, will allow comparative examination of the variation in how six groups of students paraphrased the same texts differently. To support teachers in responding to student comprehension,

researchers must be able to show them the many ways it can appear—especially because research has shown that there are often meaningful insights in what initially appear to be partial or nonstandard interpretations (Aukerman, 2007, 2008). While local variation in discourse will undoubtedly influence students' paraphrases, helpful patterns in student responses may exist that could help teachers recognize and build on student thinking. Research has not documented such patterns.

On the other hand, longitudinal analysis might show how individual students' paraphrases evolved over time. Video evidence may be able to show increased coherence in students' paraphrases, or may be able to show increasingly sophisticated language use. Adams (2010) claimed that increasingly complex text provides students the "language, information, and modes of thought they need most to move up and on" (p. 6), but what happens when students are in direct contact with that language in the complex texts? The video data from this intervention might be able to answer a research question like: During direct interaction with complex texts over a period of eight lessons, how does students' language use respond to the complexity of the texts and the scaffolding of tutors?

Potential avenues for future research. This dissertation study represents a small step toward a larger program of research aimed at improving high school students' reading. Certainly, research must avoid the inoculation fallacy (Snow & Moje, 2010) and continue to investigate how the specific demands of comprehension at the late high school level can be supported by instruction and intervention. While researchers might have hoped that the inoculation metaphor's approach would be effective—for example, Connor and colleagues' intervention (2013) specifically looked for inoculation effects—it appears more and more that a better metaphor for reading instruction would be an exercise approach: rather than a single

shot of medicine protecting students for life, regularly exercising (that is, reading) with students, helping them with their unique struggles, leveraging their existing strengths, pushing them to work hard, and accounting for their changing developmental needs may be a better conceptualization of the task facing reading researchers and educators.

The design of this study responds to recent calls for investigating contingent scaffolding by holding planned scaffolding fixed to isolate the effect of interventional scaffolding (Reynolds & Daniel, 2017; Rodgers et al., 2016). Future studies evaluating the implementation of reading curricula might consider gathering data on the interactional scaffolding that happens within a curriculum and considering the variation in that scaffolding not as implementation failure but as adaptation to student needs. These designs are just emerging in the scaffolding literature, and more are needed to more deeply understand the phenomenon of contingent scaffolding in reading.

Another potential research design might incorporate measures of academic language such as the CALS (Uccelli et al. 2015) as additional outcomes, investigating the links between interventions specifically tailored to capitalize on academic language findings. In fact, if in future designs, students are pre-assessed for their academic language skills, tutors might have a clearer sense of how to respond to their students' needs. Or, a measure like the CALS might serve as a near-transfer outcome measure—discussion of complex texts might first support students' emerging academic language before supporting their reading comprehension. Such investigations might clarify the relationships between comprehension, text complexity, and academic language.

Conclusion

This study offers rigorous, experimental evidence of how a scaffolding-based

intervention using complex texts and paraphrasing supported high school students' reading comprehension. Findings also showed that structure and syntax scaffolding were meaningful factors in students' reading comprehension growth, and that academic register and rereading scaffolds predicted comprehension decline. While research has somewhat neglected late high school students' reading comprehension, these results can help researchers design effective instructional models and teachers find ways to honor their students' existing means of comprehending complex texts and scaffold them towards ever greater understanding.

REFERENCES

ACT Inc. (2006). Reading between the lines. Iowa City, IA: Author. Retrieved from http://eric.ed.gov/?id=ED490828.

ACT Inc. (2014). *Technical Manual: The ACT.* Retrieved from https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

Adams, M. J. (2011). Advancing our students' language and literacy: The challenge of complex texts. *American Educator*, *34*(4), 3.

Alfassi, M. (1998). Reading for meaning: The efficacy of reciprocal teaching in fostering reading comprehension in high school students in remedial reading classes. *American Educational Research Journal*, *35*(2), 309-332.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P.D. Pearson, R. Barr, M. Kamil & P. Mosenthal (Eds) *Handbook of Reading Research* Vol 1, (pp. 255-291).

Applebee, A. N., & Langer, J. A. (1983). Instructional scaffolding: Reading and writing as natural language activities. *Language Arts (60)* 2, 168-175.

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal*, *40*(3), 685-730.

Athanases, S. Z., & de Oliveira, L. C. (2014). Scaffolding versus routine support for Latina/o youth in an urban school: tensions in building toward disciplinary literacy. *Journal of Literacy Research*, *46*(2), 263-299.

Aukerman, M. (2007). When reading it wrong is getting it right: Shared evaluation pedagogy among struggling fifth grade readers. *Research in the Teaching of English*, 56-103.

Aukerman, M. (2008). In praise of wiggle room: Locating comprehension in unlikely places. *Language Arts*, *86*(1), 52-60.

Aukerman, M. (2013). Rereading comprehension pedagogies: Toward a dialogic teaching ethic that honors student sensemaking. *Dialogic Pedagogy,1* (pp. A1-A31).

Aukerman, M. (2015). How should readers develop across time? Mapping Change without a deficit Perspective. *Language Arts*, *93*(1), 55.

Aukerman, M., Martin, P. C., Gargani, J., & McCallum, R. D. (2016). A randomized control trial of Shared Evaluation Pedagogy: The near-term and long-term impact of dialogically organized reading instruction. *L1 Educational Studies in Language and Literature*, *16.* 1-26.

Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition–Implications for the design of computer-based scaffolds. *Instructional Science*, *33*(5), 367-379.

Bakhtin, M.M. (1994) *The Bakhtin Reader*. Ed. P. Morris. London: Edward Arnold.

Belland, B. R., Kim, C., & Hannafin, M. J. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist*, *48*(4), 243-270.

Belland, B. R. (2014). Scaffolding: Definition, current debates, and future directions. In Spector, J. M., Merrill, M. D., Elen, J., & Bishop, M. J. (Eds.). *Handbook of Research on Educational Communications and Technology*. (pp. 505-518). New York: Springer.

Billings, L., & Fitzgerald, J. (2002). Dialogic discussion and the Paideia seminar. *American Educational Research Journal*, *39*(4), 907-941

Boulay, B., Goodson, B., Frye, M., Blocklin, M., & Price, C. (2015). *Summary of Research Generated by Striving Readers on the Effectiveness of Interventions for Struggling Adolescent Readers.* NCEE 2016-4001. National Center for Education Evaluation and Regional Assistance. Retrieved from https://ies.ed.gov/ncee/pubs/20164001/pdf/20164001.pdf

Boyd, M., & Rubin, D. (2006). How contingent questioning promotes extended student talk: A function of display questions. *Journal of Literacy Research, 38*(2), 141-169.

Brabham, E., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94*(3), 465.

Brush, T. & Saye, J. W. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development, 50*(3), 77-96.

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology, 76*(4), 683-696.

Cantrell, S. C., Almasi, J. F., Carter, J. C., Rintamaa, M., & Madden, A. (2010). The impact of a strategy-based intervention on the comprehension and strategy use of struggling adolescent readers. *Journal of Educational Psychology, 102*(2), 257.

Celani, K., McIntyre, E., & Rightmyer, E. C. (2006). Knowing the text, knowing the learner: Literature discussions with fifth grade struggling readers. *Reading Horizons, 47*(2), 97.

Chall, J. S., Bissex, G. L., Conard, S. S., & Harris-Sharples, S. H. (1996). *Qualitative assessment of text difficulty: A practical guide for teachers and writers*. Boston: Brookline Books.

Christoph, J. N., & Nystrand, M. (2001). Taking risks, negotiating relationships": One teacher's transition toward a dialogic classroom". *Research in the Teaching of English*, 249-286.

Clark, K. F., & Graves, M. F. (2005). Scaffolding students' comprehension of text. *The Reading Teacher*, *58*(6), 570-580.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale. NJ: Lawrence Erlbaum Associates,

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.

Crain-Thoreson, C., Lippman, M. Z., & McClendon-Magnuson, D. (1997). Windows on comprehension: Reading comprehension processes as revealed by two think-aloud procedures. *Journal of Educational Psychology*, *89*(4), 579.

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students?. *Journal of Research in Reading*, *36*(3), 241-260.

Crowe, L. K. (2003). Comparison of two reading feedback strategies in improving the oral and written language performance of children with language-learning disabilities. *American Journal of Speech-Language Pathology*, *12*(1), 16-27.

Daniel, S. M., Martin-Beltrán, M., Peercy, M. M. and Silverman, R. (2015), Moving beyond *Yes or No*: Shifting from over-scaffolding to contingent scaffolding in literacy instruction with emergent bilingual students. *TESOL Journal*. *7*(2), 393-420.

Dimitrov, D. M., Frye, M., Jurich, S., Sayko, S., & Lammert, J. (2014). A study on immediate effects of the first-year intervention under the Illinois Striving Readers project. *Journal of Education and Human Development*, *3*(2), 121-154.

Duke, N. K., & Pearson, P. D. (2008). Effective practices for developing reading comprehension. *The Journal of Education*, *189*(1/2), 107-122.

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.

Faddis, B. J., Beam, M., Maxim, L., Gandhi, E. V., Hahn, K., & Hale, R., (2011). *Portland Public School's Striving Readers Program: Year 5 Evaluation Report.* Prepared for the U.S. Department of Education Office of Elementary and Secondary Education.

Fisher, D., & Frey, N. (2014). Addressing CCSS anchor standard 10: Text complexity. *Language Arts*, *91*(4), 236-250.

Frey, N., & Fisher, D. (2010). Identifying instructional moves during guided learning. *The Reading Teacher*, *64*(2), 84-95.

Frankel, K. K., Becker, B. L., Rowe, M. W., & Pearson, P. D. (2016). From "What is Reading?" to “What is Literacy?”. *Journal of Education*, *196*(3) 7-17.

Frantz, R. S., Starr, L. E., & Bailey, A. L. (2015). Syntactic complexity as an aspect of text complexity. *Educational Researcher*, *44*(7), 387-393.

Gaskins, I. W., Rauch, S., Gensemer, E., & Cunicelli, E. (1997). Scaffolding the development of intelligence among children who are delayed in learning to read. In K.E. Hogan and M.E. Pressley (Eds.) *Scaffolding Student Learning: Instructional Approaches and*

Issues. (pp. 43-73)

Glenberg, A. M. (2011). How reading comprehension is embodied and why that matters. *International Electronic Journal of Elementary Education*, *4*(1), 5-18.

Glover, J. A., Plake, B. S., Roberts, B., Zimmer, J. W., & Palmere, M. (1981). Distinctiveness of encoding: The effects of paraphrasing and drawing inferences on memory from prose. *Journal of Educational Psychology*, *73*(5), 736.

Goldman, S. R., & Lee, C. D. (2014). Text complexity. *The Elementary School Journal*, *115*(2), 290-300.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, *60*(2), 183-208

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257-285.

Hacker, D. J., & Tenent, A. (2002). Implementing reciprocal teaching in the classroom: Overcoming obstacles and making modifications. *Journal of Educational Psychology*, *94*(4), 699.

Hagaman, J. L., & Reid, R. (2008). The effects of the paraphrasing strategy on the reading comprehension of middle school students at risk for failure in reading. *Remedial and Special Education*, *29*(4), 222-234.

Hagaman, J. L., Casey, K. J., & Reid, R. (2012). The effects of the paraphrasing strategy on the reading comprehension of young students. *Remedial and Special Education*, *33*(2), 110-123.

Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect (20)*1. 6-30.

Handsfield, L. J., Crumpler, T. P., & Dean, T. R. (2010). Tactical negotiations and creative adaptations: the discursive production of literacy curriculum and teacher identities across space-times. *Reading Research Quarterly*, *45*(4), 405-431.

Harris, K. & Pressley, M (1991). The nature of cognitive strategy instruction: Interactive strategy construction. *Exceptional Children,* (57)5, 392-404.

Haynes, J. E., & Fillmer, H. T. (1984). Paraphrasing and reading comprehension. *Literacy Research and Instruction*, *24*(1), 76-79.

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology*, *108*(5) 609-629.

Hedin, L.R. & Gaffney, J.S. (2013). Tutoring sixth graders who struggle with reading: illustration of Wood's contingent interventions. *Reading Psychology, 34*(3), 207-256.

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.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172-177.

Hirsch, E. D. (2003). Reading comprehension requires knowledge—of words and the world. *American Educator*, *27*(1), 10-13.

Hofstetter, C.H., Strick, B., Der-Martirosian, C., Ong-Dean, C., Long, P., & Lou, Y. (2011). Implementation and Impact of the Targeted and Whole School Interventions,

Summary of Year 4 (2009-2010). San Diego Unified School District. San Diego, CA. Report submitted to the U.S. Department of Education, Office of Elementary and Secondary Education

Hua, Y., Woods-Groves, S., Ford, J. W., & Nobles, K. A. (2014). Effects of the paraphrasing strategy on expository reading comprehension of young adults with intellectual disability. *Education and Training in Autism and Developmental Disabilities*, *49*(3), 429.

Jadallah, M., Anderson, R. C., Nguyen-Jahiel, K., Miller, B. W., Kim, I. H., Kuo, L. J., ... & Wu, X. (2011). Influence of a teacher's scaffolding moves during child-led small-group discussions. *American Educational Research Journal*, *48*(1), 194-230.

Jadallah, M., Anderson, R. C., Nguyen-Jahiel, K., Miller, B. W., Kim, I. H., Kuo, L. J., ... & Wu, X. (2011). Influence of a teacher's scaffolding moves during child-led small-group discussions. *American Educational Research Journal*, *48*(1), 194-230.

Katims, D. S., & Harris, S. (1997). Improving the reading comprehension of middle school students in inclusive classrooms. *Journal of Adolescent & Adult Literacy*, *41*(2), 116-123.

Kong, A., & Pearson, P. D. (2003). The road to participation: The construction of a literacy practice in a learning community of linguistically diverse learners. *Research in the Teaching of English*, 85-124.

LaHuis, D. M., Hartman, M. J., Hakoyama, S., & Clark, P. C. (2014). Explained variance measures for multilevel models. *Organizational Research Methods*, *17*(4), 433-451.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics(33)* 1, 159-174.

Lauterbach, S. L., & Bender, W. N. (1995). Cognitive strategy instruction for reading comprehension: A success for high school freshmen. *The High School Journal, 79*(1), 58-64.

Lee, C. D. (1995). A culturally based cognitive apprenticeship: Teaching African American high school students skills in literary interpretation. *Reading Research Quarterly*, 608-630.

Lee, C.D. (2000) Signifying in the zone of proximal development: literate cultural competencies in language minority classrooms. In Lee, C. D., & Smagorinsky, P.(Eds). *Vygotskian perspectives on literacy research: Constructing meaning through collaborative inquiry*. Cambridge: Cambridge University Press.

Lee, C. D. (2001). Is October Brown Chinese? A cultural modeling activity system for underachieving students. *American Educational Research Journal*, *38*(1), 97-141.

Lin, T. J., Jadallah, M., Anderson, R. C., Baker, A. R., Nguyen-Jahiel, K., Kim, I. H., ... & Wu, X. (2015). Less is more: Teachers' influence during peer collaboration. *Journal of Educational Psychology, 107*(2), 609.

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.

Maloch, B. (2002). Scaffolding student talk: One teacher's role in literature discussion groups. *Reading Research Quarterly*, *37*(1), 94-112.

Maloch, B. (2004). On the road to literature discussion groups: Teacher scaffolding during preparatory experiences. *Literacy Research and Instruction*, *44*(2), 1-20.

Maloch, B. (2008). Beyond exposure: The uses of informational texts in a second grade

classroom. *Research in the Teaching of English 42*(3), 315-362.

Many, J. E. (2002). An exhibition and analysis of verbal tapestries: Understanding how scaffolding is woven into the fabric of instructional conversations. *Reading Research Quarterly, 37*(4), 376-407.

Mariage, T. V. (1995). Why students learn: The nature of teacher talk during reading. *Learning Disability Quarterly*, *18*(3), 214-234.

McElhone, D. (2012). Tell us more: Reading comprehension, engagement, and conceptual press discourse. *Reading Psychology*, *33*(6), 525-561.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276-282.

Mehan, H. (1979). *Learning lessons*. Cambridge, MA: Harvard University Press.

Meskill, C. (2005). Triadic scaffolds: Tools for teaching English language learners with computers. *Language Learning & Technology*, *9*(1), 46-59.

Mertzman, T. (2008). Individualising scaffolding: teachers' literacy interruptions of ethnic minority students and students from low socioeconomic backgrounds. *Journal of Research in Reading, 31*(2), 183-202.

Meyer, B. J., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 16* (1)72-103.

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

Moss, B., Lapp, D., & O'Shea, M. (2011). Tiered texts: Supporting knowledge and language

learning for English learners and struggling readers. *English Journal (100) 5*, 54-60.

Murnane, R. J., & Willett, J. B. (2010). *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. Oxford: Oxford University Press.

Murphy, P. K., Wilkinson, I. A., 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.

National Center for Education Statistics (2015) *The Nation's Report Card*. Washington DC: US Department of Education, Institute of Education Sciences. Retrieved from http://nces.ed.gov/nationsreportcard/reading

National Governors Association for Best Practices & Council of Chief State School Officers (2010) Appendix A: Common Core State Standards for English language arts in history/social studies, science, and technical subjects. Washington, DC.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (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: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Nystrand, M. (2006). Research on the role of classroom discourse as it affects reading comprehension. *Research in the Teaching of English 40*(4),, 392-412.

Palincsar, A.S. & Brown, A.L. (1984) Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction (1)*2, 117-175.

Palincsar, A. S. (1986). The role of dialogue in providing scaffolded instruction. *Educational Psychologist*, *21*(1-2), 73-98.

Paige, D. D., Rasinski, T. V., & Magpuri-Lavell, T. (2012). Is fluent, expressive reading important for high school readers?. *Journal of Adolescent & Adult Literacy*, *56*(1), 67-76.

Paul, S. A. S., & Clarke, P. J. (2016). A systematic review of reading interventions for secondary school students. *International Journal of Educational Research*, *79*, 116-127.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences*, *13*(3), 423-451.

Pearson, P. D., & Gallagher, G. (1983). The gradual release of responsibility model of instruction. *Contemporary Educational Psychology*, *8*(3), 112-123.

Pearson, P. D., & Hiebert, E. H. (2014). The state of the field: Qualitative analyses of text complexity. *The Elementary School Journal*, *115*(2), 161-183.

Poole, D. (2008). Interactional differentiation in the mixed-ability group: A situated view of two struggling readers. *Reading Research Quarterly*, *43*(3), 228-250.

Poulsen, M., & Gravgaard, A. K. (2016). Who did what to whom? The relationship between syntactic aspects of sentence comprehension and text comprehension. *Scientific Studies of Reading*, *20*(4), 325-338.

Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, *40*(1), 1-12.

Rasinski, T., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for the primary grades?. *Literacy Research and*

Instruction, *48*(4), 350-361.

Reed, D. K., Sorrells, A. M., Cole, H. A., & Takakawa, N. N. (2013). The ecological and population validity of reading interventions for adolescents: Can effectiveness be generalized?. *Learning Disability Quarterly, 36*(3), 131-144.

Reisman, A. (2015). Entering the historical problem space: Whole-class text-based discussion in history class. *Teachers College Record, 117*(2),1-44.

Reynolds, D., & Goodwin, A. P. (2016a). Making complex texts a reality for all students: Dynamic scaffolding that bridges the gaps between student and text. *Voices from the Middle*, *23*(4), 25.

Reynolds, D., & Goodwin, A. (2016b). Supporting students reading complex texts: Evidence for motivational scaffolding. *AERA Open*, *2*(4) 1-16,

Reynolds, D. (2017) Interactional scaffolding for reading comprehension: A systematic review. *Literacy Research: Theory, Method, and Practice*. doi: 10.1177/2381336917718820

Reynolds, D., & Daniel, S. (2017). Toward contingency in scaffolding reading comprehension: Next steps for research. *Reading Research Quarterly*. doi: 10.1002/rrq.200.

Rodgers, E., D'Agostino, J. V., Harmey, S. J., Kelly, R. H., & Brownfield, K. (2016). Examining the nature of scaffolding in an early literacy intervention. *Reading Research Quarterly*, *51*(3), 345-360.

Roehler, L. R. & Cantlon, D.J. (1997) Scaffolding: A powerful tool in social constructivist classrooms. In Hogan, K. E., & Pressley, M. E. (Eds). *Scaffolding Student Learning: Instructional Approaches and Issues* (pp. 6-42) Cambridge: Brookline Books.

Rosenblatt, L. M. (1982). The literary transaction: Evocation and response. *Theory into Practice*, *21*(4), 268-277.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, *64*(4), 479-530.

Sánchez, E., & García, J. R. (2009). The relation of knowledge of textual integration devices to expository text comprehension under different assessment conditions. *Reading and Writing*, *22*(9), 1081-1108.

Sanden, S. (2014). Text complexity: A new way forward or two steps back? *Language Arts*, *(25) 2*, 7-14.

Scammacca, N., Vaughn, S., Roberts, G., Wanzek, J., & Torgesen, J. K. (2007). Extensive reading interventions in grades K-3: From research to practice. *Center on Instruction*. Retrived from http://files.eric.ed.gov/fulltext/ED521573.pdf

Schiller, E., Wei, X., Thayer, S., Blackorby, J., Javitz, H., & Williamson, C. (2012). A randomized controlled trial of the impact of the fusion reading intervention on reading achievement and motivation for adolescent struggling readers. *Society for Research on Educational Effectiveness*. Retrieved from http://files.eric.ed.gov/fulltext/ED535544.pdf.

Scott, C. M. (2009). A case for the sentence in reading comprehension. *Language, Speech, and Hearing Services in Schools*, *40*(2), 184-191.

Searle, D. (1984). Scaffolding: Who's building whose building?. *Language Arts*, *61*(5), 480-483. Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen’s f2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology, 3*, 1-6.

Segal, A., Snell, J., & Lefstein, A. (2016). Dialogic teaching to the high-stakes standardised test?. *Research Papers in Education*, 1-15. doi:10.1080/02671522.2016.1225803

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, *43*(3), 290-322.

Smit, J., AA van Eerde, H., & Bakker, A. (2013). A conceptualisation of whole-class scaffolding. *British Educational Research Journal*, *39*(5), 817-834.

Snijders, T., & Bosker, R. (2012). Multilevel analysis: An introduction to basic and applied multilevel analysis. London: Sage.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In Olson, D.R, & N. Torrance (Eds.) *The Cambridge Handbook of Literacy*, (pp. 112-133). Cambridge: Cambridge University Press.

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

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, *328*(5977), 450-452.

Sprague, K., Zaller, C., Kite, A., & Hussar, K. (2011). *Springfield-Chicopee School Districts Striving Readers (SR) program year 4 report: Evaluation of implementation and impact (pp. 1-130)* (United States, Institute of Education Sciences, Elementary and Secondary Education). Providence, RI: The Educational Alliance at Brown University. Retrieved from http://www2.ed.gov/programs/strivingreaders/springfieldeval42011.pdf

StataCorp (2015) *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Stone, C. A. (1998). The metaphor of scaffolding its utility for the field of learning disabilities. *Journal of Learning Disabilities*, *31*(4), 344-364.

Swanlund, A., Dahlke, K., Tucker, N., Kleidon, B., Kregor, J., Davidson-Gibbs, D., & Hallberg, K. (2012). Striving Readers: Impact Study and Project Evaluation Report: Wisconsin Department of Public Instruction (with Milwaukee Public Schools). Naperville, IL: American Institutes for Research.

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, *50*(3), 337-356.

US Department of Education, Institute of Education Sciences (2017) *What Works Clearinghouse.* Retrieved from http://whatworks.ed.gov.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*(3), 271-296.

Vygotsky, L.S. (1978). *Mind in Society*. Cambridge: Harvard UP.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., & Pearson, N.A. (2010). *Test of silent reading efficiency and comprehension*. Austin, TX; Pro-ED, Inc.

Wanzek, J., Vaughn, S., Scammacca, N., Metz, K., Murray, C., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for older struggling readers: Implications from research. *Review of Educational Research, 83*, 163-195.

Welie, C., Schoonen, R., Kuiken, F., & Bergh, H. (2016). Expository text comprehension in secondary school: for which readers does knowledge of connectives contribute the most? *Journal of Research in Reading*. doi:10.1111/1467-9817.12090

Wertsch, J. V. (1984). The zone of proximal development: Some conceptual issues. *New*

Directions for Child and Adolescent Development, *1984*(23), 7-18.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100.

Wortham, S. E. (1995). Experiencing the great books. *Mind, Culture, and Activity*, *2*(2), 67-80.

Wood, D., & Wood, H. (1996). Contingency in tutoring and learning [Commentary]. *Learning and Instruction*, 6(4), 391–397.

Appendix A: Sample Lesson Plan (Lesson #3)

LESSON PLAN 3

BUILD RAPPORT, REVIEW NAMES/NORMS 8:33-8:37 FIDELITY CHECK: ☐

Hey everyone! Welcome to the third session—over the last two sessions we've read about dinosaurs' intelligence, about an amazing jazz music performance, and today we're going to read about landforms on the bottom of the ocean. But before we get to that, let's do a quick name doublecheck: how fast can you name everyone in the group? I'll go first [point to each kid and name them]. **Now your turn!**

[Take 1min and call on each kid to name the other kids in the group. It's okay if they can't do it, but remind them that the goal is to know everyone in the group]

Think back to yesterday and the norms about talking out the text that we talked about. They're important that we stick to these so we can have a group that works together well and supports each other in figuring out hard texts.

Have the students read all three aloud, and then take 30s to discuss all three, emphasizing the ones you think are the most important. Leave these out on the table as you work together.

- Everyone takes turns talking out text.
- Taking guesses and making mistakes is part of learning.
- No talking down others' responses: build on them.

INTRODUCE TEXT 3 8:37-8:40 FIDELITY CHECK: ☐

OK team, we're going to get started with today's question. (Have the whole group read the question of the day aloud: *What landforms exist on the bottom of the ocean?)*. **Today's text is going to answer that question—it's from a nonfiction book called *Great Waters: An Atlantic Passage* by Deborah Cramer.**

Before we get started, what do we already know about landforms on the bottom of the ocean?

[2min discussion using any of the following questions. Goal is to activate their prior knowledge.]

- What kinds of landforms exist on the bottom of the ocean?
- Are the landforms on the bottom of the ocean the same as those on land, or are they different?
- How do we even know what's down there?
- Who are the experts who study undersea landforms?
- What happens when a volcano erupts under the ocean?

REALLY BRIEF PARAPHRASING MODELING 8:40-8:41 FIDELITY CHECK: ☐

OK, let's get started with the reading. Think back to the paraphrasing we did yesterday. What does it mean to paraphrase? (Listen to their responses)**.**

Note the little blurb at the top—"The Sargasso Sea is part of the northern Atlantic Ocean". Sometimes these texts give you these little bits of information so you can make sense of what the whole thing is about. So let's remember that about the Sargasso Sea. Now let's get started!

[Have a student read the 1st sentence: *As the Cramer* idles through the Sargasso Sea, waiting for the wind to rise, the sea is flat and empty.]

Well, there's something in the ocean, I wonder why the word *Cramer* is in italics, but if it's waiting for wind, I bet it's a boat, and there isn't much going on out in the middle of the ocean.

STUDENTS **READ AND PARAPHRASE** TEXT 2 8:41-9:04 FIDELITY CHECK: ☐

OK, now that you are all getting better at paraphrasing, you guys are going to take over the reading as well. So one of you will read a sentence aloud, and the next student will paraphrase that. Once we've all gotten to a good paraphrase, then another

[Have a student read: *Nothing demarcates or divides the smooth expanse of water dissolving into the horizon",* and then have another student paraphrase it. Scaffold as necessary.]

> **At this point, you are in control of the reading flow. If your students are struggling **a lot**, then you can take over the reading yourself and give them a little more breathing room. You should, however, err on the side of letting them struggle—that's a key part of learning to read complex texts. The goal is to *transfer responsibility* to them as quickly as you can.

[Students read to you and paraphrase. You retain the control over whether a reading or a paraphrase is acceptable enough for the group to continue.

Try to be as *contingent* as possible—for groups that are really doing a great job making coherent sense of these texts, you can push them on hard things, like analyzing what "demarcates" means in the second sentence. On the other hand, for groups that are struggling to get to coherence, you can skip "demarcates" and just focus on "divides". You should have about 20 minutes to work your way through the text, so modulate the difficulty as you find appropriate to your group—maximize the challenge without causing frustration!

CLOSURE 9:04:9:07 FIDELITY CHECK: ☐

So now that we've talked our way through the whole thing, what happened? What the heck is actually on the bottom of the ocean floor?

[Discuss the passage, including any details you think relevant to answering that big question. You can use some or all of the following questions.

- What did you think of this passage?
- How have people learned about landforms on the bottom of the ocean—both historically and currently?
- What exactly is down there?
- Is the stuff down there similar to the stuff that's up here?
- Were you surprised to find out that there are giant mountains and really long trenches there?
- Does the author do an effective job at building suspense by contrasting the calm sea at the beginning with the spectacular stuff underneath?
- Do you think fish appreciate just how cool the ocean is, or do they just blub blub blub all day long?

OK, so we've practiced paraphrasing and talking through the text, and we've now learned about dinosaurs, jazz music, and the bottom of the ocean. Next session we'll be learning about how a writer works to mix cultural traditions in her stories. It's going to be great!

[Say goodbye to each student by name, reinforcing your personal connection with them. Something like "_____, great work today, see you tomorrow!"

FIDELITY NOTES ON THIS SESSION

of lines of the text read _____

Was all of the text read? Yes No

Was there a concluding discussion after all the text was read? Yes No

Who did the reading aloud today? (1-10, 1=teacher read all aloud, 10=students read all aloud) _____

Did you have to skip/modify any of the parts of the lesson plans? If so, why?

What did you talk about during the concluding discussion?

APPENDIX B: Table of Texts and Sample Text from Lesson #3

Lesson	Genre	Text Excerpt From	Big question*
1	Natural science	Stephen Jay Gould "Were dinosaurs dumb?"	Were dinosaurs dumb?
2	Prose fiction	Paule Marshall *The Fisher King*	How does a jazz pianist unexpectedly deliver an amazing performance?
3	Social science	Deborah Cramer *Great Waters, An Atlantic Passage*	What landforms are on the bottom of the ocean?
4	Humanities	Bharati Mukherjee "A Four-Hundred-Year-Old Woman"	How does this writer deal with transnational cultural identities?
5	Natural science	Oliver Sacks "An Anthropologist on Mars"	(How do people see color?)
6	Prose fiction	Cristina Perri Rossi "The Threshold"	(How do these two friends talk about their dreams?)
7	Social science	Joseph Ellis *American Sphinx: The Character of Thomas Jefferson*	(Did Thomas Jefferson plagiarize the Declaration of Independence?)
8	Humanities	Jon Pareles "India Resounding in New York"	(How does Indian and American music mix?)

**Note: "Big questions" were only required to be used in sessions 1-4, and were used contingently thereafter.

Sample Text Used in Lesson 3

SOCIAL SCIENCE: This passage is adapted from Great Waters: An Atlantic Passage by Deborah Cramer (©2001 by Deborah Cramer). The Sargasso Sea is a part of the northern Atlantic Ocean.

As the Cramer idles through the Sargasso Sea, waiting for the wind to rise, the sea is flat and empty. Nothing demarcates or divides the smooth expanse of water dissolving into the horizon. This vast, unroughened surface, this breadth of uniform sea, deceives. But for a few lonely oceanic islands, the unperturbed surface offers no hint of the grand and sweeping energies hidden below.

Only one thousand miles offshore, the Cramer has already sailed through some of Atlantic's deepest waters. Contrary to what one might guess, Atlantic's deepest waters, like those in other oceans, are along her edges. As we continue east, toward the middle of the sea, the bottom rises. The unmarked plains of the abyss, here flattened by layers of sediment, give way to rising foothills and then to mountains. The first maps of Atlantic seafloor noted, albeit crudely, this rise. Early efforts to plumb Atlantic's depths proved outrageously inaccurate: one naval officer paid out eight miles (thirteen kilometers) of hemp rope from a drifting ship and concluded the sea had no bottom. Eventually, sailors more or less successfully calculated depth by heaving overboard cannonballs tied to bailing twine. When they hit bottom, the sailors measured and snipped the twine and then moved on, leaving a trail of lead strung out across the seafloor. These crude soundings, forming the basis of the first map of Atlantic's basin, published in 1854, identified a prominent rise halfway between Europe and America.

For many years no one could explain why the basin of Atlantic, unlike a bowl, deepened at its edges and shoaled in its center. People assumed that this "Middle Ground," "Telegraph Plateau," or "Dolphin Rise," as it was variously called, was an ancient and

drowned land bridge, or a lost continent, but sailors repairing transatlantic telegraph cable unknowingly produced evidence to prove otherwise. Wrestling with the broken cable, they accidentally twisted off a piece of the "plateau" and dredged up a twenty-one-pound (ten-kilogram) chunk of dense black volcanic rock. It was some of the youngest, freshest rock on earth, and it was torn not from a piece of continent sunk beneath the waves, but from the very foundation of the sea.

Today, highly sophisticated sound waves bring the hazy images of those early soundings into sharp focus, revealing that one of the largest and most salient geographic features on the planet lies on the floor of the ocean. Hidden beneath the waves is an immense submerged mountain range, the backbone of the sea. More extensive, rugged, and imposing than the Andes, Rockies, or Himalayas, it covers almost as much of earth's surface as the dry land of continents. Winding like the seam of a baseball, it circles the planet in a long, sinuous path, running the entire length of Atlantic, slashing the basin neatly in two. Its mountains are stark and black, as black as the sea itself, lit only at their peaks by a thin, patchy covering of white, the skeletal remains of tiny microscopic animals that once lived at the sur- face. Peaks as high as Mount St. Helens sit in a watery world of blackness, more than a mile below the surface, beyond the reach of light, beyond the sight of sailors.

A great valley, eclipsing any comparable feature on dry land, runs through these mountains. Arizona's Grand Canyon, one of earth's most spectacular places, extends for about 280 miles (450 kilometers). A lesser known canyon of similar depth but considerably greater length lies hidden in the mountains of the ridge. Although offset in many places by breaks in the mountains, the rift valley, as the canyon is called, extends the length of Atlantic for 11,000 miles (17,700 kilometers). Here in this bleak and forbidding place, where the water is almost

freezing, subterranean fires have lifted mounds of fresh lava onto the seafloor. Scientists visiting the rift valley for the first time named the volcanic hills in this otherworldly setting after distant, lifeless planets.

Yet, what had seemed so foreign to scientists is an integral part of earth's very being, for at the ridge our own planet gives birth. The floor of the rift valley is torn; from the gashes has sprung the seafloor underlying all of Atlantic. Here the youngest, newest pieces are made. Earth is still cooling from her tumultuous birth four and a half billion years ago. Heat, leaking from the molten core and from radioactive decay deep inside the planet, rises toward earth's surface, powering the volcanoes that deliver the ridge to the sea.

APPENDIX C: Full Scaffold List

Extending talk (72% agreement, κ=0.44, category discarded due to insufficient reliability)

1. Tell me more. Extend your paraphrase.
2. Ask a student: "Tell me more. Extend your paraphrase. "
3. How can we use what we know to figure this out?
4. What is going on here?
5. Where is the evidence for your paraphrase?
6. Can you incorporate (textual detail) to your paraphrase?
7. Ask another student to comment on the first student's paraphrase.
8. Ask students to summarize paragraph(s)

Background knowledge (79% agreement, κ=0.54)

9. Providing definitions of vocabulary words
10. Providing a related word (e.g. *immigration* to help with *emigration*)
11. Tutor tells students his/her interpretation of a sentence
12. Tutor tells students his/her interpretation of a paragraph
13. Tutor tells students his/her interpretation of the passage

Morphology (CALS skill: Unpacking complex words –87% agreement, κ=0.57)

14. What do you think this word means? (point out a morphologically complex word)
15. What do the parts of that word tell you?
16. Point/box/mark particular morphemes
17. Ask student to point/box/mark particular morpheme

Academic register (CALS skill: Recognizing academic register – 82% agreement, κ=0.57

18. What do you think this word means? (point out a rare/discipline-specific word but not morphologically complex)
19. Point out key punctuation (e.g. italics, quotation marks)
20. What do you think this word means in *this context*? (point out a word that is not rare or discipline-specific, but that has alternate/tertiary meaning in context)
21. Does _____ mean __(primary meaning of the word)__ here?
22. What connotation/feeling does this word have here?
23. Highlight metaphorical/symbolic language—i.e. "Is __ literally ___ here?"

Syntax (CALS skill: Unpacking complex sentences—80% agreement, κ=0.51

24. What does that sentence mean? (point out a syntactically complicated sentence)
25. Ask student to break down sentence into parts and paraphrase each part
26. How do these different parts of the sentence work together?

Connections (CALS skill: Connecting ideas logically—82% agreement, κ=0.55)

27. How does this fit with the previous sentence/paragraph?
28. How has ____ changed? (since an earlier point in the text)
29. (point to transition word/phrase) What does this word tell you?

Tracking participants (CALS skill: Tracking participants and themes—82% agreement, κ=0.55)

30. Who is this/who did that? (Point to verb)
31. What does this pronoun/nominalization refer to? (Point to noun/pronoun)
32. How does this echo (insert earlier theme)?

Structure (CALS skill: organizing argumentative texts—90% agreement,κ=0.46)

33. What is the structure of this text?
34. How does this sentence fit in the structure?

35. How does this paragraph fit in the structure?
36. What does that detail do? Why do you think the author included that detail?

Motivation (Reynolds & Goodwin, 2016b—84% agreement, κ=0.59)

37. General praise, high-fives
38. Expression of confidence in student's ability ("You can do this")
39. Expression of student's independence ("You can do this on your own"/"You don't need me")
40. Praise for particular thinking move ("I like how you did X")____
41. Create competitions/races
42. Use humor
43. Refocusing students/calling to attention

Using other semiotic mediators as vectors towards intersubjectivity (Wertsch, 1984; 91% agreement, κ=0.55)

44. Act the text out with body or hands
45. Draw something out in a sketch
46. Map it out in a concept map or diagram
47. Use symbol markers to represent ideas or characters
48. Bring in pictures, video, or external media to help explain
49. Invent an analogy (X is like Y) to explain an idea

Rereading (84% agreement, κ=0.64)

50. Tutor re-reads a sentence to students.
51. Tutor asks students to re-read a sentence.

Appendix D: Training Materials

TRAINING DAY 1: LEARNING AND DELIVERING THE SCAFFOLDS

WELCOME, BREAKFAST, AND TEAM INTRODUCTIONS

Come on in, make a name tag, make a name card! Sit with some people you don't know and introduce yourself.

STUDY OVERVIEW: UNDERSTANDING SCAFFOLDING

Here's what we're doing: learning how to *contingently scaffold* students' comprehension while they are *paraphrasing complex texts*. In a way, we're using a hybrid of everyday language and *academic language.* Let's define the following terms.

Scaffolding:

Contingency:

Paraphrasing:

Academic language:

Complex texts:

Six dimensions of complexity:

STUDY MATERIALS: LESSON PLANS AND TEXTS

- Discuss example lesson plan
- Review example text from *Great Waters: An Atlantic Passage* by Deborah Cramer
- Read through scaffolds list

Discussion questions

- **What seems clear/unclear from the lesson plans?**
- **What scaffolds seem useful?**

SAMPLE TEXT: IDENTIFY OPPORTUNITIES

- Dan models: what does this look like?
- Look @ lines 1-40, and look at the scaffold list that you have.
    - Focus on the text-
- Think: take 7-10 minutes and read the text and try to annotate potential scaffolding opportunities.
- Pair: discuss with a partner for 5-10 minutes.
- Share: discuss with the whole group.

VIDEO EXAMPLE WITH TRANSCRIPT

- Watch the video and read the transcript. As you're watching, try to identify the scaffolds!

[00:15:57.23] Dan: [reading] "Today sophisticated sound waves bring those early sound waves into sharp focus, revealing that one of the largest and most salient geographic features on the planet lies on the floor of the ocean.".
[00:16:13.00] [pause]
[00:16:15.07] Dan: What's that mean?
[00:16:17.09] A: Means a large something under the ocean--
[00:16:21.02] Dan: There's a large somethin'...
[00:16:24.24] Dan: .. good.... what about that "Today"? What about the sound waves?
[00:16:29.20] L: They were highly sophisticated
[00:16:30.29] Dan: Yeah... what does that mean? Let's explain this to a six-year-old. A six-year old might not know what "sophisticated" means
[00:16:37.16] K: Interesting, smart, educated
[00:16:39.12] Dan: Yeah, really complicated, alright... so let's put it all together, we got some sound waves, we got the hazy images of those early soundings into sharp focus and we revealed there's somethin' ...
[00:16:50.24] Multiple students: [overlapping]
[00:16:52.13] L: More developing
[00:16:55.08] Dan: Okay, more developed sound waves, what about them?
[00:16:56.04] L: - brought-
[00:16:58.19] A:- revealed the image-
[00:17:01.15] L: - not so clear images
[00:17:05.00] T: [inaudible, to me at least]
[00:17:05.17] D: So the sound waves are making unclear images?
[00:17:06.26] L: Mmmhmm
[00:17:09.10] M: They send sound waves into the ocean, I mean they send like
[00:17:14.11] Dan: So when it says "those early soundings", what are they talking about there?
[00:17:18.29] M: I guess like maybe the--
[00:17:21.19] T: --rocks and stuff a long time ago--
[00:17:21.26] M: --Like when they were just like trying to [inaudible to me]
[00:17:24.07] Dan: Those early those early soundings is talking about the cannonballs, right, like the early a sounding is the dep- so when it says "Today's sound waves bring the hazy images of those early sound waves into sharp focus" they mean that the new methods of sound waves are making the images clearer...

1. What scaffolds do you see used here? Try to mark them at specific lines.

2. How are the scaffolds contingent (or not) on students' audible comprehension?

3. How effective do you think these scaffolds are?

PRACTICE TEACHING 9

- Partner up with someone you didn't know before today.
- Practice teaching with line 45-85
    - **Tutor** should read text aloud, then listen as student paraphrases the text, and scaffold as necessary.
    - Student tries to paraphrase the text and responds to scaffolds. Try to pretend you're a high school student with some success and some struggle.

What scaffolds did you use? Why?

What seems easy about this kind of teaching? What seems hard?

HOMEWORK: VIDEO PRACTICE

- Find someone else willing to work with you and act like a high schooler for 10 minutes.
- Using the humanities passage of "A Four-Hundred-Year-Old Woman" by BHarate Mukherjee, do ten minutes of practice scaffolding, and record it on video.
- Send Dan the video
    - Upload to Box folder
    - Send it to Box via email: Videos_.7c610hsgfd5wdyq2@u.box.com
- Watch your video and identify the scaffolds that you used. We'll take a look at some of the videos on Friday and see whether we agree on them!

DAY 2: DELIVERING THE INTERVENTION WITH FIDELITY

WELCOME AND TEAMBUILDER NAME GAME

Come on in, pick up your handouts, get your donut, and get your index card out!

DISCUSS HW:

Let's watch two videos of your practice sessions, Try to:

- Identify which scaffolds were used
- Assess their contingency—did they respond to student knowledge?
- Assess their effectiveness—did they work well?

Liz Hadley:

Dongxin Li:

REVIEW FROM DAY 1 TRAINING

Scaffolding

Contingency

Paraphrasing

Academic language

Complex texts

NEW TERMS

Transfer of responsibility:

Fading:

Comprehension as

outcome,

as process,

and as sensemaking:

INTERVENTION DESIGN FOR TRANSFER AND FADING

Key features of transfer and fading are being built into the scripts!

Design element	Sessions 1+2	Sessions 3+4	Sessions 5+6	Sessions 7+8
Who does the read aloud?	After learning to paraphrase (students read and teacher paraphrases), **teacher reads text** and students paraphrase	**Students read to tutor** (with teacher taking over as necessary), students paraphrase for tutor	**Students read to tutor**, students paraphrase for tutor	**Students read and paraphrase to each other** –tutor listens and scaffolds as necessary
Big question provided for students	Yes	Yes	No	No
Attention to conversation norms and positive relationships	Lots (esp. in session 1)	Some at beginning of lesson	Quick reminder at beginning of lesson	Quick reminder at beginning of lesson if tutor thinks necessary
Text complexity	Moderate (per ACT)	Complex	Complex	Complex

TEXTS

Lesson	Genre	Text Excerpt From	Study text title	Big question*
1	Natural science	Stephen Jay Gould "Were dinosaurs dumb?"	Dumb dinos	Were dinosaurs dumb?
2	Prose fiction	Paule Marshall *The Fisher King*	Jazz show	How does a jazz pianist unexpectedly deliver an amazing performance?
3	Social science	Deborah Cramer *Great Waters, An Atlantic Passage*	Atlantic seafloor	What landforms are on the bottom of the ocean?
4	Humanities	Bharati Mukherjee "A Four-Hundred-Year-Old Woman"	Indian-American writer	How does this writer deal with transnational cultural identities?
5	Natural science	Oliver Sacks "An Anthropologist on Mars"	Color theory	(How do people see color?)
6	Prose fiction	Cristina Perri Rossi "The Threshold"	Dreaming woman	(How do these two friends talk about their dreams?)
7	Social science	Joseph Ellis *American Sphinx: The Character of Thomas Jefferson*	Jefferson	(Did Thomas Jefferson plagiarize the Declaration of Independence?)

8	Humanities	Jon Pareles "India Resounding in New York"	Musical fusion	(How does Indian and American music mix?)

**Questions in parentheses only provided if tutor deems necessary.

FIDELITY

- **Lesson fidelity**: we use the same texts and same activities and same lesson plans.
    - You will track your own directly on the lesson plans
    - ~30% of sessions will be observed by another tutor, who will track them on a second copy of the lesson plan sheet.
- **Scaffold fidelity:** we can reliably work together to decide which scaffolds were used and what adaptations were made. This is the key measurement of the entire study, so we simply can't mess this up!
    - You will track your own on your copy of the scaffolds sheet
    - ~30% of sessions will be observed by another tutor, who will track them on a second copy of the scaffold sheet.

- PLEASE ASK about whether a certain part of the intervention is malleable or not! I appreciate initiative, but we have to ensure some consistency.

PARTNER PRACTICE TEACHING

1. Using the last half of the Mukherjee passage (i.e. after line 44), practice the instruction and paraphrasing with a partner for about ~10:00.
2. Then, both the teacher and the partner stop and try to record the scaffolds that you used.
3. Compare notes with your partner
4. Then switch, with your partner starting with the beginning of the Dumb dinos passage

Tips for scaffolders:

- Try to find the balance between scaffolding and over-scaffolding
- Try to refine the students' sensemaking by drawing in their participation, extending their ideas, and pushing them for evidence WITHOUT forcing them to read it the way you do.
- Experiment with allowing a student to make a "mistake"
- Encourage maximum participation from students

Notes on partner teaching: What was easy? What was hard?

FORMATIVE ASSESSMENT: CODE THIS SEGMENT

[00:26:44.28] Dan and near half: "Here in this bleak and forbidding place, where the water is almost freezing, subterranean fires have lifted mounds of fresh lava onto the seafloor."
[00:26:33.25] A: "A forbidden place"
[00:26:38.27] D: What's "a forbidding place" mean?
[00:26:39.10] T: Are they referring to [overlapping talk]
[00:26:41.24] Dan: [using high voice to talk to A] I'm a little six-year-old, I don't understand what forbidding means
[00:26:44.24] K: An unknown place
[00:26:46.08] Dan: okay, so it's unknown place, alright
[00:26:46.08] [?] : [abyss]
[00:26:49.29] K: [an unheard of placae]
[00:26:49.29] L: [where the water's freezing]
[00:26:50.00] Dan: an unheard of place, okay, really cold [dramatic voice] subterranean fires have lifted mounds of fresh lava onto the seafloor"
[00:26:58.10] L: Okay, an unknown place, where it's really cold-
[00:26:59.28] Dan: -yeah-
[00:26:59.28] L: -fires have raised fresh lava onto the sea floor
[00:27:07.06] Dan: fires? we talkin about --hold on [turns to look at L] -- you told me something about the ocean what you talkin about subterrain you talkin about fire
[00:27:12.13] L: subterranean fires
[00:27:13.24] Dan: what kind what does subterranean mean?
[00:27:15.09] Al: uhhh
[00:27:16.15] T: isn't that a sea-
[00:27:17.29] [several students' talk overlaps]
[00:27:17.29] Dan: yeah sub sub terra see terra means earth we got two parts to this so under the earth, what would be fire under the earth?
[00:27:29.01] [overlapping talk, subterranean fires]
[00:27:29.01] Dan: yeah, volcanoes and stuff

HOMEWORK: PRACTICE FULL LESSON

Using the lesson plan, text, and scaffolds list, teach a 35-minute practice lesson to a sample student pretending to be a high schooler.

Upload the video of the lesson and your coded scaffolds to Box by 11:59 PM Monday. Or upload video to YouTube and send Dan the link to it. *****

DAY 3: STRIVING FOR CONSENSUS ON CODING THE SCAFFOLDS

WELCOME AND TEAMBUILDER NAME GAME

Come on in, pick up your handouts, make your name tag, get your donut, and get your index card out!

WARM-UP: PRACTICE CODING

[00:26:44.28] Dan and near half of the students read aloud: "Here in this bleak and forbidding place, where the water is almost freezing, subterranean fires have lifted mounds of fresh lava onto the seafloor."
[00:26:33.25] A: "A forbidden place"
[00:26:38.27] D: What's "a forbidding place" mean?
[00:26:39.10] T: Are they referring to [overlapping talk]
[00:26:41.24] Dan: [using high voice to talk to A] I'm a little six-year-old, I don't understand what forbidding means
[00:26:44.24] K: An unknown place
[00:26:46.08] Dan: okay, so it's unknown place, alright
[00:26:46.08] [?] : [abyss]
[00:26:49.29] K: [an unheard of place]
[00:26:49.29] L: [where the water's freezing]
[00:26:50.00] Dan: an unheard of place, okay, really cold [dramatic voice] subterranean fires have lifted mounds of fresh lava onto the seafloor"
[00:26:58.10] L: Okay, an unknown place, where it's really cold-
[00:26:59.28] Dan: -yeah-
[00:26:59.28] L: -fires have raised fresh lava onto the sea floor
[00:27:07.06] Dan: fires? we talkin about --hold on [turns to look at L] -- you told me something about the ocean what you talkin about subterrain you talkin about fire
[00:27:12.13] L: subterranean fires
[00:27:13.24] Dan: what kind what does subterranean mean?
[00:27:15.09] Al: uhhh
[00:27:16.15] T: isn't that a sea-
[00:27:17.29] [several students' talk overlaps]
[00:27:17.29] Dan: yeah sub sub terra see terra means earth we got two parts to this so under the earth, what would be fire under the earth?
[00:27:29.01] [overlapping talk, subterranean fires]
[00:27:29.01] Dan: yeah, volcanoes and stuff

CODING CONTINGENCY

- What does it mean if something is “contingent”? How will we separate contingency?
    - Goes beyond the extremes of the continuum of the support scale—doesn’t do the task for the student, but doesn’t leave them to do it entirely independently.
    - It is *contingent* if it responds to students’ needs in some way:
        - Might be motivational—does it encourage a student to participate even if it does not focus on particular textual elements?
        - Might be informational—does it provides information or highlight certain words for students to help them paraphrase
        - Might be challenging—simply asking students to go beyond their initial approach

- Might be hard to see contingency—letting a student make a "mistake" for engagement reasons
- **Rate contingency 1-3**
  - **1 not clear what it is contingent on, 2 =possibly contingent, 3 completely contingent**

CODING EFFECTIVENESS

- What will we do to rate the effectiveness of a scaffold?
  - Is it taken up by the students—do they respond to it?
    - Wait time will be key—tutors have to give students the opportunity to
  - Does it enable the students to do a more effective paraphrase?
  - Is the student internalizing the suggestion and using it later on his or her own? (This will be really hard to see unless tutor or student articulated is)

LOGISTICS AND SCHEDULING

- **Appearance**
  - Teacher dress: professional in the same dress code as you would use for student teaching.
  - Exceed the expecctations of everyday teachers—not because we need to look fancy but because we need to communicate the purpose of our professionalism to students.
- **Timing and transportation**:
  - meet at Wyatt at 7:45 if you want a ride, meet in the **school office** at 8:10 if you drive yourself.
  - This is for tutors and coders!
- **Emergencies**:
  - contact Dan immediately cell is fastest (512) 774-9394, or email.
  - I will give as much advance notice as possible
- **Schedule**
  - set for cycle 1,
  - almost set for cycle 2-4 (had new teacher contact me yesterday, still trying to hire 3 more tutors from Meharry MC).
  - Hope to have that finalized by the end of the week so you have ~2 weeks notice
- **Video recording**
  - Video recording protocols and data security—forthcoming. Goal will be to take video and immediately transfer it at the end of session to secure backup devices
- **Data entry**
  - Every day you're teaching, you will get a folder of the lesson plans and materials that day when we meet in the HHS office. That will also have your scaffolds list, which you will return immediately to your folder.
- **Protocols for pretesting and posttesting**
  - Will be in folders for those who are distributing those that days. Protocols must be followed!
- **Timesheets**
  - Issues with timesheets? How to report your hours? Questions about pay or HR?

FINAL "EXAM"

- We're going to watch a full 37-minute lesson here so you can see me do it live!
- Use paper form to follow lesson plan and track the scaffolds—you're a coder here.

- Then we will use the laptops and the Google form to submit our results
- Take time-stamped notes about what you think is important—effectiveness, contingency, etc.