2023

Supporting multilingual learners' reading competence: a multiple case study of teachers' instruction and student learning and motivation

Melissa A. Gallagher

Jori S. Beck

Erin M. Ramirez

Ana Taboada Barber

Michelle M. Buehl

Check for updates

Supporting multilingual learners' reading competence: a multiple case study of teachers' instruction and student learning and motivation

Melissa A. Gallagher[1]*[†], Jori S. Beck[2†], Erin M. Ramirez[3], Ana Taboada Barber[4] and Michelle M. Buehl[5]

[1]College of Education, University of Houston, Houston, TX, United States, [2]Department of Teaching and Learning, Old Dominion University, Norfolk, VA, United States, [3]College of Education, California State University Monterey Bay, Seaside, CA, United States, [4]College of Education, University of Maryland, College Park, MD, United States, [5]College of Education and Human Development, George Mason University, Fairfax, VA, United States

Supporting students' reading competence (i.e., their comprehension and vocabulary) is complex, particularly when working with multilingual learners, and involves implementing instructional practices to support their behavioral engagement in reading as well as their reading motivation. The purpose of this mixed methods case study was to examine changes in multilingual learners' reading comprehension, academic vocabulary, reading engagement, and reading motivation after participating in a 7-week intervention called United States History for Engaged Reading (USHER) and then examine qualitative data to explain why these changes may have occurred. We found changes in the reading comprehension of MLs across all four teachers' classes, and variable changes in academic vocabulary, reading engagement, and reading motivation. We highlight specific instructional practices that may have led to these changes, including engaging students in discussions during explicit vocabulary instruction, allowing students choice and the opportunity to collaborate, and making the content relevant by relating it to students' lives, among others.

KEYWORDS

comprehension, English language learners, vocabulary, motivation/engagement, instruction, multilingual learners

Introduction

Teachers' instructional practices can shape students' motivation, engagement, and their reading competence (Guthrie et al., 2012; Guthrie and Klauda, 2015). In the middle grades, reading competence is typically measured by standardized reading comprehension assessments and is predicted by students' motivation (Mucherah and Yoder, 2008; Kim et al., 2017; Taboada Barber et al., 2018), students' comprehension strategy use (Vaughn et al., 2011, 2017), and students' vocabulary knowledge (National Institute of Child Health and Human Development (NICHD), 2000; Lesaux et al., 2010; Proctor et al., 2014). Teachers can support students' motivation, comprehension strategy use, and vocabulary knowledge through specific teaching practices (Lesaux et al., 2010; Vaughn et al., 2011, 2017; Taboada Barber et al., 2018; Crosson

et al., 2021). Teachers of linguistically diverse students must also implement teaching practices that support their language development (Goldenberg, 2020). In order to understand the complexity that is the reality in classrooms, educational researchers must investigate multiple factors that influence reading competence for different groups of students to get a more complete picture of classroom instruction.

The connection between reading instruction, motivation, behavioral engagement, and achievement in reading is well documented (Guthrie et al., 2012; Wigfield et al., 2016), and other researchers have explored academic vocabulary supports for multilingual learners (MLs; Vaughn et al., 2009; Verhoeven et al., 2019); however, we do not know of a study that connects reading comprehension, reading motivation, academic vocabulary, and MLs within the context of literacy instruction in history. In the current study, we attempted to shed light on the complexity of reading instruction through case studies of four teachers' implementation of the USHER intervention. The purpose of the current study was to examine changes in reading comprehension, academic vocabulary, motivation, and engagement by teacher for middle school MLs, and then examine a case study of each teacher's instructional practices to try to explain which instructional practices might have led to changes in each teacher's classes. The following research questions guided this study:

1. For each case, was there a difference in MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, and reading motivation after participating in USHER?
2. How might the instructional practices that each teacher implemented in their classroom explain any changes in their MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, or reading motivation?

Theoretical framework

Given the complexity of reading instruction, we framed this study within the *model of reading engagement processes within classroom contexts* (Guthrie et al., 2012; Figure 1). This model synthesizes research on the relationship between instruction, motivation, behavioral engagement, and reading competence. We define reading competence as reading comprehension and academic vocabulary knowledge.

Previous research indicates that behavioral engagement influences reading competence (Guthrie et al., 1999; Guthrie and Klauda, 2015) and that motivation to read, including self-efficacy, influences behavioral engagement (Cho et al., 2021). There is also evidence that instructional practice has direct effects on reading competence and behavioral engagement (Guthrie et al., 2012; Rosenzweig et al., 2018). However, the relationships between instructional practice, motivation, and reading competence have not been investigated with students from different linguistic backgrounds (Guthrie et al., 2012; Proctor et al., 2014). In particular, MLs may require different or additional instructional practices and conditions to support their growth in reading motivation, behavior, and competence in English. Thus, we framed our case study within the model of reading engagement processes to foster reading comprehension and academic vocabulary but sought to explore a different context than has been explored in the

past: linguistically diverse classrooms. Moreover, using a case study approach allowed us to examine teachers' practices holistically through qualitative analysis of field notes. Examining both teachers' implementation of the practices written into the USHER intervention, as well as the adaptations they made to the intervention, allowed us to contribute instructional practices that may support MLs within this existing model of reading engagement.

Reading comprehension

Reading comprehension has been referred to as the "essence of reading" (Durkin, 1993, p. 4). Durkin defined reading comprehension as "intentional thinking during which meaning is constructed through interactions between text and reader" (pp. 4–5). As students progress through school, the texts they are expected to comprehend increase in complexity (Cho et al., 2021). Moreover, as students get into upper elementary and middle school, they must read and comprehend complex texts not just in their English classes, but also in their content area classes (e.g., social studies and science texts). This creates a need for reading comprehension instruction not just in the elementary grades, but also in middle school and beyond.

## Instructional practices to support reading comprehension

Efforts to improve reading comprehension have usually consisted of direct strategy instruction. Numerous experimental and quasi-experimental studies indicate that cognitive and metacognitive strategy instruction effectively increases student strategy use and text comprehension for both readers in general education settings who struggle (Okkinga et al., 2018; Cho et al., 2021) and for students with learning disabilities (National Institute of Child Health and Human Development (NICHD), 2000; Vaughn et al., 2009; Capin and Vaughn, 2017). Intervention research has also documented the influence of comprehension strategy instruction on the content vocabulary of MLs (Crosson et al., 2021), on researcher-developed comprehension measures (Taboada Barber et al., 2015, 2017; Okkinga et al., 2018), and on standardized comprehension measures (Vaughn et al., 2011, 2017; Okkinga et al., 2018).

Academic vocabulary

Vocabulary is a key component of reading competence and is significantly related to reading comprehension (Lesaux et al., 2010; Crosson et al., 2021). Academic vocabulary includes the general academic and discipline-specific words that are used in academic texts (Green and Lambert, 2018). As students are asked to read more complex academic texts, they also need instruction in academic vocabulary.

Instructional practices to support academic vocabulary

Vocabulary instruction is multifaceted, including explicitly teaching specific words or teaching word-learning strategies (e.g.,

FIGURE 1
Model of reading engagement (Adapted from Guthrie et al., 2012).
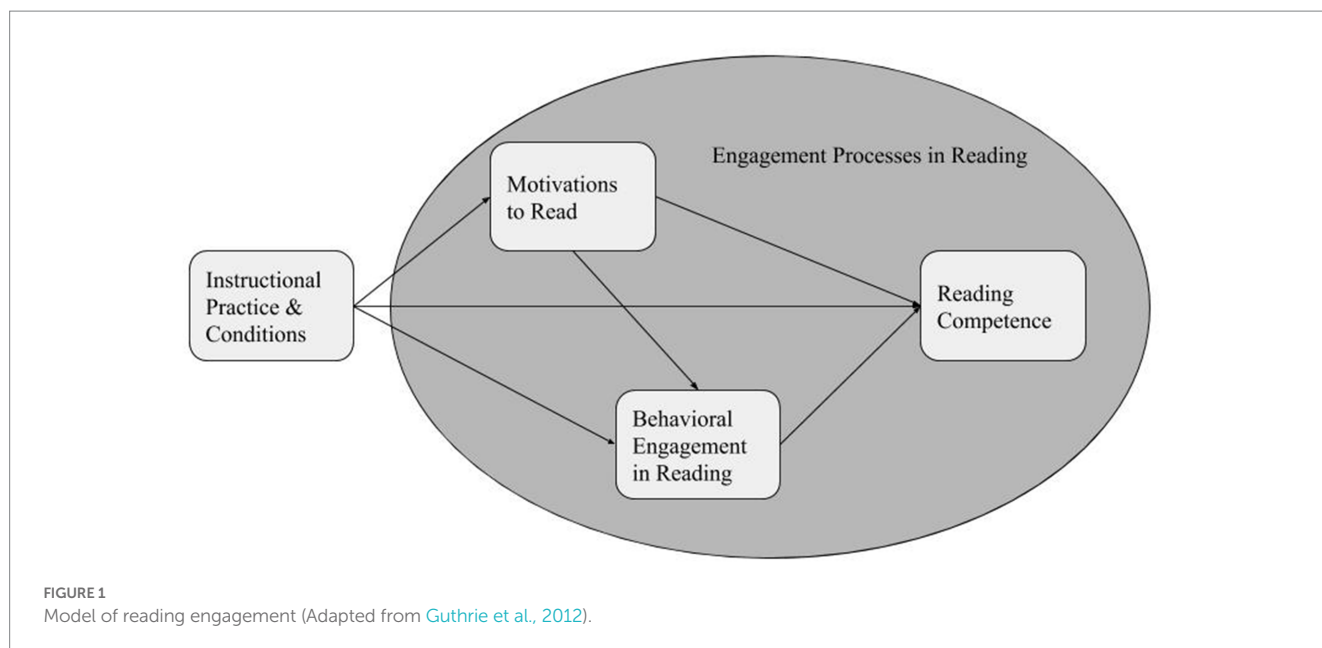
morphology, identifying cognates, using resources, using context clues; Manyak et al., 2021). Many intervention studies have shown that MLs have improved their knowledge of words that are explicitly taught (Carlo et al., 2004; Lesaux et al., 2010; Vaughn et al., 2017; Gallagher et al., 2019; Crosson et al., 2021), particularly when academic vocabulary is appropriately integrated into content learning. Although teachers should teach both specific words and word learning strategies, the benefits for MLs have been mixed. For instance, in our previous research, we found that MLs only improved on words they were explicitly taught and did not improve on words they encountered incidentally, which may have been because of the uneven teaching of word learning strategies (Gallagher et al., 2019). However, Crosson et al. (2019) gave all MLs in their study two interventions: an intervention focusing on explicitly teaching words and an intervention focused on explicitly teaching words and their Latinate roots, although in a different order (i.e., one group received explicit teaching first and Latinate roots after whereas the other received instruction in the reverse order). They found that all MLs improved in the words they were explicitly taught, but that after participating in the Latinate roots intervention, students were more successful in figuring out the meanings of unfamiliar words. These findings suggest that teachers should explicitly teach academic vocabulary. Additionally, there is a need for more research on how MLs take up word learning strategies, although there is evidence that explicitly teaching these strategies may benefit MLs.

Behavioral engagement for reading

Behavioral engagement for reading includes the observable behaviors that students exhibit when they are engaged in reading, such as effort and persistence, and spending sustained amounts of time reading (Guthrie et al., 2012; Guthrie and Klauda, 2015). Students' behavioral engagement in reading predicts their reading comprehension for both native English-speakers (NESs) as well as for

MLs (Barber et al., 2020). Thus, teachers should strive to foster their students' interest in reading.

Instructional practices to support behavioral engagement for reading

Research suggests that student-teacher relationships (Lee, 2012; Quin, 2017), specific classroom tasks (Lutz et al., 2006; Parsons et al., 2015; Taboada Barber et al., 2016), and setting classroom goals (Lutz et al., 2006; Skinner and Pitzer, 2012) are among some of the instructional practices that support student engagement in reading. Lutz and colleagues found that most students were engaged when the teacher provided: "(a) knowledge goals for tasks, (b) availability of multiple texts well matched to content goals, (c) strategy instruction, (d) choices of texts for reading, and (e) collaborative support" (p. 13). Parsons et al. (2015) also found that providing choice and student collaboration yielded higher student behavioral engagement in reading. In addition, they found that tasks that were authentic, challenging, and sustained over time were more engaging for on or above-grade level readers. In previous research, we found that MLs had the highest levels of cognitive and affective engagement during collaborative activities, such as guided or partner reading, as compared to during whole class discussion, teacher modeling, or teacher-led guided reading (Taboada Barber et al., 2016). Taken together, these findings suggest that teachers should provide students with choices and opportunities for collaboration to support their behavioral engagement in reading.

Reading motivation: self-efficacy for reading

Reading motivation refers to "the individual's personal goals, values, and beliefs with regard to the topics, processes, and outcomes of reading" (Guthrie and Wigfield, 2000, p. 405) and is important to

maintaining behavioral engagement in reading, especially when activities are cognitively demanding (Wigfield et al., 2016). Reading motivation is a predictor of reading comprehension for both NESs (Guthrie and Wigfield, 2000; Proctor et al., 2014) as well as MLs (Hwang and Duke, 2020).

In the current study, we focused on one key social cognitive variable central to reading motivation (Guthrie et al., 2004; Schunk and Zimmerman, 2007): self-efficacy for reading. Self-efficacy refers to one's perceived capabilities for learning or performing actions at designated levels in specific situations (Pajares, 1996; Bandura, 1997) and is one of the strongest motivational predictors of future achievement (Usher et al., 2019). Students with a high sense of self-efficacy are motivated and engaged in learning (Schunk and Mullen, 2012), persevere with challenging tasks, search for deeper meaning across learning tasks, report lower anxiety, and have higher academic achievement (Usher et al., 2019). Moreover, correlational research suggests that self-efficacy is a relatively stronger predictor of reading achievement in middle school as compared to other motivations, such as reading for social reasons, grades, or curiosity (Mucherah and Yoder, 2008; Proctor et al., 2014). Reading self-efficacy also impacts use of reading strategies and improved comprehension (Shehzad et al., 2019). Students with higher self-efficacy for reading interpret challenge as a learning opportunity and put in more effort, leading to developing advanced literacy skills required in the secondary grades (Cho et al., 2021). Additionally, students with higher self-efficacy for reading have stronger growth in reading comprehension (Unrau et al., 2018; Fitri et al., 2019), whereas students with low self-efficacy in reading avoid challenging reading tasks leading to fewer opportunities to improve their reading comprehension (Solheim, 2011; Unrau et al., 2018; Cho et al., 2021). Like motivation, self-efficacy is a construct that is domain-specific and, thus, students may feel efficacious for one task or domain (e.g., mathematics) but not for others (e.g., reading; Bandura, 1997).

Instructional practices to support reading motivation

Teacher supports for reading motivation, and specifically self-efficacy, include: (a) competence support, (b) content and strategy instruction relevance, and (c) student collaboration. Teacher supports for student competence include teachers being explicit about the types of comprehension strategies and their purposes (Schunk and Miller,

2002; McLaughlin, 2012), and modeling and parsing out the strategies so that students visualize the steps in the process (Schunk and Zimmerman, 2007; Magnusson et al., 2019). It is important for teachers to provide specific feedback to students regarding whether their comprehension skills need improvement or have improved over past performance (Pintrich and Schunk, 1996; Fitri et al., 2019). Fostering relevance is a way of supporting students' autonomy and motivation because teacher explanations about the relevance of schoolwork for students help them grasp its contribution toward the realization of their personal goals, interests, and values (Magnusson et al., 2019). Another teacher support for self-efficacy is student collaboration, which is based on the premise that students enjoy working together (Proctor et al., 2014) and are likely to become more motivated to do a task when they are able to exchange ideas. Collaborative activities also promote oral language learning when MLs interact with both NESs and other MLs (Ingrid, 2019).

The purpose of this study was to examine changes in MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, and reading motivation after participating in USHER, and then to analyze teachers' instructional practices qualitatively to better understand the complexity of literacy instruction for the MLs in those teachers' classrooms.

Methods

We used a mixed methods multiple case study design (Stake, 2006) to explore our research questions (see Table 1). Specifically, we chose a partially mixed concurrent dominant status design (Leech and Onwuegbuzie, 2009) in which the quantitative data carried more weight. Both the quantitative data and qualitative data were collected concurrently during the intervention including before and after the intervention as described below. The quantitative data and qualitative data were analyzed separately and then synthesized for reporting to draw inferences. The quantitative findings are reported first within each case followed by qualitative findings to provide nuance to the quantitative results.

Setting

This study was conducted in Grade 6 classrooms in a school in a suburban district. Within this school, history and English Language

TABLE 1 Alignment of research questions, data collection, and analysis.

Research question	Relevant methods of data collection	Methods of data analysis
1. For each case, was there a difference in MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, and reading motivation after participating in USHER?	• History Reading Comprehension Measure • Academic Vocabulary Measure • Reading Engagement Index • Reading Self-Efficacy Beliefs Scale	• Paired samples *t*-tests (Maria's, Jerome's, and Barbara's classes) • Wilcoxon signed rank tests (Carol's class) • Benjamini-Hochberg correction
2. How might the instructional practices that each teacher implemented in their classroom explain any changes in their MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, or reading motivation?	• Pre- and post-implementation interviews • Weekly teacher debriefs • Observation protocol	• Four rounds of data reduction • Memoing • Case study matrices

TABLE 2 Case demographics.

	Maria	Jerome	Barbara	Carol
Gender	Female	Male	Female	Female
Race/ethnicity	Latina	African American	White	White
Years of teaching experience	28	4	4	6
Number of classes	2	2	2	1 (self-contained special education)
Number of students*	31 MLs	29 MLs	26 MLs	8 MLs
	14 NESs	16 NESs	22 NESs	3 NESs

*Native English-speaking (NES) students are reported here to give a sense of the classroom makeup, but only MLs are included in the data analysis.

TABLE 3 Instructional practices included in the USHER instructional materials.

Instructional practices to support reading comprehension	Instructional practices to support reading motivation	Instructional practices to support MLs' reading competence
• Activating background knowledge through the use of text features • Generating text-based questions • Organizing information graphically • Identifying main idea and supporting details • Monitoring comprehension while reading	• Competence for reading • Relevance • Student collaboration	• Plan for purposeful grouping • Elaborate on students' responses • Repeat important information • Provide prompts for student discussion • Use word maps to explicitly teach vocabulary • Post new vocabulary on word walls

Arts (ELA) are taught by one teacher throughout the year in which history class is held every day for 45 min and ELA for 90 min. The fourth author negotiated access to the school site through the superintendent as part of a multi-year federal grant. The research team worked with school site principals to coordinate the interventions each year, supported by the endorsement of the superintendent. By the third year of the project, the team had developed robust relationships with the four teachers who participated. Many of the graduate research assistants (GRAs) had worked on the project for 2–3 years and had been immersed at the site along with the principal investigators.

Teacher participants

The cases in this study included four Grade 6 history and ELA teachers (Maria, Jerome, Carol, and Barbara[1]) who were chosen because they had taught different units in the USHER intervention using these same comprehension, academic vocabulary, engagement, and self-efficacy practices during the previous 2 years (see Table 2 for demographic information on the case teachers and classes). Thus, they had knowledge of and familiarity with the practices and were encouraged to make adaptations to the lessons as they saw fit. Furthermore, the cases were bound by the teachers during the 7 weeks of the USHER intervention; thus, whether each teacher taught one or two classes of students, data from all classes were used to understand the teachers' implementation of USHER. Each teacher was paired with a GRA who was thoroughly familiar with the development and implementation of USHER design elements.

Of the MLs ($n = 94$) in these teachers' classes, 90% were native Spanish-speakers and 47% ($n = 44$) identified as female. With regard to race or ethnicity, 90% ($n = 85$) identified as Hispanic, 5% ($n = 5$)

identified as Asian, and 1% ($n = 1$) as White. They ranged in age from 11 years, 2 months old to 13 years, 5 months old, with a median age of 11 years, 9 months old. The school district shared the World-Class Instructional Design and Assessment (WIDA; Board of Regents of the University of Wisconsin System on behalf of the WIDA Consortium, 2014) levels of English proficiency for 70 of the MLs. These data indicated that one student was at a level 1 Entering, 18 were at a level 3 Developing, 23 were at a level 4 Expanding, five were at a level 5 Bridging, and 23 were at a level 6 Reaching.

USHER implementation

Data for this multiple case study came from Grade 6 MLs and their teachers in year 3—the final year of the intervention. In year 3, explicit instructional practices to support the literacy development of MLs were added to the instructional materials, including a guide in the handbook for teachers outlining practices to support MLs in literacy, as well as specific instructional practices to support MLs included in the lessons.

USHER year 3 implementation lasted 7 weeks and included five reading comprehension strategies, three supports for reading motivation, and six practices to support MLs (see Table 3). We designed the lessons based on the state history standards for four history units and grounded them in research-based reading comprehension, motivation, and ML instructional practices. In addition, USHER included the use of trade books to allow for deep and extensive reading in history and the provision of vocabulary supports for MLs through text features such as glossaries, bolded terms, and conceptual illustrations. For each unit, teachers were provided with class sets of two to ten different trade books of varying reading levels. One on-grade text was designated as the whole class book. The other books were provided to support differentiated reading instruction targeting the same history content. Lesson plans included small group and partner reading activities and teachers were

---

1 The name of the intervention and all four teachers are pseudonyms.

encouraged to choose appropriate-level texts for students. All of the books provided were relevant to the lessons but may or may not have been used depending on the reading levels of the students.

Lesson plans were provided to teachers and they were encouraged to modify the lesson plans according to their knowledge of their students. Each of the 35 lessons included an outline of activities for each day (i.e., warm-up, whole class, small group/partner reading, and closure activities). Reading comprehension strategies were taught via the gradual release of responsibility model (Pearson and Gallagher, 1983) following modeling, guided practice, and then either small group reading (i.e., groups of three to four students) or partner reading. The three teacher motivation supports were detailed in the lessons via examples of possible teacher actions/statements or embedded in activities so that teachers were prompted to use particular supports at various points. Academic vocabulary words were preselected from texts and included as part of the focus through explicit vocabulary instruction. We held one full day and one half-day professional development sessions before implementation that included time for teachers to learn about supporting MLs in literacy instruction, review materials, and provide feedback on the lessons.

Data collection

We used pre- and post- quantitative data to evaluate whether or not students improved in reading comprehension, academic vocabulary, reading engagement, and reading motivation, and we used qualitative data to document teachers' implementation of instructional practices in an effort to explain our quantitative findings. Salomon (1991) argued that this distinction transcends independent quantitative and qualitative paradigms, allowing both types of data to be integrated to interpret the possible influence of the intervention.

History reading comprehension

Given the focus of USHER on reading in history, we created a measure of history reading comprehension where all the texts were non-fiction, historical passages. To assess students' history reading comprehension, students read three passages excerpted from trade books and responded to eight multiple-choice questions per passage targeting vocabulary, text-based (literal) understanding, and local and global inferencing. All four types of items were included twice for each passage. The passages pertained to topics not taught during the intervention and ranged in length from 512 to 616 words. One passage was on grade level, one was below grade level, and one was above grade level according to the Flesch–Kincaid reading index (Kincaid et al., 1975). Students' responses were scored as 1 for a correct response and 0 for an incorrect response, and a total history comprehension score was created (i.e., maximum score of 24). Two alternative forms were created. Students' history comprehension scores were correlated with scores on the Gates-MacGinitie Reading Comprehension Assessment (MacGinitie et al., 2000) given at the same time (i.e., pre-implementation: $r = 0.68$; post-implementation: $r = 0.74$), and the data were reliable (i.e., pre-implementation: $\alpha = 0.85$ [Form A]; post-implementation: $\alpha = 0.85$ [Form B]). There is evidence that the two forms are reliable, as they are statistically significantly correlated, $r = 0.76$, $p < 001$.

Academic vocabulary

To assess students' academic vocabulary, we designed a 24-item measure of academic vocabulary words that were either explicitly taught or incidentally encountered in the trade books as part of the intervention instruction (Gallagher et al., 2019). This measure was adapted from a task created by Schoonen and Verhallen (1998) designed to measure students' deep word knowledge. Students were asked to circle the two words that *always went* with the vocabulary word to assess their deep understanding of the meaning of the word. Responses were scored as 1 if both of the correct words were circled and a 0 for any other response. A total academic vocabulary score was created (i.e., maximum score of 24). Students' performance on the academic vocabulary measure was correlated with their performance on the Peabody Picture Vocabulary Test given at the same time (i.e., pre-implementation: $r = 0.59$, $p < 0.001$; post implementation: $r = 0.66$, $p < 0.001$) and the data from the academic vocabulary test were reliable (pre-implementation: $\alpha = 0.60$; post implementation: $\alpha = 0.68$). There was just one version of this assessment, as it was intended to capture change in students' knowledge of specific words: those explicitly taught and those incidentally encountered throughout the intervention.

Behavioral engagement in reading

In order to evaluate students' behavioral engagement in reading, teachers completed the Reading Engagement Index (REI; Guthrie et al., 2004) for each student at pre and at post. The REI includes eight items that measure the outward manifestation of students' engagement while reading. For instance, "This student often reads independently" or "This student works hard in reading." Items were scored on a 5-point scale from 1 (*not true*) to 5 (*very true*). The REI exhibited reliability, Cronbach's $\alpha = 0.82$ at pre and at post.

Reading self-efficacy beliefs

To assess students' reading self-efficacy beliefs, we adapted a measure from Shell et al. (1995) to assess students' perceptions of their ability to read in the domain of history. Students responded to 19 items related to different reading-related tasks (e.g., "Learn about history by reading books"; "Find supporting details for a main idea on a page in a history book") on a scale from 0 (*cannot do at all*) to 100 (*completely certain I can do*). This measure also demonstrated reliability (pre-implementation: $\alpha = 0.92$; post-implementation: $\alpha = 0.92$).

Teacher pre- and post-implementation interviews

Interview questions targeted teachers' implementation of the intervention, adaptations to the lessons, and evolving understanding of student motivation and reading comprehension. Interviews lasted from 12 to 19 min at pre-implementation, and from 21 to 33 min at post-implementation and were audio recorded and transcribed verbatim.

Teacher debriefs

Weekly teacher debriefs were designed to allow the researchers to follow up on teacher adaptations to the curriculum. We created interview guides for these debriefs to ensure some similarity in the data across cases, but researchers were encouraged to ask follow-up questions. Thus, these interviews were semi-structured (Merriam,

2009) and provided latitude to GRAs to encourage teachers to expand on their adaptations to lessons. Example interview questions included, "I noticed that _____ was a change to the USHER lesson. What prompted you to make this change?" and, "What USHER-related practices do you feel you need more help with?" (see Supplementary materials for the full debrief guide). All debriefs were audio recorded and transcribed verbatim. Twenty-nine debriefs were conducted throughout the implementation: seven each with Barbara, Maria, and Jerome, and eight with Carol.

Observations

An observation protocol was designed to identify whether the instructional practices included as part of USHER were enacted in each classroom (rather than to measure fidelity of implementation) and, if not, how they were adapted. The protocol mirrors the format for the USHER lessons: there are four segments for warm-up, whole class, small group reading, and closure. USHER lessons did not necessarily have all four segments in a given lesson; therefore, the observer had to use the USHER lesson to identify which segments to code. Within each of the lesson segments, quantitative and qualitative data were collected. Researchers captured whether the purpose for lesson content and purpose for reading strategy were explained by teachers, as well as the teachers' use of motivation practices. The student participation portion of the protocol was designed to capture how many students were completing the task at hand as well as contributing to small group, partner reading or whole class discussion. Researchers also captured how much of each lesson teachers implemented (e.g., 0 steps, 1–49% of steps, 50–99% of steps, or 100% of steps) and noted what adaptations to the lesson they made. The instructional practices described earlier were included throughout the protocol. Four GRAs conducted 107 observations across all four teachers' classes. Specifically, they conducted 30 observations of Barbara, 28 of Maria, 33 of Jerome, and 16 of Carol, who taught only one class.

Data analysis

We first analyzed the quantitative data to see if there were changes from pre- to post-intervention in MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, and reading motivation for each teacher. We then examined the qualitative data to try to explain why those changes might have occurred for each case. Lastly, we examined similarities and differences across the cases.

Analysis of changes from pre- to post-intervention

To determine if there were changes in MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, and reading motivation we conducted paired samples t-tests for three teachers (Maria, Jerome, and Barbara) and Wilcoxon signed rank tests for Carol because of the small sample size ($n = 8$). Because we conducted so many analyses (12 t-tests and 4 Wilcoxon signed rank tests), we ran an increased chance of Type 1 error, so we used a Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) to determine significance. We did not use an analysis of covariance (ANCOVA) because these data violated the assumption of heterogeneity of regression slopes.

Analysis of instructional practices and changes to the intervention

To understand which USHER practices were implemented and/or changed by the teachers, we analyzed the data that were collected *during* the seven-week implementation (i.e., teacher interviews, observation protocol and field notes data, and teacher debriefs). We began with approximately 400 pages of raw data (interview transcripts and field notes), not including data tables from the observation protocol, and engaged in four rounds of data reduction (see Figure 2): (1) data cycle memos, (2) case study matrices, (3) classroom characterization memos, and (4) summaries. Qualitative data analysis was an iterative process of coding, analyst triangulation, and memo writing. We describe this process in detail below.

Data cycle memos

First, we organized the interviews and observations into data cycles for each teacher, where one data cycle began and ended with an interview or a debrief. There were eight data cycles each for Barbara, Maria, and Jerome and nine for Carol. Data within a cycle were organized around three main categories: implemented practices in line with USHER lessons, changed or adapted practices deviating from the USHER lessons, and teacher rationales for changed or adapted practices. Within a cycle, each GRA first examined the observation protocol to determine lesson elements that were implemented as prescribed; second, they reviewed instances of reading comprehension, academic vocabulary, and motivation practices that the teacher adapted or were different from those specified in the lessons; and third, they analyzed the open observation notes. Next, each researcher transcribed the weekly teacher debriefs to better ensure accuracy of the data and did a line-by-line coding (Corbin and Strauss, 1990) using HyperRESEARCH (ResearchWare Inc., 2013). These data were used to write memos.

Memoing, the process of writing notes to oneself regarding coding (Saldaña, 2009), helped the researchers move beyond the raw data and to delve deeper into themes that were emerging from the data and led to secondary coding. Two faculty members on the research team read and commented on all of the GRAs' memos from the first two data cycles. After revisions were applied based on this feedback, GRAs were paired up to provide ongoing feedback to one another on memos from subsequent data cycles. All researchers met regularly to discuss teachers' implementation of the intervention, and shared comments on data cycle memos.

Case study matrices

We kept track of the findings from the data cycle memos by using an adapted version of Stake (2006) cross-case matrix. We created a matrix for each teacher, noting adaptations to comprehension instruction, ML practices, and motivation practices, as well as rationales for these adaptations and unanticipated effects in the first row of the matrix. In the first column of the matrices, we noted each interview and observation in chronological order. We then made an X in the necessary cells to denote which adaptations were observed in each observation and which rationales were mentioned in which interviews. These matrices allowed us to examine patterns in teachers' adaptations throughout the course of the USHER intervention.

Classroom characterization memos

In the third round of data analysis, the team began to highlight the most important facets of teachers' instruction in three areas:

FIGURE 2
Process of data analysis and reduction.
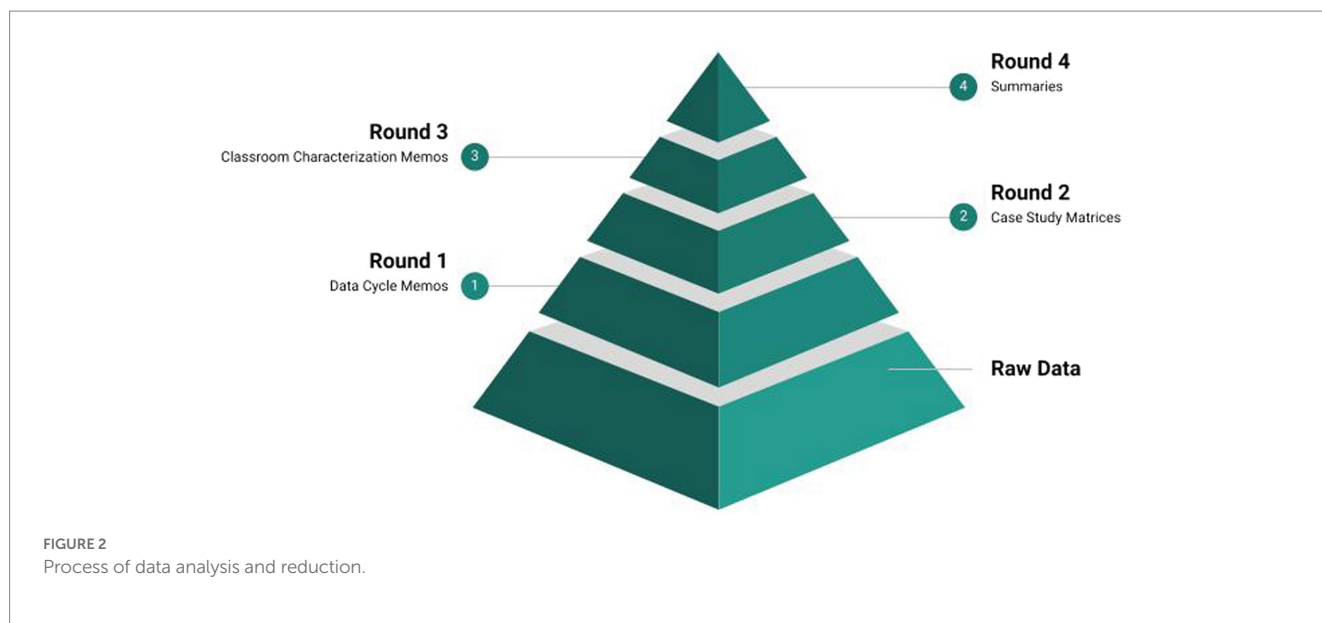
comprehension instruction, motivation practices, and vocabulary instruction. We wrote a memo for each teacher for each of the three areas mentioned above summarizing: (1) strengths, (2) challenges, (3) changes to the intervention, and (4) rationales for changes to the intervention across all data cycles.

Summaries

The classroom characterization memos were then subjected to a final round of data reduction in which the second author wrote up the most significant findings, as demonstrated by patterns in the qualitative data, within each case as well as across the four teachers. The first author read through this final analysis and ensured its accuracy.

Validity and reliability

Two methods were used to ensure the reliability of the quantitative data: standardized administration procedures and interrater reliability (IRR). All instruments were administered by trained GRAs who read from a script during administration. Two GRAs were positioned in each classroom so that one GRA could conduct the administration while the second GRA could circulate to answer clarifying questions. All measures were read aloud to ensure clarity. For purposes of interrater reliability (IRR) with the observation protocol, four GRAs conducted the observations. They received training by first watching and coding videos from previous years of USHER implementation and comparing their codes to the master codes for those videos. Where coding was discrepant, they engaged in discussion to better understand the master codes. Because the protocol was going to be used in live observations rather than video, they also engaged in multiple rounds of observations in classrooms. The GRAs went in pairs to visit classrooms, observed one lesson together, and coded that lesson independently. Immediately after the observation, they met to compare codes and resolve discrepancies. They continued observing in pairs until each pair coded at least 70% of the protocol items exactly the same. IRR continued during implementation to account for drift

and each GRA participated in IRR at least twice during implementation or until an acceptable level of IRR was reached. Reliability information (including Cronbach's alpha and correlations with existing measures) for the other quantitative measures is discussed in the "data collection" section above.

Two primary validity threats in qualitative research include researcher bias and participant reactivity (Maxwell, 2013). With regard to researcher bias, each GRA reflected on her own researcher subjectivity—an element of research design that is omnipresent but sometimes overlooked. Through the process of data analysis, each GRA became aware of the experiences, goals, and subjectivities that she brought to the study. These subjectivities were included in the memos, were discussed as the GRAs reviewed each memo in pairs and reported here when relevant. With regard to participant reactivity, we engaged in rich data collection (e.g., verbatim transcription of interviews, transcribing as much as possible during observations) as well as triangulation (e.g., collecting multiple types of data on multiple occasions) to mitigate this validity threat. Because we had also been working at this research site for 3 years, we had established productive and transparent relationships with the participating teachers. Finally, we use numbers in this manuscript (see "findings" section below) to convey how often teachers implemented particular practices to support our claims about the strength of their implementation and assess how much evidence we had to support these claims.

Findings

Below we present our findings by case, first examining changes in MLs' reading comprehension, academic vocabulary, behavioral engagement in reading, and reading motivation, then strengths and struggles in each teacher's instructional practices. Within each case, we provide percentages and claims regarding the relative strength of aspects of teachers' implementation based on our observation coding (Maxwell, 2013). The percentages indicate how often practices or strategies were used compared to how often they were indicated in the lessons (e.g., 50% of the time) that we were able to observe. Finally,

we summarize how these instructional practices may explain the changes in MLs' achievement and motivation in each of these teachers' classes.

Maria

Students' achievement and motivation

The paired-samples *t*-tests indicated that Maria's MLs had statistically significant improvements from pre- to post- on reading comprehension, $t(29) = 3.31$, $p < 0.01$ with a large effect size, $d = 0.60$ (Cohen, 1988). There were no other statistically significant differences for Maria's students, although the gains in their academic vocabulary and reading self-efficacy were approaching statistical significance (see Table 4).

Instructional practices

Maria was strong in implementing ML practices during the warm-up (i.e., repeating important information [92.3%] and elaborating on student responses [46.2%]) and relatively strong at implementing these practices during comprehension instruction (i.e., repeating important information [38%] and elaborating on student responses [26%])—whether they were indicated in the lessons or not. Maria often noted that she made changes based on what her students needed. She sometimes communicated with students in Spanish to effectively meet the needs of her MLs. We observed that Maria was skilled in building students' self-efficacy for reading which was illustrated when she made students feel like experts during vocabulary instruction. This was particularly evident in how she drew connections between the text and students' background knowledge which she implemented between 28 and 83.3% of the time across observed lesson segments. She also used relevance as a motivation practice effectively to tap into students' background knowledge—particularly during the warm-up segments of the lessons (44% of observations).

Maria struggled to use the gradual release of responsibility model to hand over responsibility to her students, "Honest [*sic*], the releasing of responsibility model, because I do not trust them yet." As a result, she adapted the lessons to be teacher-centered rather than student-centered. For example, in one observation Maria changed a think-pair-share activity to be a teacher-led discussion and, in another observation when students were supposed to work together to find the meaning of words in context, Maria instead told students the meanings. Other teacher-centered changes Maria made were to replace student volunteers for reading during guided practice with a teacher read aloud, telling students the answers to word maps instead of letting them fill them out themselves, and reading aloud the reading quizzes which were meant to be done independently. Moreover, Maria admitted that she struggled with implementing collaborative activities (e.g., small group reading), but analysis of her instruction across the USHER intervention showed that she gradually began to implement collaborative activities more as the intervention progressed which showed her growth in motivation practices. Maria also struggled to adhere to lesson pacing (she went over time in her warm-ups in 68% of observations and ran short on guided reading and closure during 57.1 and 58.8% of observations respectively), to use specific praise for strategy use (only 2 to 4% of observations), and to tie history vocabulary instruction to essential history content.

Given that Maria's students improved on reading comprehension, it is possible that her use of ML practices, including using Spanish to explain concepts, as well as using lower-level texts may have supported her students' learning of reading comprehension strategies. In spite of the adaptations she made to read aloud quizzes or passages to students and not gradually releasing the cognitive work of reading and comprehending to the students, her MLs nevertheless improved in their reading comprehension.

Jerome

Student achievement and motivation

The paired samples *t*-tests indicated that Jerome's MLs' had statistically significant improvements in reading comprehension, $t(27) = 3.98$, $p < 0.001$, with a large effect size, $d = 0.75$, and reading self-efficacy beliefs, $t(26) = 2.81$, $p < 0.01$, with a large effect size, $d = 0.54$. There were no statistically significant gains for his MLs on academic vocabulary or behavioral engagement in reading.

Instructional practices

Jerome, the only male teacher and a three-year veteran of the USHER intervention, had different strengths and weaknesses as compared to the other three teachers. He excelled in his implementation of activating background knowledge using text features and finding main idea and supporting details, and he believed strongly that these comprehension strategies helped students to increase their reading comprehension. The changes that Jerome made to USHER were based on his beliefs about what would help students grow as readers—in particular, he often made lessons more student-centered. One example occurred when Jerome modeled comprehension monitoring through the use of the mnemonic GAUGE (Nuland et al., 2014) by asking students to tell him how he used the A in GAUGE (i.e., ask questions): "What questions did I ask?" and "What was the answer to my question?" When asked about involving students during this think-aloud, Jerome explained his rationale for this change to the intervention:

> One to see if they were paying attention; two, because the poster was on the board. I wanted them to be able to see me modeling the procedure and be able to point out the specific things they were noticing that were on the poster that I was doing … I feel like it was more effective for them to actually know what I was doing because they got the chance to see and hear me modeling the procedure and then they were able to point out the specific things I was doing, rather than me just telling them.

Jerome's strongest motivation and ML practices were making connections (27.1 to 84.4% of the time across observed lesson segments), providing students with opportunities to collaborate (25 to 98.4% of the time across lesson observed segments), providing students with choice (6.3 to 69.3% of the time across observed lesson segments), establishing relevance for instruction (53.3 to 81.3% of the time across observed lesson segments), elaborating on student responses (36.7 to 75% of the time across observed lesson segments), and repeating important information (60.4 to 90.6% of the time across observed lesson segments).

TABLE 4 Means and standard deviations for reading comprehension and academic vocabulary.

Teacher	History reading comprehension				Academic vocabulary				Reading self-efficacy beliefs				Behavioral engagement in reading			
	Pre M (SD)	Post M (SD)	Test statistic	Effect size	Pre M (SD)	Post M (SD)	Test statistic	Effect size	Pre M (SD)	Post M (SD)	Test statistic	Effect size[b]	Pre M (SD)	Post M (SD)	Test statistic	Effect size
Maria	8.10 (4.71)	10.33 (5.73)	t(29) = 3.31[a]	0.60	9.18 (4.25)	10.43 (3.96)	t(27) = 1.97		61.33 (15.90)	67.35 (22.92)	t(27) = 1.99		3.11 (0.87)	3.26 (0.78)	t(29) = 1.27	
Jerome	7.57 (4.67)	10.68 (4.36)	t(27) = 3.98[a]	0.75	10.14 (3.06)	11.11 (3.33)	t(27) = 1.55		65.59 (20.43)	74.46 (15.11)	t(26) = 2.81[a]	0.54	3.24 (0.68)	3.24 (0.71)	t(27) = 0.04	
Barbara	11.00 (5.91)	13.65 (4.93)	t(25) = 3.58[a]	0.70	10.65 (3.30)	13.27 (4.15)	t(25) = 4.15[a]	0.81	67.86 (14.06)	75.06 (12.10)	t(25) = 3.18[a]	0.62	3.88 (0.79)	3.51 (0.84)	t(23) = 4.11[a]	0.84
Carol	5.00 (2.14)	7.25 (2.96)	Z = 2.38[a]	0.70	8.38 (2.87)	9.50 (2.33)	Z = 0.85		39.88 (16.26)	53.10 (9.18)	Z = 2.10		3.00 (0.18)	3.63 (0.42)	Z = 2.52[a]	0.72

[a]Indicates statistical significance after the Benjamini-Hochberg correction was applied.

[b]Cohen's d was used to determine the effect sizes for Maria, Jerome, and Barbara; $r_{equivalent}$ was used to determine effect sizes for Carol (Rosenthal and Rubin, 2003).

Jerome struggled to connect literacy instruction to history content, which was particularly evident during vocabulary instruction; specifically, he had challenges linking new vocabulary to the theme questions defining a weekly unit. Jerome also struggled with student-led, text-based questioning as a comprehension strategy and setting a clear purpose for either strategy use or content to be learned. It was difficult for Jerome to maintain the indicated pacing during vocabulary instruction. For instance, although he taught all the word maps as directed in the lessons, he did not spend much time discussing the words with the students, but rather told them what to write in the word maps and moved on. Similar to the other teachers, he found it difficult to provide specific feedback in his lessons (implemented only 0 to 23.8% during observations).

Jerome's stated belief in the reading comprehension strategies seemed to be related to the effort he put into making them relevant to students and to both modeling these strategies and creating space for students to collaborate and practice the comprehension strategies, particularly using text features to activate background knowledge, monitoring comprehension, and identifying main ideas and details. This emphasis in his instruction may be related to his students' improvement from pre- to post-implementation in both reading and history comprehension. Furthermore, when teaching vocabulary, Jerome made connections to students' lives and helped them see the relevance in the tasks they were doing. Perhaps because of his difficulty balancing the pacing of vocabulary instruction, his students did not make statistically significant improvements. His students' self-efficacy for reading also improved despite his difficulty with providing specific feedback; this may be because of his strengths in making instruction relevant and fostering collaboration.

Barbara

Student achievement and motivation

The paired samples t-tests indicated statistically significant improvement for Barbara's MLs in reading comprehension, $t(25) = 3.58$, $p < 0.01$, and a large effect size, $d = 0.70$ (Cohen, 1988), academic vocabulary, $t(25) = 4.15$, $p < 0.001$, and a large effect size, $d = 0.81$, and self-efficacy beliefs, $t(25) = 3.18$, $p < 0.01$, and a large effect size, $d = 0.64$. However, her MLs' behavioral engagement in reading statistically significantly decreased from pre- to post-, $t(23) = 4.11$, $p < 0.001$ and a large effect size, $d = 0.84$ (Cohen, 1988).

Instructional practices

Barbara's strengths were her use of the gradual release of responsibility model (she implemented more than 50% of the steps indicated in the lesson 68% of the time she was observed), use of supports for MLs' vocabulary development, and establishment of relevance through drawing connections between content and students' lives (implemented 15.8 to 69.2% of the time during observations across lesson segments). Within the gradual release of responsibility model, Barbara was particularly strong in her modeling of strategies such as finding the main idea and supporting details and comprehension monitoring. During her instruction, she would often revisit previously taught comprehension strategies. She also implemented supports for MLs consistently during comprehension instruction including repeating important information (implemented 63.6 to 92.3% of the time observed across lesson segments),

emphasizing key vocabulary words, and elaborating or expanding on students' responses (implemented 20 to 88.5% of the time observed across lesson segments).

Barbara was reluctant to implement small group reading; however, she did implement partner reading and even adapted word maps from a whole class setting to partner work to allow students to "bounce ideas" off one another. Barbara's rationale for adapting small group reading activities to partner reading was that students could not work together effectively in groups larger than two students. During her pre-implementation interview, nearly every weekly debrief, and her post-implementation interview, she expressed reservations about implementing small group reading. Specifically, she was concerned with managing and keeping her students engaged during small group reading. She also expressed a concern that if she did not manage the groups, "you could have a dominant person and you could have someone that's just kinda quiet, off to themselves, not really participating, and not getting what they can out of it." Instead, Barbara seemed to favor partner reading which she expressed in at least two debriefs.

Barbara had difficulty adhering to the pacing of the USHER lessons, often rushing through the closure (which ran short 31.8% of the time during observations), much like Maria and Jerome. Barbara also struggled with allowing students choice (implemented only 10.6 to 16.7% of the time observed across lesson segments). Additionally, she struggled to provide students with specific feedback on strategy use (implemented 0 to 11.5% of the time observed across lesson segments). When asked about her use of specific feedback in a debrief, Barbara commented:

> I do tend to say you know, "Excellent or good job." But it is literally if they looked up a main idea and they had a main idea. "Good job, you are right there" or "You're right where you need to be," so that they know it is tied to that particular strategy that they were working on that they did well with.

Her confusion about providing specific feedback was evident in this explanation. "You're right where you need to be" was not specific feedback because it did not name what the student did accurately. However, Barbara seemed to think it was specific because she assumed students would know it was tied to the strategy they used.

Besides lessening the emphasis on small group work, Barbara also adapted USHER by supplementing some of the non-fiction texts provided as part of USHER with fiction materials. Some of the texts she added were nicely aligned with the non-fiction texts, whereas others were not and were significantly below grade level.

Barbara's students improved on reading comprehension, academic vocabulary, and reading self-efficacy in spite of her reticence to use small group reading. It is possible that her strong modeling of comprehension strategies through think alouds and her scaffolding of student use of the strategies through the gradual release of responsibility model contributed to her students' success in reading comprehension. Her addition of historical fiction texts may have also supported students' comprehension. Moreover, her adaptation of word maps to be more student-centered and her emphasis on key vocabulary may have contributed to her students' improvement in their academic vocabulary. Her use of relevance and collaboration, albeit only in partners, may have led to students' improvement in their reading self-efficacy. Her students' behavioral engagement for reading

may have decreased because she did not allow them to make many choices.

Carol

Student achievement and motivation

For Carol's class, we conducted Wilcoxon signed rank tests to evaluate changes in her MLs' reading comprehension, academic vocabulary, self-efficacy, and behavioral engagement. We found that her students statistically significantly improved in reading comprehension, $Z = 2.38$, $p < 0.05$, and a large effect size, $r = 0.70$ (Rosenthal and Rubin, 2003), and in behavioral engagement, $Z = 2.52$, $p < 0.05$, and a large effect size, $r = 0.72$. Her MLs did not make statistically significant improvements in their academic vocabulary or reading self-efficacy (although self-efficacy was approaching significance and would have been significant without the Benjamini-Hochberg correction).

Instructional practices

Carol, the special education teacher with a self-contained class, had particular strengths and weaknesses that cut across her comprehension instruction, use of motivation practices, and vocabulary instruction. Her comprehension instruction was the strongest of these three areas, and she often emphasized this to the neglect of other types of instruction—for example, she often cut vocabulary activities from her lessons because small group reading ran over time (86% of observed instances). Additionally, she would take the time to reteach strategies such as identifying the main idea and supporting details if she thought her students would benefit from this emphasis. Perhaps because many of her students were also identified as MLs, Carol excelled at implementing ML practices such as repeating important information (94 to 100% of the time across observed lesson segments), elaborating on student responses (56 to 93% of the time across observed lesson segments), and providing prompts for student discussion (63 to 100% of the time across observed lesson segments). She was also very consistent in providing students with feedback on their use of comprehension strategies. Finally, Carol demonstrated growth in gradually releasing responsibility to her students to collaborate in small groups. She had been reluctant to do this at the beginning of implementation and indicated in a debrief, "If I want more *accurate* information, I know that I need to read it aloud" (original emphasis). By the end of implementation, Carol provided little scaffolding for students during small group reading and allowed them more autonomy during this activity. In her final interview she reflected, "I wasn't afraid this time to put them in small group reading groups. I wasn't afraid to say, 'You need to try this. If you need help reading it, we will help reading it.'"

Although Carol struggled at first to implement guided reading, she did allow students to work collaboratively nearly every time it was indicated in the lessons. Her students professed to enjoy working in small groups as well, "[Student A] was like, 'Are we getting in groups again?" And I'm like, 'Yeah, why?' He was just like, 'Oh, it was fun yesterday.'" Thus, her students seemed to enjoy the collaborative activities that USHER provided and Carol implemented.

The pacing of lessons was difficult for Carol and she often omitted closure or shortened it (69% of the time observed) while

whole class or small group reading portions of lessons ran over the allotted time (50 and 86% of the time observed respectively). Although Carol did a good job of tying content to students' lives to make lessons more engaging, she struggled for half of the implementation to integrate literacy instruction with history content—including during vocabulary instruction. Carol made changes to her lessons based on her perceptions of what her students needed with regard to their different abilities; for example, she asked them to draw pictures or diagrams of abstract vocabulary words (e.g., civilization) during vocabulary instruction, which seemed to work especially well for her MLs.

It is important to note that Carol only had 8 MLs and, thus, finding any statistically significant differences was unlikely. Given that her students statistically significantly improved in their reading comprehension and reading engagement after the 7 weeks USHER intervention is a testament to her excellent teaching. Throughout the intervention she modeled the comprehension strategies and adapted the intervention by making space to reteach when necessary. She also implemented collaborative groups per USHER, which may have contributed to students' improved behavioral engagement in reading. In spite of the fact that she adapted the word maps to include students' drawings of the words, her students may not have improved on academic vocabulary because she frequently cut word maps from her instruction altogether to make time for comprehension.

Discussion

In this study, we drew on a mixed methods approach to delve into four grade 6 teachers' instructional practices in an effort to explain changes (or lack thereof) in their MLs' reading comprehension, academic vocabulary, behavioral engagement, and reading self-efficacy during their implementation of USHER, an intervention based on the model of reading engagement processes within classroom contexts (Guthrie et al., 2012). Moreover, we explored the model of reading engagement with a sample of MLs in an effort to better understand which instructional practices contribute to MLs' reading engagement, motivation, and reading competence.

We found that MLs across all four classes improved in their reading comprehension. All four teachers were strong at explicitly modeling the use of comprehension strategies. Maria was multilingual and many of Carol's special education students were also MLs. Thus, they may have developed expertise in serving this student population. Jerome expertly activated students' background knowledge using text features and scaffolded them in finding main ideas and supporting details. Barbara consistently modeled reading comprehension strategies such as finding the main idea and supporting details and comprehension monitoring. Additionally, all four teachers also regularly elaborated on student responses and repeated important information, which are strategies that support MLs' content area learning (Goldenberg, 2020). Moreover, the teachers made important adaptations to the USHER lesson plans that may have supported the development of their MLs' reading comprehension, including reviewing previously introduced strategies (Barbara), using Spanish to support understanding (Maria), extending time for guided reading (Carol), and making comprehension instruction student-centered (Jerome). These practices are aligned with research on supporting

MLs' language development [i.e., reviewing previously introduced content (Goldenberg, 2020), translanguaging to help students make connections to their home languages (García and Kleifgen, 2010), and increasing peer-to-peer interaction (Washington-Nortey et al., 2022)]. The findings of this study add evidence that these practices not only support MLs' oral language development, but also their reading comprehension.

We took a different approach to explaining why students whose teachers implemented USHER improved in their reading comprehension, academic vocabulary, reading self-efficacy beliefs, and behavioral engagement in reading, as compared to traditional mediational and correlational models (Proctor et al., 2014; Usher et al., 2019; Cho et al., 2021). Instead of explaining our outcomes through participation in the intervention, we chose to explain them by teachers' implementation of the intervention components as well as their enacted changes to the intervention to hypothesize how their implementation may have supported or inhibited their MLs' reading comprehension, academic vocabulary, behavioral engagement, and reading self-efficacy. To be clear, our findings regarding which instructional practices may have influenced the changes in reading comprehension, academic vocabulary, behavioral engagement, and reading self-efficacy in each case are not generalizable beyond these four cases. Our findings do, however, mirror those of Magnusson et al. (2019) who found that teachers find their own ways and methods to integrate reading strategy instruction; thus, in their study the teaching of strategies was adaptive to classroom settings. However, these findings can nevertheless inform the field and suggest directions for future research.

Limitations

The limitations we encountered were due, in part, to the fact that we did not have available to us all pertinent information to fully understand these students' language competencies in English or enough knowledge of their cultural resources to situate the relevant aspects of USHER in more direct alignment with their goals for learning. A second limitation came from our constrained measurement of motivation components of the students in USHER. We limited our focus to self-efficacy for reading, with the express goal of focusing on teacher application of specific motivation practices. Future researchers could look at targeting specific practices in teachers and aligning those with their influence on student outcomes (e.g., how students perceive relevance supports). We hope some of the principles learned through teacher insights, beliefs, and practices in this study can inform such work. Additionally, the use of a teacher-report measure of students' behavioral engagement in reading may have been biased against MLs (Garcia et al., 2019).

Conclusion

The U.S. student population is rapidly diversifying in language, socioeconomic status, and ethnicity. To better understand the complexity of reading instruction for these diverse learners, educational researchers must examine multiple aspects of instruction, including the instructional practices that teachers use to support students' motivation, engagement, comprehension

strategy use, and vocabulary. Mixed methods research lends itself to this type of complex investigation. Our study leaves us with the promise that teachers can and are most likely willing to implement and adapt motivation supports if these are well-embedded in content- and comprehension-rigorous instruction (e.g., Taboada Barber et al., 2015, 2018). Teachers' decisions about changes to instruction were illuminating in showing ways in which researcher-led interventions can be morphed to transform teacher practices, instead of being mere models to be implemented by teachers at the request of researchers. Further, the findings that emerged can inform not only practice, but research aimed to navigate the mutually beneficial collaborations that seek to produce rigorous research about implementing practices to support MLs' reading competence, engagement, and motivation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

This study involved human participants was reviewed and approved by the George Mason University Institutional Review Board. The teacher participants provided their written informed consent to participate in this study. The student participants provided written assent and their parents provided written consent.

Author contributions

AB and MB conceptualized the study, secured the funding, and led the data collection and analysis. ER supported data collection, analysis, and manuscript revisions. JB and MG supported data collection and analysis and conceptualized and wrote most of the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1085909/full#supplementary-material

References

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York, NY: Freeman.

Barber, A. T., Cartwright, K. B., Stapleton, L. M., Klauda, S. L., Archer, C. J., and Smith, P. (2020). Direct and indirect effects of executive functions, reading engagement, and higher order strategic processes in the reading comprehension of dual language learners and English monolinguals. *Contemp. Educ. Psychol.* 61:101848. doi: 10.1016/j.cedpsych.2020.101848

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300.

Board of Regents of the University of Wisconsin System on behalf of the WIDA Consortium. (2014). *2012 amplification of the English language development standards, Kindergarten–Grade 12*. Available at: https://wida.wisc.edu/sites/default/files/resource/2012-ELD-Standards.pdf.

Capin, P., and Vaughn, S. (2017). Improving reading and social studies learning for secondary students with reading disabilities. *Teach. Except. Child.* 49, 249–261. doi: 10.1177/0040059917691043

Carlo, M. S., August, D., Mclaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., et al. (2004). Closing the gap: addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Read. Res. Q.* 39, 188–215. doi: 10.1598/RRQ.39.2.3

Cho, E., Kim, E. H., Ju, U., and Lee, G. A. (2021). Motivational predictors of reading comprehension in middle school: role of self-efficacy and growth mindsets. *Read. Writ.* 34, 2337–2355. doi: 10.1007/s11145-021-10146-5

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences. 2nd* Edn. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Corbin, J., and Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qual. Sociol.* 13, 3–21. doi: 10.1007/BF00988593

Crosson, A. C., McKeown, M. G., Lei, P., Zhao, H., Li, X., Patrick, K., et al. (2021). Morphological analysis skill and academic vocabulary knowledge are malleable through intervention and may contribute to reading comprehension for multilingual adolescents. *J. Res. Read.* 44, 154–174. doi: 10.1111/1467-9817.12323

Crosson, A. C., McKeown, M. G., Moore, D. W., and Ye, F. (2019). Extending the bounds of morphology instruction: teaching Latin roots facilitates academic word learning for English learner adolescents. *Read. Writ.* 32, 689–727. doi: 10.1007/s11145-018-9885-y

Durkin, D. (1993). *Teaching them to read. 6th* Edn. Des Moines, IA: Allyn and Bacon.

Fitri, D. R., Sofyan, D., and Jayanti, F. G. (2019). The correlation between reading self-efficacy and reading comprehension. *J. Engl. Educ. Teach.* 3, 1–13. doi: 10.33369/jeet.3.1.1-13

Gallagher, M. A., Taboada Barber, A., Beck, J., and Buehl, M. M. (2019). Academic vocabulary knowledge of middle schoolers of diverse language backgrounds. *Read. Writ. Q.* 35, 84–102. doi: 10.1080/10573569.2018.1510796

García, O., and Kleifgen, J. A. (2010). *Educating emergent bilinguals: policies, programs, and practices for English language learners*. New York, NY: Teachers College Press.

Garcia, E. B., Sulik, M. J., and Obradović, J. (2019). Teachers' perceptions of students' executive functions: disparities by gender, ethnicity, and ELL status. *J. Educ. Psychol.* 111, 918–931. doi: 10.1037/edu0000308

Goldenberg, C. (2020). Reading wars, reading science, and English learners. *Read. Res. Q.* 55, S131–S144. doi: 10.1002/rrq.340

Green, C., and Lambert, J. (2018). Advancing disciplinary literacy through English for academic purposes: discipline-specific wordlists, collocations and word families for eight secondary subjects. *J. Engl. Acad. Purp.* 35, 105–115. doi: 10.1016/j.jeap.2018.07.004

Guthrie, J. T., and Klauda, S. L. (2015). "Engagement and motivational processes in reading" in *Handbook of individual differences in reading*. Ed. P. Afflerbach (New York, NY: Routledge), 59–71.

Guthrie, J. T., and Wigfield, A. (2000). "Engagement and motivation in reading" in *Handbook of reading research*. eds. M. L. Kamil and P. B. Mosenthal, vol. *III* (New York, NY: Routledge), 403–422.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *J. Educ. Psychol.* 96, 403–423. doi: 10.1037/0022-0663.96.3.403

Guthrie, J. T., Wigfield, A., Metsala, J. L., and Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Sci. Stud. Read.* 3, 231–256. doi: 10.1207/s1532799xssr0303_3

Guthrie, J. T., Wigfield, A., and You, W. (2012). "Instructional contexts for engagement and achievement in reading" in *Handbook of research on student engagement*. eds. S. Christenson, A. L. Reschly and C. Wylie (Boston, MA: Springer), 601–634.

Hwang, H., and Duke, N. K. (2020). Content counts and motivation matters: reading comprehension in third-grade students who are English learners. *AERA Open* 6:233285841989907. doi: 10.1177/2332858419899075

Ingrid, I. (2019). The effect of peer collaboration-based learning on enhancing English oral communication proficiency in MICE. *J. Hosp. Leis. Sports Tour. Educ.* 24, 38–49. doi: 10.1016/j.jhlste.2018.10.006

Kim, J. S., Hemphill, L., Troyer, M., Thomson, J. M., Jones, S. M., LaRusso, M. D., et al. (2017). Engaging struggling adolescent readers to improve reading skills. *Read. Res. Q.* 52, 357–382. doi: 10.1002/rrq.171

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., and Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count, and Flesch reading ease formula) for navy enlisted personnel [Research Branch Report 8–75]*. Millington, TN: Chief of Naval Technical Training.

Lee, J. S. (2012). The effects of the teacher–student relationship and academic press on student engagement and academic performance. *Int. J. Educ. Res.* 53, 330–340. doi: 10.1016/j.ijer.2012.04.006

Leech, N. L., and Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Qual. Quant.* 43, 265–275. doi: 10.1007/s11135-007-9105-3

Lesaux, N. K., Kieffer, M. J., Faller, S. E., and Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Read. Res. Q.* 45, 196–228. doi: 10.2307/20697183

Lutz, S. L., Guthrie, J. T., and Davis, M. H. (2006). Scaffolding for engagement in elementary school reading instruction. *J. Educ. Res.* 100, 3–20. doi: 10.3200/joer.100.1.3-20

MacGinitie, W. H., MacGinitie, R. K., Maria, K., and Dreyer, L. G. (2000). *Gates-MacGinitie reading tests (4th ed.)*. Itasca, IL: Riverside.

Magnusson, C. G., Roe, A., and Blikstad-Balas, M. (2019). To what extent and how are reading comprehension strategies part of language arts instruction? A study of lower secondary classrooms. *Read. Res. Q.* 54, 187–212. doi: 10.1002/rrq.231

Manyak, P. C., Manyak, A. M., and Kappus, E. M. (2021). Lessons from a decade of research on multifaceted vocabulary instruction. *Read. Teach.* 75, 27–39. doi: 10.1002/trtr.2010

Maxwell, J. A. (2013). *Qualitative research design: an interactive approach*. 3rd Edn. Thousand Oaks, CA: SAGE Publications, Inc.

McLaughlin, M. (2012). Reading comprehension: what every teacher needs to know. *Read. Teach.* 65, 432–440. doi: 10.1002/trtr.01064

Merriam, S. B. (2009). *Qualitative research: a guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Mucherah, W., and Yoder, A. (2008). Motivation for reading and middle school students' performance on standardized testing in reading. *Read. Psychol.* 29, 214–235. doi: 10.1080/02702710801982159

National Institute of Child Health and Human Development (NICHD). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (No. NIH Publication No. 00-4769). Government Printing Office. Available at: https://www.nichd.nih.gov/publications/pubs/nrp/pages/smallbook.aspx

Nuland, L. R., Ramirez, E. M., Taboada Barber, A., Sturtevant, E., and Kidd, J. K. (2014). *"It is a tough nut to crack!" case studies of two social studies teachers' experience with enacting content area literacy instruction*. [Poster presentation]. American Educational Research Association, Philadelphia, PA.

Okkinga, M., van Steensel, R., van Gelderen, A. J., van Schooten, E., Sleegers, P. J., and Arends, L. R. (2018). Effectiveness of reading-strategy interventions in whole classrooms: a meta-analysis. *Educ. Psychol. Rev.* 30, 1215–1239. doi: 10.1007/s10648-018-9445-7

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* 66, 543–578. doi: 10.3102/00346543066004543

Parsons, S. A., Malloy, J. A., Parsons, A. W., and Burrowbridge, S. C. (2015). Students' engagement in literacy tasks. *Read. Teach.* 69, 223–231. doi: 10.1002/trtr.1378

Pearson, P. D., and Gallagher, G. (1983). The gradual release of responsibility model of instruction. *Contemp. Educ. Psychol.* 8, 112–123.

Pintrich, P. R., and Schunk, D. H. (1996). *Motivation in education: Theory, research and applications*. Englewood Cliffs, NJ: Prentice Hall Merrill.

Proctor, C. P., Daley, S., Louick, R., Leider, C. M., and Gardner, G. L. (2014). How motivation and engagement predict reading comprehension among native English-speaking and English-learning middle school students with disabilities in a remedial reading curriculum. *Learn. Individ. Differ.* 36, 76–83. doi: 10.1016/j.lindif.2014.10.014

Quin, D. (2017). Longitudinal and contextual associations between teacher–student relationships and student engagement: a systematic review. *Rev. Educ. Res.* 87, 345–387. doi: 10.3102/0034654316669434

ResearchWare Inc. (2013) *HyperRESEARCH*. Availabel at: http://www.researchware.com/products/hyperresearch.html

Rosenthal, R., and Rubin, D. B. (2003). r(equivalent): a simple effect size indicator. *Psychol. Methods* 8, 492–496. doi: 10.1037/1082-989X.8.4.492

Rosenzweig, E. Q., Wigfield, A., Gaspard, H., and Guthrie, J. T. (2018). How do perceptions of importance support from a reading intervention affect students' motivation, engagement, and comprehension? *J. Res. Read.* 41, 625–641. doi: 10.1111/1467-9817.12243

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Thousand Oaks, CA: SAGE.

Salomon, G. (1991). Transcending the qualitative-quantitative debate: the analytic and systemic approaches to educational research. *Educ. Res.* 20, 10–18. doi: 10.3102/0013189X020006010

Schoonen, R., and Verhallen, M. (1998). Kennis van woorden; de toetsing van diepe woordkennis [Knowledge of words; Testing deep word knowledge]. *Paedagog. Stud.* 75, 153–168.

Schunk, D. H., and Miller, S. D. (2002). "Self-efficacy and adolescents' motivation" in *Academic motivation of adolescents*. eds. F. Pajares and T. C. Urdan (Greenwich, CT: Information Age Publishing Inc.), 1–28.

Schunk, D. H., and Mullen, C. A. (2012). "Self-efficacy as an engaged reader" in *Handbook of research on student engagement*. eds. s. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 219–236.

Schunk, D. H., and Zimmerman, B. J. (2007). Influencing children's self- efficacy and self- regulation of reading and writing through modeling. *Read. Writ. Q.* 23, 7–25. doi: 10.1080/10573560600837578

Shehzad, M. W., Alghorbany, A., Lashari, S. A., and Lashari, T. A. (2019). Self-efficacy sources and reading comprehension: the mediating role of reading self-efficacy beliefs. *3L: Lang. Linguist. Lit.* 25, 90–105. doi: 10.17576/3l-2019-2503-07

Shell, D. F., Colvin, C., and Bruning, R. H. (1995). Self-efficacy, attributions, and outcome expectancy mechanisms in reading and writing achievement: grade-level and achievement-level differences. *J. Educ. Psychol.* 87, 386–398. doi: 10.1037/0022-0663.87.3.386

Skinner, E. A., and Pitzer, J. R. (2012). "Developmental dynamics of student engagement, coping, and everyday resilience" in *Handbook of research on student engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 21–44.

Solheim, O. J. (2011). The impact of reading self-efficacy and task value on reading comprehension scores in different item formats. *Read. Psychol.* 32, 1–27. doi: 10.1080/02702710903256601

Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: The Guilford Press.

Taboada Barber, A. M., Buehl, M. M., and Beck, J. S. (2017). Dynamics of engagement and disaffection in a social studies classroom context. *Psychol. Sch.* 54, 736–755. doi: 10.1002/pits.22027

Taboada Barber, A., Buehl, M. M., Beck, J. S., Ramirez, E. M., Gallagher, M., Richey Nuland, L. N., et al. (2018). Literacy in social studies: the influence of cognitive and motivational practices on the reading comprehension of English learners and non-English learners. *Read. Writ. Q.* 34, 79–97. doi: 10.1080/10573569.2017.1344942

Taboada Barber, A., Buehl, M. M., Kidd, J. K., Sturtevant, E. G., Richey Nuland, L., and Beck, J. (2015). Reading engagement in social studies: exploring the role of a social studies literacy intervention on reading comprehension, reading self-efficacy, and engagement in middle school students with different language backgrounds. *Read. Psychol.* 36, 31–85. doi: 10.1080/02702711.2013.815140

Taboada Barber, A., Gallagher, M. A., Buehl, M. M., Smith, P., and Beck, J. (2016). Examining student engagement and reading instructional activities: English learners' profiles. *Lit. Res. Instr.* 55, 209–236. doi: 10.1080/19388071.2016.1167987

Unrau, N. J., Rueda, R., Son, E., Polanin, J. R., Lundeen, R. J., and Muraszewski, A. K. (2018). Can reading self-efficacy be modified? A meta-analysis of the impact of interventions on reading self-efficacy. *Rev. Educ. Res.* 88, 167–204. doi: 10.3102/0034654317743199

Usher, E. L., Li, C. R., Butz, A. R., and Rojas, J. P. (2019). Perseverant grit and self-efficacy: are both essential for children's academic success? *J. Educ. Psychol.* 111, 877–902. doi: 10.1037/edu0000324

Vaughn, S., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., et al. (2011). Efficacy of collaborative strategic reading with middle school students. *Am. Educ. Res. J.* 48, 938–964. doi: 10.3102/0002831211410305

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., and Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: findings from two experimental studies. *J. Res. Educ. Effect.* 2, 297–324. doi: 10.1080/19345740903167018

Vaughn, S., Martinez, L. R., Wanzek, J., Roberts, G., Swanson, E., and Fall, A. M. (2017). Improving content knowledge and comprehension for English language learners: findings from a randomized control trial. *J. Educ. Psychol.* 109, 22–34. doi: 10.1037/edu0000069

Verhoeven, L., Voeten, M., and Vermeer, A. (2019). Beyond the simple view of early first and second language reading: the impact of lexical quality. *J. Neurolinguistics* 50, 28–36. doi: 10.1016/j.jneuroling.2018.03.002

Washington-Nortey, P. M., Zhang, F., Xu, Y., Ruiz, A. B., Chen, C. C., and Spence, C. (2022). The impact of peer interactions on language development among preschool English language learners: a systematic review. *Early Childhood Educ. J.* 50, 49–59. doi: 10.1007/s10643-020-01126-5

Wigfield, A., Gladstone, J. R., and Turci, L. (2016). Beyond cognition: reading motivation and reading comprehension. *Child Dev. Perspect.* 10, 190–195. doi: 10.1111/cdep.12184