Teaching Academic Vocabulary to Sixth Grade Students with Disabilities

Rollanda E. O'Connor, Ph.D. (rollanda.oconnor@ucr.edu)

University of California at Riverside

Graduate School of Education, 1207 Sproul Hall, Riverside, CA 92521

Kristen D. Beach, Ph.D. (kristen.beach@uncc.edu)

University of North Carolina, Charlotte

Department of Special Education and Child Development

9201 University City Blvd., Charlottle, NC 28223

Victoria M. Sanchez, Ph.D. (victoria.sanchez@ucr.edu), Joyce J. Kim, Ph.D. (joycejungheekim@ gmail.com), Kerri Knight Teague, Ph.D. (kerri.knightteague@gmail.com), Guadalupe Orozco, M.A. (gguzm008@ucr.edu), and Brian T. Jones, MA (bjone011@ucr.edu)

University of California at Riverside
Graduate School of Education, 1207 Sproul Hall, Riverside, CA 92521

This research was supported in part by Grant R324A160019 from the U.S. Department of Education, Institute of Education Sciences, to the University of California at Riverside. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Submitted July 20, 2018 to *Learning Disability Quarterly*
Revision submitted November 1, 2018
Final accepted November 18, 2018

Contact Author:
Rollanda E. O'Connor
GSOE, 1207 Sproul Hall
Riverside, CA 92521
Rollanda.oconnor@ucr.edu
951-827-6052

Abstract

Adolescents with disabilities have great difficulty with academic content in middle school, and their teachers have difficulty teaching them to understand and use academic language. We taught teachers of 6th grade students with Learning Disabilities, over half of whom were also English Language Learners (ELL), to implement about 15 minutes of daily interactive vocabulary instruction in their intact special education English/Language Arts classes. Three schools were assigned randomly to treatment (2 schools) or control conditions (1 school; 52 students total). We developed instructional routines to introduce four new words per week in three 4-week units to test for replicability. ANCOVAs (with each cycle's pretest and intelligence quotient (IQ) as covariates) were conducted on taught vocabulary, all of which favored the treatment condition with effect sizes ranging .6 to .7 per cycle. Near-transfer effects to vocabulary usage were weaker, with significant effects in the last two cycles. Effects were similar for students with disabilities who were ELL and native English speakers. Treated students maintained their knowledge of words 4 to 24 weeks following the close of treatment.

Keywords: vocabulary instruction, academic language, learning disabilities, English learners, teacher implementation, replication, CHAAOS

Teaching Academic Vocabulary to Sixth Grade Students with Disabilities

Students with Learning Disabilities (LD) and other disabilities often have difficulty succeeding in middle and high school courses due to poor reading and language skills. Data from The Nation's Report Card (NAEP, 2017) report that only 10% of students eligible for special education in 8th grade scored as proficient in reading. Many of these students who receive special education services also have English Language Learner (ELL) designations, and in the same NAEP report only 5% of students who were ELL scored in the proficient range or above.

Although many aspects of reading and language difficulties constrain students' ability to engage in content area courses, one of the most difficult to shift meaningfully has been vocabulary knowledge (Scammaca et al., 2007). Knowledge of academic vocabulary--the types of words used in school and in textbooks--impacts reading comprehension in late elementary and on through middle and high school (LaRusso et al., 2016). In their theory of reading comprehension, Perfetti and Stafura (2014) place word knowledge (i.e., linking decoding with the meanings and usage of words) at the center of reading comprehension. Rising standards for academic literacy (e.g., the Common Core State Standards, CCSS, National Governors Association, 2010; CCCSSI, 2017) create an urgency for students to learn the meanings of academic words and use them appropriately.

The types of words used in school shift from conversational English toward the language of textbooks around the fourth or fifth grade (Chall, 1983; Hiebert, Goodwin, & Cervetti, 2017). Most vocabulary growth from fourth grade on stems from wide and varied reading (Anderson, Wilson, & Fielding, 1988). As examples, Cipielewski and Stanovich (1992) followed students from grades 3 to 5 and found the time students spent reading text predicted vocabulary and reading comprehension growth. Echols, West, Stanovich, and Zehr (1996) studied 5th and 6th grade students with similar findings, focused in particular on growth in verbal cognitive skills linked with wide reading. Unfortunately, many students with LD read little in or out of school. Compounding the problem, the level of the materials students with disabilities may read tends to be several years below their grade-level placement, which offers few opportunities to learn the meanings of academic words independently and makes teaching vocabulary crucial.

Teaching words well requires multiple exposures and opportunities for students to use the words in their conversations and writing (Lawrence, Crosson, Pare-Blagoev, & Snow, 2015; McKeown, Crosson, Moore, & Beck, 2018). Students with LD who are poor readers need more explicit instruction and opportunity than do typical learners (Bulgren, Graner, & Deshler, 2013; Elleman, Lindo, Morphy, & Compton, 2009; Kamil et al., 2008). Moreover, many students with LD are also ELL, whose English language difficulties complicate learning advanced vocabulary in English (Bunch, Walqui, & Pearson, 2014). Shifting teachers' instruction to include sufficient time and practice on word meanings and usage could be difficult. For example, Cunningham, Zibulsky, Stanovich, and Stanovich (2009) surveyed teachers regarding their preferences for teaching particular aspects of reading. Of the 13 instructional categories, teachers responded they would spend 1.6% of their time on vocabulary, which translates to two minutes per day in a 90-minute reading block.

Analysis of transcripts from 26 business-as-usual classrooms (Apthorpe et al., 2012) failed to find examples of multiple exposures to word meanings or opportunities to practice and use words and their meanings in typical English/Language Arts instruction. Words introduced to students tend to be tied to a particular passage, taught "just in time" to help students understand a particular event as they read about it. The words taught might not be reviewed, incorporated in new contexts, or accumulated over time.

Recently several teams have conducted large-scale research on vocabulary instruction in classrooms (e.g., Lawrence et al., 2015; Lesaux, Kieffer, Faller, & Kelley, 2010; McKeown, Crosson, Moore, & Beck, 2018)) and these routines have been somewhat successful with typical learners, but less effective for students with poor reading skills than for average readers. Although vocabulary researchers have analyzed data from interventions for the students with disabilities or for status as an ELL, rarely have they included analyses for students who were LD and ELL.

Regarding ELL, LaRusso et al. (2016) found these students grow slower than their typically developing peers in English vocabulary and reading comprehension. Coleman and Goldenberg's (2010) summary of research with ELL stresses that academic vocabulary is "one of the most pressing needs faced by English Language Learners." (p. 61) In a meta-analysis of reading interventions for middle grade students who were ELL, Hall et al. (2017) found positive effects for reading overall; however, few studies generated positive effects in vocabulary, and no studies in that sample explored improvements for students who were ELL and also had disabilities.

Working in low-performing general education English/Language Arts (ELA) classes, Lesaux et al. (2010) taught teachers to implement Academic Language Instruction for All Students (ALIAS), an 18-week instructional program. Importantly, these classes included high proportions of ELL. Teachers taught eight-to-nine academic words during each of a series of two-week units. At the close of treatment, students receiving ALIAS outperformed control students on meanings of the taught words, but not on standardized measures of vocabulary, and those who were ELL made similar gains to students who were native English speakers. These results were replicated in a larger study (Lesaux, Kieffer, Kelley, & Harris, 2014); however, neither study analyzed data for students with disabilities.

To stimulate better application of word learning, several researchers have tied vocabulary instruction to particular academic content. For example in history contexts, researchers have tested procedures for vocabulary instruction with middle and high school students with low reading ability (O'Connor, Sanchez, Beach, & Bocian, 2017; Swanson, Wanzek, Vaughn, Roberts, & Fall, 2015). In a complex treatment that included instruction in essential vocabulary in history classes, Vaughn, Martinez, and Wanzek (2017) found improvements for students who were ELL over similar students in untreated classes; however, as the proportion of students who were ELL in these classes increased, class averages decreased. This finding matters because many classes currently include large proportions of students who are ELL. Results for students who received special education services were not evaluated in this study.

Nevertheless, tying words closely to a single content area may limit generalization to new contexts and situations. Lawrence et al. (2015) addressed this issue by extending their research on *Word Generation*, a program with instruction that crosses content areas to include ELA, mathematics, and science. In a large, randomized trial, the percent of students with disabilities who participated in the general education classes was used as a control variable; however, their scores were not analyzed separately. The positive effect of *Word Generation* on taught vocabulary was small (.25), and the intervention's effect was moderated by the quality of discussion among teachers and their students. Better readers made stronger gains and the percent of special education in the classes had a small negative effect on outcomes. Special education teachers were not participants in delivering the intervention.

The most recent in McKeown and colleagues' long line of vocabulary research in general education classes, McKeown, Crosson, Moore, and Beck (2018) conducted a two-year study of their Robust Academic Vocabulary Encounters (*RAVE*) in 6^{th} and 7^{th} grade classrooms. Teachers introduced nine words during each 7-day cycle over 22 weeks of total instruction. Neither students receiving special education nor ELL were mentioned. Fidelity of implementation was high and *RAVE* students made significant gains on taught words.

Most instruction in these studies occurred in general education classes; however, many students with LD receive their English/Language Arts instruction (ELA) with special education teachers, who have not been included in most experimental studies of teaching academic language. Thus, most special education teachers lack access to instructional materials designed to teach academic vocabulary to students who have learning difficulties. Importantly, to date, few studies have produced lasting vocabulary improvement for students with disabilities, and fewer have considered interaction effects for students with disabilities who were also learning the English language.

Studies are needed that evaluate results of vocabulary intervention for students with disabilities who are or are not ELL. In a rare example that analyzed potential effects of both disability and ELL status, Wanzek, Swanson, Vaughn, Roberts, and Fall (2016) tested an intervention that included instruction on 15 academic words focused on social studies content in general education classes. On a multiple choice test of questions on taught content, students with disabilities (primarily LD) in these classes scored higher than students with LD in comparison classes. ELL status did not interact with these positive results. Nevertheless, on questions that required reading paragraphs about taught content, treated students with disabilities did not outperform comparison students with disabilities. Thus application of learned vocabulary was still difficult for these students, regardless of native language. Importantly, maintenance of vocabulary learning was not evaluated.

The Current Study

Our purpose was to develop an intervention to teach academic words to students in special education classes in middle school that included high proportions of students who were ELL. We reasoned that to be feasible and effective, teachers would need to be able to implement short (i.e., 15 minutes), focused instruction on important academic words in their special education ELA classes, and students should both learn the words and retain them over time. We assisted special education teachers to implement the instruction with their intact classes. Our research questions were:

1. Will 6th grade students with disabilities who receive 12 weeks of vocabulary instruction outperform students with disabilities who do not receive this instruction on academic word meanings and usage?
2. How well do students maintain their learning of vocabulary after the close of instruction?
3. Do gains differ between students with disabilities who are ELL and native English speakers?
4. To what degree of fidelity do teachers in special education ELA classrooms implement the lessons?

Method

This study was conducted in one urban school district in southern California with just under 20,000 students and four middle schools, each including 11-13% of students receiving special education services. Each school had one 6th grade English/Language Arts special education teacher. The principals and special education teachers in three of the middle schools agree to participate.

The teachers were assigned randomly to treatment (2 classes) or business-as-usual (BAU) control (1 class) conditions. Treatment teachers had 4 and 28 years of teaching experience and the BAU teacher had 7 years of experience. All were female, Caucasian, and held special education credentials.

Their 52 participating students were all eligible for special education (83% under LD eligibility), 70% were Latino/a, and 60% of these students with disabilities were ELL. These percentages were similar for the district overall. On the California English Language Development Test (CELDT), administered by the district to all ELL, English reading and writing scores range from 1-5, with 5 being high. Scores for the participating students who were ELL also ranged from 1-to-5 and are shown by class in Table 1. Language preference, as identified by their parents, is also shown in Table 1.

Insert Table 1 about here

Special Education Classroom Context

Although all of these students participated in some general education classes, this district assigned students with LD whose reading scores were more than 2.5 years below grade level to receive ELA in a special education class. The only significant difference on student demographics for condition was a higher proportion of males in treatment classes.

All classes followed the same district schedule of 55 minutes of ELA daily. During the first 12 to 15 minutes of class, teachers conducted warm-up activities that included group-led "find the errors" in a paragraph projected on a screen, work sheets on which students circled correct choices to fill in sentences, and correction of the previous night's homework. Teachers in all three classes used primarily *Corrective Reading Decoding Strategies* (Engelmann et al., 1999) three to four days per week in September through March for approximately 30 minutes. In

January teachers set aside *Corrective Reading* and used novels (i.e., *Holes* (Sachar, 1998) in one treatment class, *The Giver* (Lowry, 1993) in the other) or mythology (i.e., MythologyTeacher.com in the BAU class) as the basis for reading instruction and practice. In all classes, brief vocabulary instruction was incorporated in the comprehension questions of *Corrective Reading* or selected from the novels and passages prior to reading the day's text. Vocabulary activities ranged from 2 to 9 minutes in any observation and averaged just under 5 minutes. In the treatment classes, teachers replaced part of their reading text time with the vocabulary lessons and continued with the warm-up activities that were part of their class routine.

Intervention

Vocabulary words. We selected words for utility in middle school and triangulated words for instruction across three major sources: (1) Coxhead's Academic Word List (2000), which contains academic words that occur frequently in textbooks across content areas; (2) Biemiller's *Words Worth Teaching* (2010), which lists academic words based on when students typically acquire them and how difficult they may be to teach; and (3) words from Tier 2 grade level lists developed for the Common Core State Standards (National Governors Association, 2010). The cross referenced lists were reviewed by the research team and 55 words were selected. A four-item multiple choice test was created by the research team as a pilot assessment to further narrow the word selection. This assessment was given to 23 sixth graders in two special education ELA classes in a different school district the year prior to the current study. We used the 48 words most frequently missed to develop the instructional routine, which we call CHAAOS (Creating Habits that Accelerate Academic language of Students). We shared this list of words with all three teachers (CHAAOS and BAU) to ensure they viewed the words as important and useful for instruction in sixth grade.

The 48 words were divided into 3 sets of 16 words each. A student friendly definition was created for each word, with the Collins Cobuild English Language Dictionary as a frequent reference. We examined the sets for words that were similar in appearance or sound (*prejudice*, *preliminary*) or meaning (*perceive*: to see or notice and *indicate*: to see and point out). These were separated into different sets and cycles of instruction.

Key lesson components. We constructed lessons based on research on adolescent learning (Kamil et al., 2008) and effective instruction for students with LD (Swanson & Deshler, 2003), as well as effective vocabulary instruction for students who are ELL (Baker et al., 2014; Carlo et al., 2004; Coleman & Goldenberg, 2010; Hall et al., 2017), which include active and interactive learning, cumulative introduction and maintenance, and ongoing formative assessment. Instructional design included the five essential components of explicit instruction as outlined by Hughes, Morris, Therrien, and Benson (2017). Specifically, instruction built understanding cumulatively across a series of lessons for each word set, teachers modeled usage of words across contexts and in writing, supports and prompts were faded as students gained skill with each word set, and students received ample opportunities for oral and written practice with feedback. To ease teacher use of lessons and students' learning, we followed a predictable routine for introducing and contextualizing the four new words per week, and for integrating novel words with known words and contexts across the 12 weeks of instruction. This predictable 15-minute routine is outlined below.

Week 1: Monday. Introduce the four new words and their synonyms or definitions in adolescent-friendly contexts, using appropriate graphics to demonstrate usage. Teachers and students discuss meanings of words across contexts. Students complete cloze sentence writing based on context word play and answer questions. In the last minutes, students write the words and meanings in their vocabulary journals.

Week 1: Tuesday. Brief review of the four new words and concentration on two of the words to explore more deeply. Students generate sentences with picture prompts and sentence stems, reinforcing reading words in isolation and context. Peer groups generate a sentence for the two words and discuss how to depict the words' meaning. Teachers continue with whole group discussion of confusions over meanings, justifications of usage, and extensions of meanings.

Week 1: Wednesday. As on Tuesday, but with the other two focus words for the week.

Week 1: Thursday. Teachers lead students through activities that cross all four words and through a small-group task in which students justify why or how particular words make sense. Students decide on the correct words to use when given a context that includes gradients of meanings (e.g., earn, acquire, attain, steal). Students use two words in sentences they write independently, and discuss as a large group.

Weeks 2-3. The pattern above continues; however, words from previous weeks are used in contexts to illustrate the new words of the week. Each Thursday includes integrative activities across the word set.

Weeks 4-12. The only changes are in the integrative activities, which include each week at least 8 words from across the previously taught words. The 12 weeks of teacher presentation PowerPoint slides and measures are available on the first author's website (add website here: blinded for review).

Teacher Training and Fidelity Tool

Teachers take varying amounts of time to learn and integrate new instructional skills, especially when they take student responsiveness and learning into account (Grierson, 2013). Professional development for teachers used an apprenticeship model and began with a two-hour review of research that supports the techniques we integrated into the lesson design and discussion of each component of our teacher observation tool. Cycle 1 was relayed to teacher participants as a pilot study where their feedback on all materials was deemed pivotal to the improvement and iterative design of the CHAAOS vocabulary intervention program. The following week, and for the first three-to-seven days of instruction, the lead researcher taught each lesson to the teacher's class as the teacher rated the presence or absence of each instructional component on the teacher observation tool. Following each day's implementation, the teacher and researcher discussed aspects of lesson delivery and student engagement. The researchers used teachers' suggestions during Cycle 1 to improve the design of the Teacher Handbook and format of Powerpoint slides used in Cycles 2 and 3.

One teacher began full implementation on Day 4, the other on Day 8. Thereafter, teachers implemented the lesson as observers used the tool to rate presence or absence of each component in the first cycle, and quality of instruction during Cycles 2 and 3, in which teachers delivered all of the instruction.

Measures

Descriptive measures were gathered prior to beginning treatment, and vocabulary measures before and after each instructional cycle and 4 to 24 weeks afterward. Descriptive measures included the Wechsler Abbreviated Scale of Intelligence (WASI-III; Wechsler, 2011) and the reading portions of the Woodcock-Johnson Tests of Achievement III (WJ-III; 2001).

The WASI-III (Wechsler, 2011) is a norm-referenced test intended to provide evidence of verbal, performance, and full-scale intelligence. We administered the two subtest version in the fall. Loadings for vocabulary on Language and matrix design on Non-Language were strong. Across subtests, reliabilities ranged from .81-.98.

We used the WJ-III as a normed measure of reading ability prior to intervention to describe students' reading ability and confirm that participants had severe reading difficulties. The Word Identification subtest requires students to read aloud from a list words that increase in difficulty. Word Attack measures decoding of increasingly difficult nonwords, Picture Vocabulary requires students to point to pictures and name objects, and Passage Comprehension requires students to read short sentences or passages (increasing in difficulty) and provide a contextually appropriate word for a deletion within the passage. Across subtests, reliabilities ranged from .81-.94. We report the Total Reading score in Table 2.

Insert Table 2 about here

Experimental Measures

For each of the three 16-word cycles, we developed a near measure of vocabulary learning, which was a 16-item multiple choice test of the taught words. The four choices included one correct answer, one morphologically related incorrect choice, one orthographically similar incorrect choice, and one unrelated incorrect choice. Test-retest reliability ranged from .88 to .96 across cycles and item difficulty is reported in the results section.

We also developed near transfer measures of vocabulary usage in contexts that were not taught, which included cloze passages and reading comprehension of passages containing some of the vocabulary words from that cycle and comprised 8 items. These measures were administered before and after each instructional cycle. The measures of vocabulary knowledge and usage are available on the first author's website.

Last, we developed an 18-word vocabulary maintenance measure that drew six words from each multiple choice measure from each of the three cycles. We selected the six words on which students scored lowest at pretest and highest on the immediate posttest, thus indicating words on which students showed the strongest growth in word knowledge for that cycle. We used this measure at the end of the school year.

Fidelity of Intervention

Intervention training in August, 2016 targeted teacher participants' understanding of the vocabulary instruction research base and intervention design. The researcher modeled the first lesson of instruction with use of the teacher manual, PowerPoint slides, and student workbook. Teachers were given copies of the observation-fidelity tool, and components on the tool were discussed in relation to the instruction they had just observed.

During the first week of Cycle 1 instruction, teacher participants used the observation-fidelity tool as they observed the lead researcher demonstrate the lessons with their students in their intact classrooms. As teachers implemented the lessons, the researchers observed their instruction daily to ensure that teachers were practicing the well-defined protocol. In this first four-week cycle, only the presence or absence of treatment components was marked on the observation-fidelity forms. Observation notes and teacher feedback on ease and difficulty of lesson components were used to revise lesson format and provide feedback to teachers prior to the next cycle of instruction.

During Cycles 2 and 3, teachers implemented lessons with new sets of 16 words per cycle. Observers during these cycles rated instruction for fidelity and student engagement. Each intervention teacher was observed eight to twelve times per cycle to document presence of instructional components (yes/no), instructional time (minutes), and quality of implementation on a scale of 1-3.

Following the conclusion of the CHAAOS intervention in March, teachers resumed the routines for ELA they had used prior to treatment. After we administered the final maintenance test of CHAAOS vocabulary, no further observations were conducted.

Results

ANOVA on pretests of WASI-III and WJ-III across classes did not differ [$F(2, 45) = 0.69$ and 1.41, respectively, $p > .10$], nor did pretests of C1 vocabulary [$F(2, 45) = 0.51$ and 0.58, respectively, $p > .10$]. ANOVA for ELL status was significant on the WJ-III and the WASI-III [$F(2, 46) = 6.41$ and 4.72, respectively, $p < .05$]; students with disabilities who were also ELL scored about 10 points (2/3 SD) lower than NES (standard scores of 75.62 and 61.27, for ELL respectively). We ran correlations between WASI Full-scale IQ and each of the vocabulary pretest variables to determine whether it was proper to use these variables as covariates in the analyses of vocabulary outcomes. These correlations ranged from .31 to .59, suggesting that each covariate could explain unique variance on the cycle posttests. Because pretests were not used to select or assign students to condition, the pretest range restriction on vocabulary was less concerning than had we used these measures for group assignment (Miciak, Taylor, Stuebing, Fletcher, & Vaughn, 2016). With just three classes in this pilot study, we analyzed individual student data rather than class-level composites (note that pretests by class did not differ). Next, we conducted a series of ANCOVA for each of the three cycles on learning taught words, including the multiple choice measures and the near-transfer measures of usage of the words. We assessed maintenance of learned words with repeated measures (pretest, posttest, follow up four weeks later). Last, we conducted two-way ANCOVAs (ELL status by treatment condition) using pretests as covariates to evaluate potential differential effects for students who were ELL, and evaluated fidelity of implementation.

Vocabulary

For the multiple-choice (MC) test in Cycle 1, ANCOVA indicated a significant main effect for treatment [F (1, 38) = 72.71, *p* = <.001, Ƞ2 = .68] and a significant effect of the pretest covariate on the posttest [F (1, 38) = 10.81, *p* = .002, Ƞ2 = .24]; however, the IQ covariate was not significant. Results were similar on the MC scores in Cycles 2 and 3, with main effects for treatment and for pretest, but not for IQ. Partial Eta Squared values for treatment were .64 and .54 in Cycles 2 and 3, respectively. Table 3 provides summaries.

Insert Table 3 about here

On tests of usage of vocabulary (near transfer), results were weaker. Note that neither the cloze passages nor the comprehension paragraphs on this measure were used instructionally. ANCOVA of Cycle 1 usage with pretest and IQ as covariates indicated nonsignificant effects for treatment (*p* = .063, Ƞ2 = .10) and IQ, but a significant effect for usage pretest (*p* <.001, Ƞ2 = .36). In Cycles 2 and 3, treatment effects were found (*p* =.008 and .002, respectively, Ƞ2 = .18 and .23), and IQ also exerted a significant effect. These statistics are shown in Table 4.

Insert Table 4 about here

Item difficulty. We calculated item difficulty coefficients for each item of the multiple-choice tests, at pre- and post- test, and for each cycle of instruction for students who received the CHAAOS intervention. Item-level difficulty coefficients represent the proportion of students with a correct response for that item. In the context of the multiple-choice test, item difficulty coefficients represent the proportion of students who selected the correct definition for the target word. Average difficulty coefficients of items at pretest decreased from Cycles 1 to 3, Cycle 1 (M = 0.40; SD = 0.08), Cycle 2 (M = 0.48; SD = 0.18), Cycle 3 (M = 0.56; SD = 0.14), suggesting that higher proportions of students answered pre-test items correctly with each subsequent pretest. The average difficulty coefficients of items at post-test were similar and high across cycles, Cycle 1 (M = 0.80; SD = 0.08), Cycle 2 (M = 0.86; SD = 0.06), Cycle 3 (M = 0.86; SD = 0.08). Thus, similarly high proportions of students (80% - 86%) answered items correctly at post-test across all three cycles of instruction.

To estimate average word learning within cycle, we calculated the difference in difficulty coefficients between pre- and post-test for each item. The resulting item-level difference coefficients describe the difference in the proportion of students who answered the item correctly at post-test compared to pre-test. Next, we calculated the mean difference coefficients across all items within each cycle. These coefficients describe the average change in the proportion of students who answered items correctly at post-test compared to pre-test. Results indicated that word learning decreased slightly over time, Cycle 1 (M = 0.41, SD = 0.08), Cycle 2 (M = 0.39, SD = .19), Cycle 3 (M = 0.30, SD = 0.11). However, this decrease is likely due to the higher proportions of students knowing words at pretest across cycles, especially because the proportions of students knowing words at post-test were similar across cycles.

We were also interested in whether certain words were more difficult to learn than other words. To this end, we examined the item-level difference coefficients between words within each cycle. Words considered “difficult” to learn had smaller than average difficulty coefficients at pretest and smaller than average difference coefficients. Thus, difficult words were known by few students at pretest, and were learned by proportionally fewer students compared to the average word in that cycle. Only 6 of the 48 words taught across cycles (12.5%) were difficult to learn: *consent*, *establish*, *predict*, *vary*, *conserve*, and *unify*. Words that were easiest to learn were those with smaller than average difficulty coefficients at pre-test and larger than average difference coefficients. These words were known by few students at pretest, and learned by proportionally more students than the average word. Nineteen of the 48 words (39.6%) were easy to learn: *emerge, alter, credible, attain, contrast, contribute, rigid, negative, interpret, restrict, prejudice, preliminary, indicate, controversy, primary, precise, unique, apprehensive*, and *stable*.

Maintenance of Vocabulary Words

The three cycles of instruction spanned October through March, and we were interested in whether treated students maintained knowledge of the words they had learned. Recall that the maintenance test was constructed from the six words per cycle that showed the most growth during each four weeks of instruction for a total of 18 words. This measure was administered to treated students four weeks following the close of the last instructional cycle, giving a range of 4 to 24 weeks of maintenance across cycles.

We conducted a repeated measures ANOVA (pretest, posttest, maintenance) to determine the number of words known initially, learned during the cycles, and whether students maintained gains on taught words. Mauchly's test showed the assumption of sphericity was upheld ($\chi2$ (2) = 3.28, p = .194). ANOVA results showed a significant main effect of time on the average number of taught words known by participants (F (2, 27) = 121.04, p < .001). Planned comparisons using the Bonferroni correction revealed that student performance on the 18-word measure did not differ significantly between posttest ($\bar{x}$ = 13.69; SD = 4.05) and maintenance ($\bar{x}$ = 13.45; SD = 4.21; p = 1.00); thus students retained the meanings of words they learned during the intervention.

ELL vs NES Comparisons

Two-way ANCOVAs (treatment/BAU by ELL/NES) were conducted to assess whether ELL status influenced effects of treatment on the multiple choice vocabulary posttest in each instructional cycle, with pretest as the covariate. We did not use IQ as a covariate for this analysis, given its emphasis on vocabulary in English. Following Cycle 1, main effects were found for treatment ($F(1,38) = 82.74$, $p < .01$), for pretest ($F(1,38) = 11.21$, $p = .002$), and for ELL status ($F(1,38) = 4.77$, $p = .036$); however, the interaction was not significant ($F(1,38) = .06$, $p = .81$). The same results were found for Cycles 2 (Interaction $F(1,38) = .104$, $p = .75$) and 3 (Interaction $F(1,38) = .244$, $p = .62$). Students who were ELL began with lower pretest scores but gained similarly to NES during instruction.

We also assessed the effect of ELL status on the near transfer measure of vocabulary usage with two-way ANCOVAs. In Cycle 1, main effects were found for pretest ($F(1,38) = 16.151$, $p <.01$) and treatment ($F(1,38) = .4.810$, $p = .035$); however, neither ELL status nor the interaction term were significant. In Cycle 2, only the treatment effect was significant ($F(1,38) = 7.736$, $p = .008$); in Cycle 3 main effects were found for pretest ($F(1,38) = 9.541$, $p = .004$) and treatment ($F(1,38) = 8.485$, $p = .006$), but not for ELL status or the ELL by treatment interaction. These results suggest that students who are ELL did not respond differentially to the vocabulary treatment as measured bv near-transfer usage measures.

Treatment Fidelity

We conducted 28 observations of treatment teachers in Cycle 2 and 18 in Cycle 3 to document presence of instructional components (yes/no), instructional time (minutes), and quality of implementation on a scale of 1-3. Inter-rater reliability (IRR) was established between two raters at 92% percent agreement on eight observations. Observers provided feedback to teachers following daily classroom observation and audio recording, which confirmed consistency of program delivery. Feedback included praise for portions of lessons implemented well, and when needed, one suggestion for improving lesson delivery or student engagement. Student workbooks were collected at the close of each cycle as evidence of student participation.

Presence of instruction was 100% for both teachers and treatment dosage (i.e., meeting the goal of at least 15 min per session) was almost universally met. Intervention time varied from 13-27 minutes per session, averaging 23 minutes, with more time spent on days with higher proportions of peer activity. Lesson time decreased as teachers became more efficient with materials and brisk in their style of teaching, such that instructional time averaged 25 minutes in Cycle 2 and 19 minutes in Cycle 3. Quality of implementation varied between teachers but was generally high for both (average 2.5 of 3). The area of most difference between teachers was in student engagement/student support, with one teacher scoring 3s consistently and the other teacher scoring a mix of 3s, 2s, and occasionally 1s (average 2.5). Teachers averaged highest scores in encouraging frequent responses from students and immediate and corrective feedback during whole-class portions of the lessons. The most difficult part, with lowest scores, was in scaffolding student responses during their construction of sentences.

Our primary emphasis was on quality of implementation and fine-tuning of lesson components, and so most observations were in treatment classes. Due to time constraints, we observed the BAU teacher formally only twice, once each during the timeframe of Cycles 2 and 3. The total time devoted to ELA in the three classes was consistent across condition; however, the amount of vocabulary focus in the BAU observations was less than in CHAAOS classes. During Cycle 2 the BAU teacher used a project-based approach, in which she selected five vocabulary words from the Greek mythology unit of study. The first listed adjectives that could be applied to Greek gods and goddesses, and students were asked to draw lines connecting appropriate adjectives to characteristics. The second reviewed a list of science terms to be used in an upcoming science class, using a word bank to match words with definitions. Following this activity, the teacher used Corrective Reading without adding vocabulary instruction. In both instances, students were encouraged to work together to complete assignments prior to the teacher providing the correct answers. These activities ranged from five to nine minutes of the ELA period.

Discussion

The aims of this study were to test the immediate and lasting effects of a structured academic vocabulary instructional routine designed for and delivered by special education teachers to students in their intact classes, to determine whether effects differed for students with disabilities who were ELL and NES, and to describe the degree to which teachers implemented CHAAOS with fidelity. The intervention development was guided by principles of explicit instruction (Hughes et al., 2017), and drew from recommendations for providing instruction for adolescents with LD (Kamil et al., 2008; Swanson & Deschler, 2003) and for ELL (Baker et al., 2014). The expectation was that by providing structured, explicit instruction with multiple opportunities to practice, repeated exposure across contexts, and embedded review, students' knowledge of words would improve.

Results of our first research question, which examined differences in vocabulary knowledge post-intervention between the intervention group and BAU, showed that students who received the intervention outperformed BAU on a multiple-choice measure of vocabulary knowledge, with large effects that were replicated across three cycles of instruction. Prior research has demonstrated that students' definitional knowledge of taught vocabulary can be improved through intervention in general education settings (Lawrence et al., 2015; Lesaux et al., 2010;2014; McKeown et al., 2018)). Results of the current study demonstrate that such gains are possible for students receiving ELA instruction from their special education teachers.

It is important to note that the design of the lessons in the current study included components found in successful general education vocabulary interventions, such as direct instruction on word meanings, incorporating contextual practice, providing multiple response opportunities, and structuring activities to promote deep processing. However, because we were working with a group of students reading far below grade level, many of whom had limited oral vocabularies, we focused instruction on a smaller set of words (i.e., 4 per week rather than 6-8), considerable oral practice and discussion, which was also found to influence results in Lawrence et al., and use of instructional texts with controlled reading levels during practice. Without these modifications, students may have not made the gains evidenced in the current study. The small standard deviations on item difficulty at posttest show that most students learned most of the words, with mean posttest scores across cycles of 80-86% correct responses.

In addition to the large effects on the multiple-choice measure, small effects were found on a usage measure, which included cloze and reading comprehension tasks that assessed deeper knowledge of words. This result is similar to earlier research with typical learners (McKeown et al., 2018) and ELL (Lesaux et al., 2010), where effects on word usage tasks were not as strong as those on multiple choice (word knowledge) tasks. Like typical learners, students with disabilities have difficulty applying learned words to novel contexts that require deeper, more nuanced word knowledge. Though students in the current study also received practice and support on tasks that required deep word knowledge during instruction (two days per word for contextual applications), the support may not have been sufficiently intensive to translate to strong gains in depth of vocabulary knowledge for these participants. Additional practice applying word meanings to novel oral and written contexts may be needed.

To answer our second research question, we examined whether students maintained knowledge of taught words 4 to 24 weeks after instruction. Results indicated strong retention. The careful lesson design, which included embedded review of words within and across cycles, may have contributed to students' ability to retain word meanings up to 6 months after the first cycle of instruction ended. Also, we commented on teachers' use of academic words following observations, in essence coaching them to expand their use of the words beyond the CHAAOS scripts. This finding suggests that students who receive special education not only can learn grade appropriate academic words as has been found in prior research, but they can also commit learned word meanings to long-term memory, thus increasing their vocabularies.

With our third research question, we examined whether there were differences in word learning between ELL and NES. We expected smaller gains for ELL due to the significant pretest differences in which ELL scored 2/3 SD lower than native English speakers as measured by the WASI-II and WJ-III. However, we found no interaction between treatment and ELL status, indicating that ELL and NES benefitted similarly from the intervention. This finding may again be a result of the design of the lessons, which included multiple supports to help ELL learn, rehearse, and use taught vocabulary. Specifically, lessons included daily oral discussion of words and contexts, and nearly all participating ELL had sufficient conversational English to participate in these discussions. These supports are similar to those provided to ELL with unknown disability status in prior research demonstrating positive intervention effects (Carlo et al., 2004; Lesaux et al., 2010; Lawrence et al., 2015). Taken together, these previous findings and the results of this study suggest that teachers should incorporate explicit, systematic instruction in high-utility academic words in the daily plans, and should incorporate modifications (e.g., fewer words; accessible contexts) to support access and learning of students with disabilities and ELL.

In our final question, we explored the fidelity of implementation of the lessons by classroom teachers. Importantly, fidelity data suggested strong implementation of instructional components and appropriate time spent on instruction. The support offered to teachers, including modeling of lessons, opportunities for reflection and feedback, and coaching likely contributed to teachers' ability to adhere to the instructional expectations.

Another potential reason why fidelity to implementation was high may be clear alignment of the teacher manuals to the lesson (e.g., Power Point) and student materials (e.g., student notebooks). In our conversations with teachers and administrators prior to the study, we learned that vocabulary instruction varied widely across teachers and that there were no clear parameters in this district for teaching academic vocabulary in special education ELA. Indeed, the few observations in BAU classrooms revealed vocabulary assignments based on the particular text that week, but sparse vocabulary instruction, although BAU instruction included more vocabulary-focused activity than found by Cunningham et al. (2009). Conversely, the vocabulary words in the intervention were drawn from researched lists of academic vocabulary common across conditions. Administration in this district had offered workshops the previous year devoted to teaching within the CCSS, which made it more likely that our target words would align with utility for learning content (free from any particular textbooks) and with the CCSS. It is possible that the teachers implementing treatment recognized the approach as advantageous over existing approaches to teaching vocabulary at their school. This perception may have strengthened teachers' investment in implementing instruction as designed.

Limitations

The positive effects of the intervention must be considered in light of several limitations. First by design, instruction was provided in the context of intact, special education ELA classes taught by a special education teacher. Teachers committed to teaching the lessons four days per week for at least 15 minutes each day. Because the instructional contexts of special education vary widely from self-contained settings, to pull-out resource support, to full inclusion with embedded support, delivery of lessons as prescribed may be difficult to emulate in other settings. Future research should explore the effectiveness of the lessons in different special education contexts to ascertain the robustness of the approach.

Second, the intensity of teacher participation in lesson implementation and redesign—especially in Cycle 1-- necessitated a small sample of teachers and their students for intervention and BAU groups. Accordingly, it was not possible to investigate effects in a multi-level model that considered the nesting of students in classrooms and potential teacher effects. Although classroom-level differences were not found on reading or vocabulary measures prior to treatment, it is possible that some effects of the intervention were due to classroom level differences that were not measured in the current study.

The third limitation concerns limited documentation of vocabulary instruction in BAU classrooms, which revealed a maximum of 9 minutes dedicated to vocabulary and no direct instruction on word meanings during these observations. Given funding constraints, we did not offer an alternative vocabulary approach in the special education BAU. More frequent observation of BAU conditions would have resulted in a more comprehensive picture of the instruction that occurred in BAU, and a clearer context in which to situate the results of the current study.

A final limitation—as in most studies of vocabulary intervention--involves the reliance on researcher-designed measures to document effects of the intervention. Such reliance precludes a conclusion that learning taught vocabulary words results in stronger general word knowledge. However, we did not expect such generalization based on prior research demonstrating non-significant effects on generalized measures of vocabulary (e.g., Lawrence et al., 2015; Lesaux et al., 2010). Additionally, we did not design instruction to address generalization to untaught words (e.g., as instruction in morphemes might). Nevertheless, the lasting effect of instruction on taught words is meaningful given the rarity of maintenance measures, the generally poor English academic vocabulary knowledge of students with LD and ELL, and the scarcity of research addressing systematic academic vocabulary instruction for these populations.

Implications for Research and Practice

Given the results of this study, we suggest implications for practice and future research. First, special educators who deliver ELA instruction to students with LD and other mild/moderate disabilities, including those who are also ELL, should consider embedding systematic instruction in academic vocabulary into their daily instructional routines. Providing multiple opportunities for students to practice and discuss word meanings, use words in conversation and writing, and apply word meanings across contexts can lead to improved knowledge and retention of taught words.

The impact of this learning can be strengthened by selecting for instruction high-utility words that are widely represented in texts and on published lists of useful academic words for instruction. Our research team (Author) has implemented similar instruction in social studies classes for students with and without disabilities; however, in those studies when we reduced the focus on academic vocabulary from 15 minutes to 7 minutes to accommodate social studies content instruction, students learned smaller percentages of the taught words. Future research efforts should continue to involve teachers in the implementation and revision or development of instructional materials. Such teacher involvement has the potential to improve feasibility of designed lessons, which may result in stronger implementation and the teacher buy-in that is necessary for instituting sustained improvement in classroom practice.

References

Anderson, R., Wilson, P., & Fielding, L. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, *23*, 285–303.

Apthorp, H., Randel, B., Chersasaro, T., Clark, T., McKeown, M., et al. (2012). Effects of a supplemental vocabulary program on word knowledge and passage comprehension. *Journal of Educational Effectiveness*, *5*, 160-188.

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school (NCEE 2014-4012)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://ies.ed.gov/ncee/wwc/publications_reviews.aspx

Biemiller, A. (2010). *Words Worth Teaching*. Columbus, OH: SRA/McGraw-Hill. Chall, Jeanne S. Literacy: Trends and Explanations. (1983). *Educational Researcher,* 12, 3-8.

Bulgren, J.A., Graner, P.S., & Deshler, D.D. (2013). Literacy challenges and opportunities for students with learning disabilities in social studies and history. *Learning Disabilities Research & Practice*, *28*, 17-27.

Bunch, G.C., Walqui, A., & Pearson, P.D. (2014). Complex text and new common standards in the United States: Pedagogical implications for English learners. *TESOL Quarterly, 48*, 533-559. Doi: 10.1002/tesq.175

Carlo, M., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., Lively, T., & White, C. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*, 188-215.

doi:http://dx.doi.org/10.1598/RRQ.39.2.3

Chall, Jeanne S. Literacy: Trends and Explanations. (1983). *Educational Researcher,* 12, 3-8.

Cipielewski, J. & Stanovich, K.E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54*, 74-89.

Coleman, R., & Goldenberg, C. (2010). What does research have to say about effective practices for English Learners? *Kappa Delta Pi Record*, Winter, 60-65.

Common Core State Standards Initiative (CCSSI, 2017). *California Common Core State Standards*. Retrieved July 18, 2018, from https://www.cde.ca.gov/be/st/ss/documents/finalelaccssstandards.pdf

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Cunningham, A.E., Zibulsky, J., Stanovich, K.E., and Stanovich, P. (2009). How teachers would spend their time teaching language arts. *Journal of Learning Disabilities, 42*, 418-430. Doi:10.1177/0022219409339063

Echols, L.D., West, R.F., Stanovich, K.E., & Zehr, K.S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88*, 296-304. Doi:10.1037/0022-0663.88.2.296

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

Engelmann, S., Johnson, G., Carnine, L., Meyer, L., Becker, W., & Eisele, J. (1999). *Corrective Reading Decoding Strategies*. Chicago: SRA, MacMillan/McGraw-Hill.

Grierson, A.L. (2013). Walking the walk: Supporting teachers' growth with differentiated professional learning. *Professional Development in Education, 39,* 401-419.

Hall, C., Roberts, G.J., Cho, E., McCulley, L.V., Carroll, M., & Vaughn, S. (2017). Reading instruction for English learners in the middle grades: A meta-analysis. *Educational Psychology Review, 29*, 763-794. Doi: 10.1007/s10648-016-9372-4

Hiebert, E., Goodwin, A., & Cervetti, G.N. (2018). Core vocabulary: Its morphological content and presence in exemplar texts. *Reading Research Quarterly, 53*, 29-49. Doi:10.1002/rrq.183

Hughes, C.A., Morris, J.R., Therrien, W.J., & Benson, S.K. (2017). Explicit Instruction: Historical and contemporary contexts. *Learning Disabilities Research and Practice, 32,* 140-148. Doi: 10.1111/ldrp.12142.

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., and Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide* (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc.

LaRusso, M., Kim, H.Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., Donovan, S., & Snow, C. (2016). Contributions of academic language perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research on Educational Effectiveness, 9*, 201-222. doi: 10.1080/19345747.2015.1116035

Lawrence, J.F., Crosson, A.C., Pare-Blagoev, E.J., & Snow, C.E. (2015). Word generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal, 52*, 750-786. Doi: 10.3102/0002830215579485

Lesaux, N., Kieffer, M.J., Faller, S.E., & Kelley, J.G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly, 45*, 196-228. Doi.org/10.1598/RRQ.45.2.3

Lesaux, N., Kieffer, M.J., Kelley, J.G., & Harris, J.R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, *51*, 1159-1194. Doi: 10.3102/0002831214532165

Lowry, L. (1993). *The Giver*. New York: Dell Laurel Leaf.

McKeown, M.G., Crosson, A.C., Moore, D.W., & Beck, I.L. (2018). Word knowledge and comprehension effects of an academic vocabulary intervention for middle school students. *American Educational Research Journal*, *55*, 572-616. Doi: 10.3102/0002831217744181

Miciak, J., Taylor, W.P., Stuebing, K.K., Fletcher, J.M., & Vaughn, S. (2016). Designing intervention studies: Selected populations, range restrictions, and statistical power. *Journal of Research on Educational Effectiveness, 9*, 556-569. Doi: 10.1080/19345747.2015.1086916

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

National Assessment of Educational Progress (2017). *The Nation's Report Card*. Retrieved July 18, 2018, from https://www.nationsreportcard.gov/

O'Connor, R.E., Sanchez, V., Beach, K., Bocian, K. (2017). Special education teachers integrating reading with 8th grade U.S. History. *Learning Disabilities Research & Practice, 32*, 99-111.

Perfetti, C. & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*, 22-37. Doi: 10.1080/10888438.2013.827687

Sachar, L. (1998). *Holes*. New York: Penguin Random House.

Scammaca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J. Reutebuch, C.K., et al. (2007) *Interventions for adolescent struggling reading: A meta-analysis with implications for practice*. Portsmouth, NH: Center on Instruction. RMC Research.

Swanson, E., Wanzek, J., Vaughn, S., Roberts, G., & Fall, A. (2015). Improving reading comprehension and social studies knowledge among middle school students with disabilities. *Exceptional Children, 81*, 426-442. Doi: 10.1177/0014402914563407

Swanson, H.L., & Deshler, D. (2003). Instructing adolescents with learning disabilities: Converting a meta-analysis to practice. *Journal of Learning Disabilities*, *36*, 124-135.

Vaughn, S., Martinez, L.R., & Wanzek, J. (2017). Improving content knowledge and comprehension for English Language learners: Findings from a randomized field trial. *Journal of Educational Psychology, 109*, 22-34. Doi:10.1037/edu0000069

Wanzek, J., Swanson, E., Vaughn, S., Roberts, G., & Fall, A. (2016). English learner and non-English learner students with disabilities: Content acquisition and comprehension. *Exceptional Children, 82*, 428-442. Doi: 10.1177/0014402915619419

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence-Second Edition*. Bloomington, MN: NCS Pearson, Inc.

Woodcock, R. W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Rolling Meadows, IL: Riverside.

Table 1

Demographic Data by School Site

	Business as Usual		Treatment Classes				Total	
Variable	(N = 20)		(N = 15)		(N = 17)		(N = 52)	
	%	n	%	n	%	n	%	n
Gender								
Male	65.0%	13	93.3%	14	82.4%	14	78.8%	41
Female	35.0%	7	06.7%	1	17.6%	3	21.2%	11
Special Education Classification								
SLD	75.0%	15	86.7%	13	88.2%	15	82.7%	43
SLI	05.0%	1	N/A		N/A		01.9%	1
Autism	10.0%	2	06.7%	1	N/A		05.8%	3
OHI	10.0%	2	06.7%	1	11.8%	2	09.6%	5
Ethnicity								
Hispanic	80.0%	16	67.7%	10	88.2%	15	78.8%	41
White	15.0%	3	26.7%	4	05.9%	1	15.4%	8
Black	05.0%	1	06.7%	1	05.9%	1	05.8%	3
Language Preference								
English	35.0%	7	46.7%	7	41.2%	7	40.4%	21
Spanish	65.0%	13	53.3%	8	58.8%	10	59.6%	31
Number of ELL Students at each CELDT Level:								
CELDT One	2		1		2			
CELDT Two	2		2		3			

CELDT Three	6	3	3
CELDT Four	2	2	1
CELDT Five	1	0	1

Note: SLD = Specific Learning Disability, SLI = Speech/Language Impairment, OHI = Other Health Impairment, CELDT = California English Language Development Test (administered only to students who are ELL) with scores ranging from One (Beginner) to Five (advanced).

Table 2

Scores on Measures by Treatment Condition

Measure	Treatment	Control
WASI.FSIQ	79.81 (15.92)	77.92 (12.36)
WJ-III Total Reading	62.77 (14.97)	60.89 (14.59)
Cycle 1		
Multiple Choice Pretest (16)	6.25 (3.17)	5.22 (2.05)
Multiple Choice Posttest (16)	13.83 (3.34)	4.71 (3.32)
Usage Pretest (8)	2.35 (1.80)	1.90 (1.70)
Usage Posttest (8)	4.08 (1.99)	2.33 (1.46)
Cycle 2		
Multiple Choice Pretest (16)	7.90 (2.66)	5.68 (2.19)
Multiple Choice Posttest (16)	14.10 (2.82)	6.33 (2.60)
Usage Pretest (8)	3.85 (1.46)	3.68 (1.60)
Usage Posttest (8)	4.71 (2.22)	3.04 (1.51)
Cycle 3		
Multiple Choice Pretest (16)	9.25 (3.86)	6.04 (2.96)
Multiple Choice Posttest (16)	14.29 (2.67)	6.85 (3.85)
Usage Pretest (8)	4.85 (2.12)	3.48 (1.59)
Usage Posttest (8)	6.48 (1.40)	4.19 (1.88)

Note: WASI.FSIQ = Wechsler Abbreviated Scale of Intelligence. WJ-III = Woodcock-Johnson Tests of Achievement, Total Reading. Total points possible for each vocabulary measure are shown in parentheses following the test name.

Table 3

ANCOVA Summary Table for Multiple Choice Test, Cycles 1-3

Source	*SS*	*Df*	*MS*	*F*	*P*	*Ƞ2*
Cycle 1						
Pretest	94.84	1	94.84	10.81	.002	.24
WASI	.052	1	.052	.006	.939	.00
Treatment	638.18	1	638.18	72.71	.000	.68
Error	307.20	35	8.78			
Total	4509.00	39				
Cycle 2						
Pretest	55.63	1	55.63	11.95	.001	.24
WASI	4.13	1	4.13	.888	.352	.02
Treatment	319.96	1	319.96	68.75	.000	.64
Error		39	4.65			
Total		43				
Cycle 3						
Pretest	93.72	1	93.72	12.53	.001	.25
WASI	2.81	1	2.81	.38	.543	.01
Treatment	331.13	1	331.13	44.26	.000	.54
Error	284.27	38	7.48			
Total	5335.00	42				

Note: WASI = Wechsler Abbreviated Scale of Intelligence. Source: *SS* = Sum of squares, *Df* = Degrees of freedom, *MS* = Mean square, *F* = F-ratio, *P* = significance level, *Ƞ2 = Effect size.*

Table 4

ANCOVA Summary Table for Usage of Vocabulary, Cycles 1-3

Source	*SS*	*Df*	*MS*	*F*	*P*	*Ƞ2*
Cycle 1						
Pretest	37.19	1	37.19	18.87	.000	.36
WASI	.40	1	.40	.204	.655	.01
Treatment	7.25	1	7.25	3.68	.063	.10
Error		34	66.99			
Total		38	498.00			
Cycle 2						
Pretest	1.38	1	1.38	.52	.475	.01
WASI	28.54	1	28.54	10.78	.002	.22
Treatment	20.74	1	20.74	7.83	.008	.17
Error	103.29	39	2.65			
Total	767.00	43				
Cycle 3						
Pretest	9.45	1	9.45	4.86	.034	.11
WASI	22.09	1	22.09	11.36	.002	.23
Treatment	22.45	1	22.45	11.54	.002	.23
Error	73.94	38	1.95			
Total	1306.00	42				

Note: WASI = Wechsler Abbreviated Scale of Intelligence. Source: *SS* = Sum of squares, *Df* = Degrees of freedom, *MS* = Mean square, *F* = F-ratio, *P* = significance level, *Ƞ2 = Effect size.*