Academic vocabulary and middle school English learners: an intervention study

Dianna Townsend · Penny Collins

Abstract   The goal of this experimental intervention study was to determine if evidence-based instructional strategies for general vocabulary words are effective with middle school English learner (EL) students and academic vocabulary words. Participants showed significantly more growth in their knowledge of academic vocabulary during the treatment condition than during the control condition. A secondary goal of this study was to examine the predictive utility of students' English language proficiency, and students' general vocabulary knowledge in English was a positive predictor for their academic vocabulary growth during the intervention. However, participants' growth during the control period had the greatest predictive utility for their growth during the intervention. Furthermore, this relationship was negative, suggesting that the intervention had the greatest benefits for students who made the least progress in English vocabulary in the absence of the intervention. Implications for instruction, policy, and future research are presented.

Keywords   Academic English · Adolescent · English learners · Middle school · Vocabulary

Introduction

The development of vocabulary in English makes an essential contribution to the academic achievement of English learners (EL); it is also a highly complex process that can vary widely among all groups of students. EL characteristically have less

D. Townsend (✉)
University of Nevada, Reno, NV, USA
e-mail: dtownsend@unr.edu

P. Collins
University of California, Irvine, CA, USA

English vocabulary knowledge than their native English speaking (NS) peers (Geva, Yaghoub-Zadeh, & Schuster, 2000), simply because of less experience with English, which can hinder reading comprehension in English. Because EL students face the dual task of developing proficiency in English while they are developing critical literacy skills, their test scores are often below state averages. For example, EL students obtain reading scores that are below the mean scores for the total representative samples tested in 2005 in the fourth, eighth, and twelfth grades in 2005 (Institute of Education Sciences, 2007). The consequences of this achievement gap include the higher drop-out rates of minority and language minority teenagers (Snow & Biancarosa, 2003). The purpose of the current research is to determine the efficacy of an after-school intervention designed to accelerate the academic vocabulary development of middle school EL students. Following is a synthesis of the relevant literature on vocabulary knowledge, academic vocabulary and academic English, and vocabulary instruction and interventions with diverse populations of students.

Vocabulary knowledge and development

Generally, there is agreement that word knowledge is representative of knowledge in general, although there has been some divergence regarding the degrees to which words are known. For example, Kame'enui, Dixon, and Carnine (1987) have suggested that receptive vocabulary knowledge may be known in the form of verbal association knowledge, partial concept knowledge, or full concept knowledge. Other conceptualizations of word knowledge include Dale's (1965) four stages of self-evaluation of word knowledge, examinations of the incremental nature of building word knowledge (Nagy & Scott, 2000; Wesche & Paribakht, 2000), and Graves' (1986) characterization of the kinds of word-learning tasks that correspond to degrees of word knowledge. Furthermore, vocabulary knowledge may be considered in terms of breadth, which is the size of one's lexicon, or depth, which refers to the quality and richness of semantic representations of those lexical entries (e.g., Ouellette, 2006).

Academic vocabulary and academic English

Academic vocabulary is one class of vocabulary that poses particular challenges due to its complex and often abstract nature. Academic vocabulary is a component of academic English, a register of English used in academic settings and in academic texts, and it is critical for academic success (Corson, 1997; Cunningham & Moore, 1993; Nation & Kyongho, 1995; Scarcella, 2003). Academic vocabulary has recently received a great deal of attention in the literature on word knowledge. For example, Beck, McKeown, and Kucan's (2002) tiers of word knowledge include Tier 3 words, or discipline-specific, low-frequency words. In addition, Tier 2 words, which are high frequency words used across disciplines, have some overlap with general (as opposed to discipline specific) academic vocabulary words. Coxhead's (2000) Academic Word List (AWL) is an important contribution that identifies general, or cross-disciplinary, academic vocabulary words. The AWL contains the 570 most frequent general academic words found in academic texts. The AWL

accounts for roughly 10% of all words in academic texts, and about 12.6 words per page (Coxhead & Nation, 2001). Academic vocabulary may be particularly challenging for students to both comprehend and use as they are mostly Graeco-Latin in origin and are "usually non-concrete, low in imagery, low in frequency, and semantically opaque" (Corson, 1997, p. 696). Additionally, Corson (1997) has demonstrated that knowledge of this language is primarily accessed through texts, not conversation. Thus, students who have gained strong conversational skills in English but who lack extensive print exposure to academic texts will likely not have the vocabulary resources for academic reading comprehension. In short, general academic vocabulary words pose unique challenges to students due to the abstract and often opaque nature of the words, yet students need to develop knowledge of these words, particularly through print exposure, in order to access academic texts and discourse.

A number of recent studies on academic English has been conducted within a functional grammar framework (Bruna, Vann, & Escudero, 2007; Coxhead & Byrd, 2007; Mohan & Slater, 2006; Schleppegrell & Go, 2007; Spycher, 2007; Zwiers, 2006). A functional grammar perspective allows for a close look at the grammatical structures of academic English (i.e. noun phrases, nominalizations, lexical density, clause-connectors). Such a perspective would discourage isolating a linguistic feature, in this case lexical items, for the purposes of improving overall academic English proficiency. However, based on the established body of research demonstrating the strong relationship between vocabulary knowledge and reading comprehension (Biemiller, 1999), there is clear rationale for isolating lexical items as an instructional objective.

English learners and vocabulary development

English learners are at particular risk for struggling with academic vocabulary. Unlike EL students' phonological and orthographic processing skills, which develop similarly to that of NS students (Chiappe, Siegel, & Wade-Woolley, 2002; Geva et al., 2000), EL students' semantic knowledge of English is often less developed than that of their NS peers (Bialystok, Luk, & Kwan, 2005; Biemiller, 1999; Droop & Verhoeven, 2003; Geva et al., 2000). To assist EL students in learning basic English skills and content, ELD instruction typically involves the use of short, simple sentences with accessible vocabulary (Guerrero, 2004). While this type of instruction may be necessary, EL students may not be receiving exposure to the academic English they will need to be successful as they transition to mainstream classes. Additionally, EL students need several years, at least, to fully develop mastery of academic English (Cummins, 1981; Hakuta, Butler, & Witt, 2000). Thus, due to less experience with English than their NS peers and to the lack of academic English exposure they receive prior to entering mainstream classrooms, EL students face numerous challenges when they encounter academic English in texts. The current study was designed to address this issue by giving middle school EL students meaningful exposures to and opportunities to practice using academic vocabulary words in an effort to accelerate their academic vocabulary development. The rationale for the instructional strategies used in the intervention for this study

comes directly from the vocabulary literature. Thus, the following section provides an overview of the research on vocabulary instruction and interventions.

Vocabulary instruction with English learners

While academic vocabulary refers to a specific subset of words, there is much to be learned from the greater literature on instruction of vocabulary in general. There is considerable support for the efficacy of rich vocabulary instruction, or a combination of evidence-based strategies including direct instruction, providing multiple opportunities to words in multiple contexts, and engaging students in active practice and personalization of word meanings (Beck et al., 2002; Beck, Perfetti, & McKeown, 1982; McKeown & Beck, 2004; McKeown, Beck, Omanson, & Perfetti, 1983; NICHD, 2000; Stahl & Fairbanks, 1986). Although rich vocabulary instruction may be effective within one's native language, one must also address the unique challenges faced by students when learning vocabulary in their L2. Such challenges include limited instructional time (August, Carlo, Dressler, & Snow, 2005), the need to teach abstract words that are not readily supported by graphics or pictures (Anderson & Roit, 1996), and teachers' tendency to overestimate EL students' knowledge of written English vocabulary because of their greater mastery of conversational English (Beck et al., 2002; Scarcella, 2003). Despite these additional challenges, there is growing evidence that the same techniques identified for NS students may benefit EL (Shanahan & Beck, 2006). However, these strategies should be supplemented with many visual aids and increased opportunities to practice using new words in various contexts (August et al., 2005). Furthermore, reading aloud from appropriate texts has also been found to be effective in building EL students' vocabulary (Coady, 1997; Houk, 2005; Stahl, Richek, & Vandevier, 1990).

Vocabulary interventions with English learners

The majority of empirical studies on vocabulary interventions have been conducted with monolingual English speakers. The National Reading Panel (2000) did not review studies based on EL or at-risk populations, and few empirical vocabulary intervention studies have been conducted with EL students (Calderon et al., 2005). Indeed, in a recent report on language minority students, Shanahan and Beck (2006) found only three experimental studies involving EL. However, there is some evidence that, at least in the elementary grades, vocabulary interventions with EL students can make a difference. In a study modeled after Beck and McKeown's rich vocabulary instruction approach, Calderon et al. (2005) found modest but significant effects for their intervention with Spanish-speaking EL students in grade three. The components of the intervention were pre-teaching vocabulary, developing vocabulary through discourse around text, and using oral language activities to build vocabulary; the effect sizes on the literacy outcome measures ranged from .11 to .21. Thus, the rich vocabulary instruction previously shown to improve the vocabulary knowledge and reading comprehension of native English-speakers was also effective with EL.

In a series of two experiments, Carlo, August, and Snow (2005) found longitudinal effects for a vocabulary intervention emphasizing deep and rich

processing of words for Spanish-speaking EL students. They conducted two studies using the Vocabulary Improvement Project, which consisted of a curriculum, an instructional routine, and professional development. The intervention promoted significant growth on almost all literacy measures, including target word mastery, knowledge of polysemy, depth of word knowledge, and reading comprehension.

Rationale and research questions

As with the intervention research on NS students, interventions that include both direct instruction and multiple opportunities for processing words in different contexts led to gains in vocabulary knowledge among EL. In addition, interventions that were designed with principles of rich vocabulary instruction resulted in reading comprehension gains for EL students. These findings are promising; evidently, there are principles of instruction and strategies that will support EL students as they attend to the dual task of learning English and learning the curriculum.

However, most of the intervention research for EL students has taken place with elementary school students. Middle school EL students, in particular, are a population of students with distinct needs, and they are under-represented in the body of empirical vocabulary research (Snow, 2006). Also, as Klinger and Vaughn (2004) assert, the handful of intervention studies on adolescent EL are not sufficient to guide significant improvements in practice. Additionally, the aforementioned evidence-based vocabulary strategies have not been empirically evaluated with general academic vocabulary words, which are abstract and challenging to learn. Thus, the current study is concerned with applying evidence-based strategies in a new context: adolescent EL students learning general academic vocabulary words. Finally, because EL students are at widely differing degrees of English development, this study also addresses the issue of "readiness to learn", or the extent to which there is a threshold of English proficiency necessary for students to learn general academic vocabulary words. The current study is driven by these issues and the following questions:

1. Can an after-school, evidence-based academic vocabulary development intervention increase the academic vocabulary knowledge of middle school EL?
2. To what extent does English learners' proficiency in English mediate their response to a vocabulary development intervention?

Method

Participants and setting

The participants, 52 EL, were recruited from English language development (ELD) classes from one middle school in Southern California. Students were enrolled in ELD classes based on their performance levels on the California ELD Test (CTB/McGraw-Hill, 2005). The mean age of the students was 12 years, 11 months and the range of the participants' ages was 11 years to 15 years, 2 months. All students who

returned parental consent forms were evenly and randomly assigned to the two treatment groups (A and B). Not all of the 52 students regularly attended the intervention, Language Workshop. Students who attended five or fewer days of the 20 days of the program were considered attrition students and were excluded from the analyses. The final number of students included in analyses was 37 students, with 20 students in group A and 17 students in group B. Table 1 presents the breakdown of the two groups and illustrates that there were no statistically significant differences between the two groups.

The middle school that served as the research site for this study is located in a suburban school district, and it is a California Distinguished School. The school serves roughly 900 students that make up an ethnically and linguistically diverse student population. In a four-block day, the EL students participated in two blocks

Table 1 Demographic information

	Language Workshop session A ($N = 20$)	Language Workshop session B ($N = 17$)	Pearson Chi-squared value
Gender			$\chi^2 (1) = .62$, *ns*
Male	8 (40%)	9 (53%)	
Female	12 (60%)	8 (47%)	
CELDT[a] level			$\chi^2 (4) = 3.59$, *ns*
Beginning	5 (25%)	3 (18%)	
Early	2 (10%)	3 (18%)	
Intermediate	6 (30%)	6 (35%)	
Early advanced	4 (20%)	5 (29%)	
Advanced	3 (15%)	0 (0%)	
Grade			$\chi^2 (2) = 0.20$, *ns*
6th grade	8 (40%)	8 (47%)	
7th grade	7 (35%)	5 (29%)	
8th grade	5 (25%)	4 (24%)	
First language			$\chi^2 (4) = 6.51$, *ns*
Spanish	12 (60%)	13 (76%)	
Vietnamese	4 (20%)	2 (12%)	
Japanese	3 (15%)	0	
Arabic	0	2 (12%)	
Gujarati	1 (5%)	0	
ELD[b] class			$\chi^2 (3) = .359$, *ns*
Level Basic	6 (30%)	4 (24%)	
Level A	8 (40%)	8 (47%)	
Level B	3 (15%)	2 (12%)	
Other interventions[c]	3 (15%)	3 (18%)	

[a] California English Language Development Test

[b] English Language Development

[c] Students who graduated out of Level B ELD, but who were not ready for the mainstream curriculum, graduated into other literacy intervention programs such as Language! and Read 180

of English language and literacy instruction. Their other classes included math, physical education, and electives.

Design

An experimental research design with two treatment groups was used for this study. Each group served as its own and the other's control group, allowing for both between and within groups analyses. This research design involved three testing phases for both groups, titled the T1, T2, and T3 phases. Participants were tested on standardized language and literacy measures, as well as other published measures of language proficiency. All students were tested in December of 2006, the T1 phase of testing. From January through the middle of February, group A participated in the treatment after school. In March, all study participants were retested, and this was the T2 testing phase. From March through the beginning of April, group B participated in the treatment. Finally, all participants were tested once again at the end of April, the T3 testing phase.

Treatment—procedures

The treatment, Language Workshop, was designed as an after-school program to accelerate the academic vocabulary development of middle school EL students. The duration of the intervention was 20 sessions, or 5 weeks for 4 days a week, with each session lasting approximately 75 min. The instructional context was comprised of a classroom setting with the first author as the intervention teacher. On most days, an instructional aide was available. Each day followed a similar routine, starting with a snack and a hook activity with the three to four target words for that day. The instructional texts and passages related to two standards-based topics (*California State Board of Education Standards and Frameworks,* 2004), the history of inventions and space and the solar system. The two instructional texts, *Great Inventions* (Wood, 1995) and *Stars and Planets* (Levy, 2003), are part of a series of large discovery books with rich graphics and short passages of academic writing.

Both groups received the same 5 week intervention during two different 5 week sessions in the winter and spring. Also, 2 days in the second session had to be cancelled; thus, group B students only had the opportunity to attend 18 days of the intervention. While one group received the intervention, students in the other group were encouraged to attend the school's after-school homework club regularly. In addition, some students were involved in other after-school programs, including a reading fluency intervention and a college preparatory program. However, attendance at these voluntary programs was minimal and sporadic, and subsequent analyses showed that these programs had no effect on the study measures.

Treatment—instructional strategies

The goals of *Language Workshop* were explicitly related to the aforementioned principles of rich vocabulary instruction advocated by McKeown and Beck (2004). According to these leading researchers, students will best learn words in a rich

verbal environment in which they experience direct instruction of target words, multiple exposures to target words in multiple contexts, and many opportunities to use and personalize word meanings. McKeown and Beck (2004) explain: "Rich instruction is particularly important for words that seem necessary for comprehension, or for words that turn up in a wide variety of contexts, or for words that are hard to get across with just a brief explanation" (p. 18). In other words, rich vocabulary instruction is a potentially strong match for general academic vocabulary words, which show up across disciplines and have abstract definitions. The principles of vocabulary instruction advocated by McKeown and Beck are echoed throughout a great deal of the vocabulary literature for both NS and EL students (August et al., 2005; Blachowicz & Fisher, 2000; Corson, 1997; Jimenez, Garcia, & Pearson, 1996; Scarcella, 2003; Stahl & Nagy, 2006). Informed by this research, the instructional practices of Language Workshop reflected two main goals. The first goal, building depth of academic word knowledge, was operationalized by providing opportunities to actively process word meanings, opportunities to personalize word meanings, and exposures to multiple contexts for the target words. The second goal was to build general breadth of word knowledge in a language-rich environment. Strategies to meet these two goals of the intervention were used each session. All direct instruction, games, and activities provided students with multiple exposures to the target academic words in multiple contexts, background knowledge on additional vocabulary words, and multiple opportunities to practice with and personalize word meanings.

Specifically, daily direct instruction and discussion of the three or four target words involved using large cards with the words, definitions, sentences containing the words, and sentences with the target words missing, as well as supporting pictures. Students regularly participated in matching games, in which students received a small card with a word, its definition, or sentence with a target word missing. Students were required to find peers with related cards, and in small groups, write a sentence, draw a picture, or design a short skit to illustrate the word. The rest of the class would then guess each group's word, and misconceptions were addressed by peers and the instructor. Other instructional activities included shared reading of passages from the discovery books that used the target words or provided useful contexts for discussing the target words, and modified versions of games such as Taboo and Pictionary, which enabled students to make connections between the target words and other words from the discovery texts and their own prior knowledge. Additionally, high-interest novels were read aloud to the students throughout the intervention and students were able to choose two of the novels to keep and continue reading on their own. Both the games and the read alouds contributed to the second goal, which was to increase general breadth of vocabulary knowledge. Appendix A presents a sample week from Language Workshop.

Treatment—target words

Table 2 presents the target academic words for the intervention, the 60 most frequent words on the Academic Word List (Coxhead, 2000). Twelve new academic words were introduced for each of the 5 weeks. Instructional texts and activities

Table 2 Target academic words for Language Workshop[a]

analysis	contract	factors	legal	research
approach	create	financial	legislation	response
area	data	formula	major	role
assessment	definition	function	method	section
assume	derived	identified	occur	sector
authority	distribution	income	percent	significant
available	economic	indicate	period	similar
benefit	environment	individual	policy	source
concept	established	interpretation	principle	specific
consistent	estimate	involved	procedure	structure
constitutional	evidence	issues	process	theory
context	export	labour	required	variables

[a] These words are the 60 most frequent words on Coxhead's (2000) academic word list

gave students multiple opportunities to practice using the target academic words. In addition to the target academic words, high-utility words used across both academic and non-academic settings received some instructional focus. These words were generally Tier 2 words, which are considered to be the best candidates for vocabulary instruction (Beck et al., 2002). The purpose of using some instructional time for these words was related to the second goal of the intervention, which was to promote general breadth of vocabulary knowledge, and to provide background knowledge for the contexts used to practice the target academic words.

Fidelity of implementation and tracking exposures to words

To insure fidelity of implementation of the intervention, a specific lesson plan was designed for each of the twenty sessions. These lesson plans were designed and implemented by the first author; thus, students in group B had the same set of experiences with the target words as students in group A. In addition, the number of exposures to each of the target words was tracked on a daily basis based on the lesson plans. Both the number and nature of word exposures were tracked; the nature of exposures ranged from "instructional mention", in which the word was mentioned briefly in an instructional context, to "active practice" in which the word was a primary focus of a game or other activity in which students were actively involved with using the words. Participants received a minimum of three exposures, with an average of five exposures, to each of the target words over the 20 sessions, with at least half of the exposures falling under the "active practice" category.

Measures

All participating students were individually tested during all three testing phases. The testing occurred in a quiet location on the campus of the school, and the measures included three measures of vocabulary knowledge. The individual testing sessions lasted 30 minutes.

Vocabulary Knowledge Scale (Paribakht & Wesche, 1997). The Vocabulary Knowledge Scale is a generic vocabulary measure that was modified for the current study and has been used previously in an interview format (see Read, 2000, for a review). The modified version of the measure was titled the Vocabulary Knowledge Scale—Measure of Academic Vocabulary (MAV). In an interview setting, students were shown a word and the word was read out loud to them. They were then asked if they have seen or heard the word before. If they had seen or heard it, they were asked if they thought they knew what the word meant and to provide a definition. Participants were then asked to use the word in a sentence. Each student answer was written down verbatim and the test administrator directly moved on to the next item.

Two parallel forms of the MAV, each with twenty items, were designed for this study. Every third word from the top 60 words (the intervention target words) and from the second 60 words in the Academic Word List were selected. From those 40 words, every other word was selected for version A of the measure and the remainder of the items made up version B of the measure. The result was two parallel forms of the MAV that were made up of ten randomly selected words from each of the top two sublists of the Academic Word List. The purpose of including words from the second sublist was to isolate the effects of the intervention from general classroom instruction which may have yielded some academic vocabulary growth. On the MAV, a systematic scoring procedure was used. Each item was awarded a maximum of 5 points, yielding a maximum raw score of 100. An item received a score of 0 if participants said they did not recognize the word. One point was awarded if participants reported they were familiar with the word, but could neither define nor use that word in a sentence. Items that were either partially defined or used somewhat accurately in a sentence were awarded two points. Items were awarded three points if they were accurately defined or if an incomplete definition was accompanied by a sentence in which the word was used partially accurately. Items received four points if an accurate definition was accompanied by a sentence in which the word was used somewhat accurately or if an incomplete definition was accompanied by a sentence in which the target word was used accurately. Finally, an item received the full five points if an accurate definition was accompanied by a sentence in which the target word was used accurately. This test was individually administered. Raw scores were obtained for each student, and the maximum raw score was 100. In a pilot administration of the MAV with a different group of students, the coefficient alphas for forms A and B, respectively, were .82 and .87. The coefficient of equivalence for the two forms was .91.

The *Vocabulary Levels Test* (Schmitt, Schmitt, & Clapham, 2001) was designed to be a diagnostic tool for EL students of all ages and provides information about their levels of both general and academic vocabulary knowledge in English. The levels of vocabulary are based on West's (1953) General Service List, which categorizes in groups of thousands the most frequent words in English. For this study, items drawn from the 2,000–2,999 most frequent words and a version of the academic vocabulary level, modified to include more of the intervention target words, were administered. In this written test, students matched vocabulary words with definitions. To minimize guessing, sets of six target words were presented with

three definitions. This individually-administered test had a maximum raw score of 60 points, or 30 points for each of the two sections administered. Cronbach's alpha for each of the three sections of the original VLT is reported as above .91 (Schmitt et al., 2001). In the current study, the test–retest reliability of the modified form of the VLT, based on group B students' performance at the beginning and end of their control period, was .95. The validity of the VLT as a measure of receptive word knowledge has been established (see Schmitt et al., 2001). Additionally, concurrent validity of the VLT as a measure of general receptive vocabulary knowledge is supported in the current study, as the correlation between the VLT and PPVT-III was strong and significant, $r = .89$, $p < .01$.

The *Peabody Picture Vocabulary Test – III* (PPVT) (Dunn & Dunn, 1997) was used as an index of students' receptive vocabulary in English. In this task, students viewed four pictures and were asked to identify which best illustrated the target word for the item. Raw and percentile scores were obtained for each student. Reliability has been established; test–retest reliabilities fall in the range of .91 to .94, and the alpha coefficient ranges from .89 to .94 for raw scores (Bessai, 2001). Additionally, both construct and concurrent validity of the PPVT-III as a measure of general receptive vocabulary knowledge have been established (Bessai, 2001).

Results

Testing differences at T1, T2, and T3 phases

The means for the two Language Workshop groups, groups A and B, at all three testing phases on all measures are presented in Table 3. Group A participated in Language Workshop between the T1 and T2 phases, and its within-subject control period was between the T2 and T3 phases. Group B participated in Language Workshop between the T2 and T3 phases, and its within-subject control period was between the T1 and the T2 phases.

For the T1 scores, a multivariate analysis of variance (MANOVA) on the vocabulary measures, which included the Peabody Picture Vocabulary Test (PPVT) raw scores, the Measure of Academic Vocabulary (MAV) scores, and the Vocabulary Levels Test (VLT) scores, revealed no statistically significant differences between groups, $F(3,33) < 1$, *ns*.

For the T2 scores, a MANOVA examining student performance on the three vocabulary measures, PPVT raw scores, MAV scores, and VLT scores, revealed no differences on total scores between groups A and B, $F(3,33) < 1$, *ns*. However, because of the small sample size and wide variation of English proficiency among participants, an analysis of covariance (ANCOVA) was conducted to determine if there were significant differences between groups on just the items on the Language Workshop target words (LW-items). In this analysis, treatment group (groups A and B) was the independent variable, performance on the LW-items at T2 was the independent variable, and baseline performance on the total MAV score was the covariate. The analysis did result in a significant difference with group A outperforming group B on the target word items, $F(1,36) = 14.014$, $p < .01$.

Table 3 Means for intervention Groups A and B at three testing periods (standard deviations in parentheses)

Test	T1 Testing phase		T2 Testing phase		T3 Testing phase	
	A	B	A	B	A	B
MAV[a] Total	27.50 (12.76)	26.18 (12.18)	36.60 (19.67)	31.59 (14.18)	34.05 (14.64)	36.53 (15.90)
MAV target word items	17.75 (9.09)	17.00 (6.45)	19.15 (10.71)	12.88 (6.33)	23.50 (10.07)	24.06 (8.72)
MAV non-target word items	9.75 (4.78)	9.35 (6.68)	17.40 (9.19)	18.71 (8.50)	10.65 (5.43)	12.82 (7.25)
VLT[b] Total	26.35 (16.64)	28.59 (13.87)	30.95 (14.04)	31.65 (12.33)	32.55 (15.25)	33.12 (13.03)
VLT level 1	15.75 (10.14)	16.35 (7.83)	18.95 (8.53)	18.76 (7.40)	19.00 (8.75)	20.12 (7.90)
VLT academic language	10.60 (7.13)	12.24 (6.45)	12.15 (5.91)	12.88 (5.83)	13.45 (6.89)	13.59 (6.08)
PPVT raw	98.15 (41.20)	97.65 (36.36)	104.25 (38.09)	101.82 (35.22)	111.05 (36.67)	105.71 (33.69)
PPVT percentile	12.45 (15.10)	9.06 (14.71)	13.60 (16.13)	9.59 (15.63)	18.00 (18.54)	10.47 (14.85)

[a] MAV = Measure of academic vocabulary

[b] VLT = Vocabulary levels test

For the T3 scores, a MANOVA examining student performance on the three vocabulary measures, PPVT raw scores, MAV scores, and VLT scores, revealed no overall group differences, $F(3,33) = 1.86$, *ns*. Similarly to the analyses conducted on the T2 scores, an ANCOVA was conducted with the treatment group as the independent variable, LW-items at T3 as the independent variable, and baseline performance on the total MAV score was the covariate. This analysis did not reveal any significant differences $F(1,36) < 1$, *ns*, which was not surprising since both groups had completed the intervention at this stage.

Can an after school, research-based, academic vocabulary development intervention increase the academic vocabulary knowledge of middle school EL?

To answer the first research question, growth scores were computed for all participants during their respective control and treatment periods. For several analyses, the data from the two groups were collapsed so that the treatment growth for all participants could be compared with the control growth for all participants. Table 4 presents the mean growth of groups A and B, as well as the mean growth of the entire sample, during the treatment and control periods. Following the computation of the mean growth of both groups and the entire sample during the control and treatment periods, a series of repeated measures analyses of variance were conducted to determine significant differences between the control and treatment periods and between groups. For these analyses, the within-groups factor was condition, or treatment versus control period. In additional to statistical significance, practical significance is also reported in these analyses as the measurement of the effect size, *partial η^2*, with .02 indicating a small effect, .15 indicating a medium effect, and .35 indicating a large effect.

Intervention period growth versus control period growth on the MAV

On the MAV total score, the measure that was aligned with the intervention, there were statistically significant within and between group differences. When participants were in the treatment condition, they made significantly more growth than when they were in the control condition, $F(1,35) = 6.09$, $p < .05$, *partial $\eta^2 = 0.15$*, demonstrating the efficacy of the intervention with both significant and practical significance. In addition, there was a between-groups interaction. Students in group A made significantly more growth, both statistically and practically, in the intervention than group B, $F(1,35) = 7.16$, $p < .05$, *partial $\eta^2 = 0.17$*. Following this analysis, the MAV was broken down into an analysis of those items that were aligned with the intervention target words and those that were not aligned with intervention target words. For the target word items, there was a statistically significant within-groups difference with subjects making significantly greater growth, both statistically and practically, during the intervention period than during the control period, $F(1,35) = 19.98$, $p < .001$, *partial $\eta^2 = .36$*. In contrast, for the non-target word items, there was no statistically or practically significant within-groups difference on growth during the treatment versus control periods,

Table 4 Mean growth of Groups A and B, and total sample, during treatment and control periods

Measures	Growth during treatment period			Growth during control period		
	A	B	Total	A	B	Total
Measure of academic vocabulary: total	9.10 (9.76)	4.94 (7.80)	7.19 (9.04)	−2.55 (6.53)	5.41 (5.98)	1.11 (7.39)
Target word items	1.40 (4.56)	11.18 (6.29)	5.89 (7.28)	4.35 (3.75)	−4.12 (3.98)	0.46 (5.72)
Non-target word items	7.65 (7.28)	−5.89 (4.54)	1.43 (9.16)	−6.75 (5.27)	9.35 (5.00)	0.65 (9.84)
Vocabulary levels test: total	4.60 (6.75)	3.06 (5.98)	3.16 (6.52)	4.60 (6.75)	3.06 (5.98)	2.27 (5.58)
Section 1 (general vocabulary)	3.20 (4.86)	1.35 (2.67)	2.35 (4.06)	0.05 (2.56)	2.41 (3.71)	1.14 (3.32)
Section 3 (academic vocabulary)	1.55 (4.29)	0.65 (4.62)	1.16 (4.43)	1.55 (4.29)	0.65 (4.62)	1.00 (3.94)
PPVT						
Raw	6.10 (12.87)	3.88 (12.05)	5.08 (12.37)	6.80 (9.90)	4.18 (9.86)	5.59 (9.83)
Percentile	1.15 (9.49)	0.88 (6.87)	1.03 (8.28)	4.40 (7.54)	0.53 (11.84)	2.62 (9.81)

$F(1,35) = .05$, *ns*. Thus, the intervention was effective in building participants' knowledge of the target words.

Analysis of the experimental measure, the MAV

A series of follow-up analyses revealed a complex pattern of between group effects. Specifically, a repeated measures ANOVA on the growth of both types of items revealed an interaction with condition, group membership, and types of items. Group A showed greater growth on non-target word items than on target word items during the intervention, while group B showed greater growth on target word items than on non-target word items during the intervention period, $F(1,35) = 103.28$, $p < .001$, *partial η^2* = 0.75. The large effect size here suggests that this was both statistically and practically significant. During the control period, the opposite was true for both groups. The two groups each took a different form of the MAV immediately prior to their respective intervention periods. Because there were no statistically significant differences between groups at the T1 phase, and because group B did not receive any treatment prior to the T2 phase, the performances of group A and group B on the two different forms of the MAV were compared. This data suggested that form A had easier target word items and more difficult non-target word items, $t(19) = 5.17$, $p < .001$; while form B had easier non-target word items and more difficult target words items, $t(16) = 4.95$, $p < .001$. To test this trend, the pilot data was re-examined. Pilot students who had form A did significantly better on target word items, $t(13) = 6.09$, $p < .001$; and students with form B did significantly better on non-target word items, $t(26) = 2.27$, $p < .05$. In other words, the two samples of students from the pilot administration of the MAV and the current study showed parallel performances on the two forms of the MAV. This suggests that the two forms, though parallel in overall scores, varied in difficulty with respect to different groups of words.

Overall growth on target word knowledge

To further test the effectiveness of the intervention, and to eliminate error caused by the two different forms of the MAV, the next step in this analysis was to determine the growth participants made on form A, which both groups took during the T1 and T3 testing phases of the study. In these analyses, practical significance is also reported, using Cohen's *d* as a measure of effect size. Unlike *partial η^2*, the guidelines for effect sizes for *d* are: .2 is a small effect, .5 is a medium effect, and .8 is a strong effect. Over the course of the entire study, group A participants made statistically significant growth on the target word items, $t(19) = 6.10$, $p < .001$, $d = .42$; but not on the non-target word items, $t(19) = 1.00$, $d = .18$. Group B participants made statistically significant growth on both the target word items, $t(16) = 6.67$, $p < .001$, $d = .71$; and on the non-target word items, $t(16) = 4.36$, $p < .001$, $d = .50$. While growth on both types of words was statistically significant for group B, the effect size for the target word growth was strong while the effect size for the non-target word growth was moderate. Additionally, the growth that both groups made on the target word items was significantly greater than the growth they made on the non-target word

items, $t_{group\ A}(19) = 5.40, p < .001, d = 1.18; t_{group\ B}(16) = 3.51, p < .01, d = .93$. Thus, all students made significantly more growth on the items that were aligned with the intervention, and the effect sizes show strong practical significance as well.

Finally, when the data on the groups was collapsed to compare intervention period growth to control period growth for the whole sample, a similar pattern was found. All participants made more statistical and practical growth during their respective intervention periods than their respective control periods on the target words, $t(37) = 2.67, p < .05, d = .83$. However, on the non-target words, there was no statistical or practical significant difference between the intervention and control periods, $t(37) = .256$, ns, $d = .08$. These findings further demonstrate the effectiveness of the intervention.

Intervention period growth versus control period growth on the VLT

A repeated measures ANOVA on growth on the VLT revealed no significant differences within groups, $F(1,35) < 1$, ns; or interactions between groups, $F(1,35) = 1.59$, ns. However, there was a statistically significant interaction on the first section, or general vocabulary section, of the VLT between condition and group membership. Group B showed greater growth on this section during the control period, and group A showed greater growth during the treatment period, $F(1,35) = 4.54$, $p < .05$, $partial\ \eta^2 = .12$. There were no statistically significant within-group differences, $F(1,35) < 1$, ns; or between-group interactions, $F(1,35) < 1$, ns; with the academic vocabulary section of the VLT. With regards to the vocabulary measures, no other effects from the intervention period were significant. An a priori power analysis with the alpha level set at .05 and the sensitivity to detect a moderate effect (.4) indicated a power of .53. Thus, it is possible that there was not sufficient power to detect significant differences on the VLT.

Intervention period growth versus control period growth on the PPVT

The T3 phase of this study was, in fact, a delayed post-testing trial for group A. Because group A participants had completed the intervention roughly 6 weeks earlier, the T3 phase allowed for an examination of any enduring effects of the intervention. On the raw scores of the PPVT, group A students made statistically significant growth following the intervention, $t(19) = .307, p < .01$. This suggests that group A's participation in the intervention may have accelerated their general vocabulary development following the intervention, although the small sample size and the small effect size ($d = .18$) would heed against any such conclusion. However, group B showed no significant growth at T2 or T3 on the PPVT, suggesting that the fact that both groups took the same form of the PPVT at all three testing phases was not the cause of group A's growth.

Additional mediating factors on vocabulary growth

A series of analyses of variance examined if other factors, such as first language or intervention attendance rates, influenced students' response to the intervention.

First, an ANOVA revealed that children's response to the intervention, as measured by their MAV scores, did not vary as a function of their home languages. Similarly, an ANOVA on days of attendance, which was recoded into three variables (6–10 days, 11–14 days, and 15–20 days), showed that attendance was not related to growth in the intervention, $F(2,34) < 1$, *ns*. The absence of attendance effects may reflect the limited duration of the study as a whole, and the high attendance rates of the students. Indeed, of the 37 students, 31 students attended Language Workshop at least 15 out of 20 days.

To what extent do language skills in English mediate middle school English learners' vocabulary growth in a vocabulary development intervention?

To answer the second research question, ANOVA, Pearson product moment correlations, and multiple regression models were used. Students' growth during the intervention differed as a function of their English proficiency, as indicated by an ANOVA on students' California English Language Development Tests (CELDT) levels. Growth on the MAV during the intervention increased with each of the five successive CELDT levels, from basic to advanced, $F(4,32) = 4.16$, $p < .01$, *partial* $\eta^2 = .34$. These statistically and practically significant differences were limited to growth during the intervention periods; growth during the control periods showed different patterns for both variables, but those differences were not statistically significant. Next, the Pearson correlations for age, CELDT level, and all T1 phase raw scores were determined. Table 5 presents these results.

Several trends emerged from the correlation analysis. Age tended to correlate negatively with performance on the measures, which might suggest that younger students performed better than older students on these measures, although this pattern was not significant. The CELDT levels and vocabulary measures showed strong and significant correlations with each other, which lends credibility to the MAV and VLT in making valid inferences about students' vocabulary knowledge from these two measures.

Following the correlation analyses, multiple regression models were computed to determine which T1 phase and demographic variables predicted growth during the intervention. The small sample size might suggest against using multiple regression, but the design of the model specifically addressed the second research question and therefore provides information relevant to the study. Because subjects' growth on

Table 5 Pearson correlations on raw scores at T1 phase testing[a]

	Age	CELDT	PPVT	MAV	VLT
Age					
CELDT	−.08				
PPVT	−.21	.83**			
MAV	.04	.83**	.82**		
VLT	−.20	.82**	.89**	.79**	

[a] CELDT = CELDT levels; PPVT = PPVT raw scores; MAV = MAV total scores; VLT = VLT total scores

Table 6 Multiple regression model on MAV intervention growth

Variables	R^2	R^2 Change	Standardized β	p
MAV growth control period	.40	.40	−.476	.000
Age	.47	.08	−.217	.054
PPVT T1 phase raw score	.64	.16	.419	.001

the MAV during their respective intervention periods was significantly different from their growth during their respective control periods, growth on the MAV during the intervention periods was used as the dependent variable. MAV growth during the control period was entered as the first step in the regression model, so that students' growth independent of the intervention could be statistically controlled. Next, age was entered as the second step of the model to account for the variance in ages that may have predicted growth. The patterns of correlations informed the choice of other variables to be included in the model. For example, the degree of correlation between CELDT levels and PPVT raw scores, $r = .83$, $p < .01$, suggested that only one of these variables be used in the multiple regression models. Indeed, preliminary analyses with these two variables showed that they shared considerable common variance in predicting growth during the intervention. However, when entered into a multiple regression model together, their variance was not significant, suggesting a problem of collinearity. When entered individually, CELDT levels explained 31%, $F(1,35) = 15.35$, $p < .000$, of the variance and PPVT levels explained 35%, $F(1,35) = 18.82$, $p < .000$, of the variance. Thus, the PPVT was determined to be a more sensitive predictor and was entered in the subsequent models.

Table 6 summarizes the results of the multiple regression analysis on the MAV intervention growth. On its own, MAV growth during the control period accounted for 40% of the variance of MAV growth during the treatment periods, $F(1,35) = 23.37$, $p < .001$. However, the beta coefficient in this relationship was negative, $\beta = -.48$, $p < .001$, which suggests that students who experienced less growth during the control period experienced more growth during the intervention period. Following the influence of MAV growth during the control period, both age and PPVT raw scores explained statistically significant and unique variance in MAV growth during the intervention, 8% and 16%, respectively. For age, the beta coefficient was negative, $\beta = -.22$, and it approached significance, $p = .054$, suggesting that younger students showed more growth on the MAV during the intervention. The PPVT raw scores, however, yielded a positive beta coefficient $\beta = .42$, $p < .01$, suggesting that children who initially had stronger receptive vocabulary in English showed greater response to the intervention.

Discussion

The results from this study suggest that an academic vocabulary development intervention with research-based instructional strategies can increase the academic

vocabulary knowledge of middle school EL students. The results also suggest that middle school EL students' general vocabulary knowledge can significantly predict their growth in such an intervention. Additionally, this study indicates that EL who make less growth in the absence of an intervention will show more growth during the intervention.

Can an after school, research-based, academic vocabulary development intervention increase the academic vocabulary knowledge of middle school EL?

Participants showed greater growth in their knowledge of academic vocabulary, as measured by the Measure of Academic Vocabulary (MAV), immediately after taking part in the intervention and for words that were taught rather than those that were not taught during the intervention. These findings showed that the instructional strategies previously found to be effective with general vocabulary words (Carlo et al., 2004; McKeown et al., 1983), can also be used to build academic vocabulary for middle school EL. Additionally, despite the differences in the two forms of the MAV, the entire sample showed more growth on the target word items immediately following the intervention. This finding strengthens the primary conclusion of this study: research-based strategies are effective for teaching the words that are the most challenging to learn, academic vocabulary words (Anderson & Roit, 1996; Corson, 1997). Because the research on vocabulary instruction suggests that learning words is not an easy process (Beck & McKeown, 2007; McKeown, Beck, Omanson, & Pople, 1985), the current study's findings are promising.

   As expected, participants did not show significant growth on the PPVT during the intervention period; this was not surprising because the PPVT is a measure of general vocabulary and the intervention focused on 60 academic target words over 5 weeks. Indeed, Carlo et al.'s (2005) 15 week vocabulary instruction program did not produce treatment effects on the PPVT. However, group A participants' apparent growth on the PPVT following the intervention was promising and somewhat unprecedented. In the 6 weeks following the intervention, group A students made significant gains on the PPVT raw and percentile scores. Although other researchers have reported delayed post-testing effects with vocabulary interventions (Bos & Anders, 1990; Nash & Snowling, 2006), these effects were generally revealed by experimental measures aligned with the intervention or instructional procedures in the study. In contrast, Margosein, Pascarella, and Pflaum (1982) found moderate growth on a standardized vocabulary measure, the Gates-MacGinitie Vocabulary subtest, after a 24 day intervention using semantic mapping, but their study did not include delayed post-testing data. Thus, the current study diverges from the literature on two counts; the measure that showed delayed post-testing effects (PPVT) was a standardized measure of general vocabulary, and the students showed no significant gains on this measure during the actual intervention. Because of the small sample size and small effect size, group A's delayed post-testing result does not merit a firm conclusion. However, it does suggest the need for further research, particularly because group B, who also took the same form of the PPVT at all three testing phases, did not make significant gains.

The final vocabulary measure for this study, the modified version of the VLT, did not result in growth that could be attributed to the intervention. Overall, the VLT was a measure of students' general vocabulary knowledge at various levels of word frequency. However, the academic vocabulary section of the VLT was modified to include more of the intervention target words. Thus, the absence of effects on this section appeared to contradict the growth shown on the MAV. However, the nature of the tasks may explain the apparent contradictory findings. Success on the VLT depended on students' skills with definitions, a skill set which Scott and Nagy (1997) found to be quite under-developed in fourth and sixth graders. While the intervention for the current study did include regular activities with dictionary definitions, students always had other tools, such as peers, graphics, and sentences, available to help them work with the target words. The format of the VLT, which measured students' understanding of brief definitions in print of the target words, may have been too challenging, thereby lacking sufficient sensitivity to show effects of students' growth. Additionally, the MAV was administered in an interview format, thereby omitting any confounds of writing or reading comprehension in English that may have affected student performance on the VLT.

Overall, the results related to the first research question indicate that the intervention was successful in helping students learn the target words. Many vocabulary studies have demonstrated that students make gains on measures aligned with the instruction (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003; Beck & McKeown, 2007; Bos & Anders, 1990; Nash & Snowling, 2006; Stahl et al., 1990). However, the unique contribution of the current study is that it demonstrates the efficacy of research-based strategies in helping adolescent EL students, an understudied population, learn academic vocabulary words, a set of words previously unexamined in an experimental study.

To what extent do language skills in English mediate middle school English learners' vocabulary growth in a vocabulary development intervention?

The results demonstrated that language skills in English did mediate participants' growth in the context of the intervention. The results also suggested that students' initial proficiency in English was related to vocabulary growth. Indeed, participants with greater English proficiency, as indicated by their CELDT levels, made greater gains during the intervention. This finding is consistent with reading comprehension research that has shown that new words will likely not be learned from a context with many new or difficult words (Adolphs & Schmitt, 2003; Biemiller, 1999). Thus, students who know more words have more tools at their disposal to learn more sophisticated words. Furthermore, because word learning for all students is incremental (Stahl & Nagy, 2006), students with less proficiency in English were likely learning in smaller increments than students with more proficiency in English. As a result, interventions addressing academic language may to require a threshold of English proficiency to make significant gains in academic vocabulary knowledge.

In particular, participants' growth on the MAV during the control period had the most predictive utility for intervention growth. Interestingly, the beta coefficient was negative, which suggested that students who made the least growth in the

absence of the intervention made the most growth during the intervention. The PPVT results showed a different trend; participants who entered the study with greater breadth of vocabulary knowledge showed more growth during the intervention than those with less breadth of vocabulary knowledge. These two findings create a complex picture. Participants who showed less growth outside of the intervention, but who had larger breadth of vocabulary, showed more growth in academic vocabulary during the intervention.

The predictive utility of vocabulary knowledge for language growth is aligned with other vocabulary research. For example, Ramirez (1986) also found that pre-intervention vocabulary knowledge promoted growth in the context of an intervention. Also, these findings relate to Stanovich's (1986) conceptualization of the Matthew Effect in which more advanced readers build their knowledge base more rapidly than less advanced readers.

In contrast to this positive relationship between breadth of vocabulary knowledge and intervention growth, the finding that participants who made less growth during the control period made greater growth within the intervention is not as prevalent in the vocabulary research. In other words, students who were presumably less successful in their day-to-day academic setting in building academic word knowledge were more successful in the context of the intervention. The nature of the instruction may be the most plausible explanation for this pattern. One goal of the intervention was to deliver a fast-paced, engaging, and highly interactive set of activities with the target words. This approach may have met the needs of the students who had difficulty staying engaged in a more passive learning environment. Indeed, Scott, Jamieson-Noel, and Asselin (2003) found that traditional, often passive, methods of vocabulary instruction dominate most classrooms. While the use of interactive strategies may help to explain this finding, a case study analysis of the school context would be necessary to compare the use of instructional strategies, and such an analysis was beyond the scope of this study.

Essentially, the components of Language Workshop better met the needs of students who were not as successful at learning academic vocabulary as their peers in the absence of the intervention. Several studies have demonstrated the effectiveness of early literacy intervention with low-achieving students (Biemiller & Slonim, 2001; D'Angiulli, Siegel, & Maggi, 2004), and other research has demonstrated the efficacy of interactive strategies for vocabulary development (Bos & Anders, 1990). Furthermore, there have been calls for supplemental support beyond traditional school hours for EL students (Hakuta et al., 2000). The current findings are convergent with this research, suggesting that after-school, research-based literacy support for adolescent EL students can both accelerate language development and meet the needs of the students making the least gains during the traditional school day.

Limitations

The current findings should be considered in light of several limitations. One such limitation was the duration of the intervention, as Nagy (2005) recommended that vocabulary instruction should be long term and intensive. Two seminal intervention

studies each involved about 15 weeks of instruction (Beck et al., 1982; Carlo et al., 2004). However, the goals of these longer interventions involved reading comprehension improvement, which was not a direct goal of the current study. The absence of comprehension measures may be considered a limitation in any vocabulary development study. However, even in studies that examine and measure reading comprehension, scholars in EL vocabulary intervention research struggle with demonstrating reading comprehension gains (Shanahan & Beck, 2006). For example, after a 14 week vocabulary development intervention, Carlo et al. (2004) found significant vocabulary gains but no reading comprehension gains. Baumann et al. (2003) found a similar pattern following an intervention lasting 25 days. In light of these findings, as well as the short and targeted nature of the current study, an examination of reading comprehension was beyond the scope of this study.

Two other limitations were the sample size and the lack of normed assessments to measure growth in academic vocabulary knowledge. Because the intervention was a voluntary after-school program, sample size was a concern from the initiation of this study. Under these circumstances, and the recruitment of only approximately 100 students, the regular attendance of 37 students was an acceptable sample size. In addition, the practically and statistically significant growth on the MAV demonstrated the efficacy of the intervention in building students' knowledge of the target words. However, the power analyses suggested that we did not have sufficient statistical power to detect growth on some of the measures, particularly the VLT or the PPVT. While we would not have expected to see growth on the PPVT as it is a measure of general vocabulary rather than academic vocabulary, the VLT was partially aligned with the intervention. However, the results on these measures do not lend themselves to strong conclusions because of the lack of power. Future research with this or a similar intervention should include more participants in order to garner the statistical power necessary for stronger conclusions on the overall efficacy of the intervention. In addition, a larger sample would allow for more advanced multivariate analyses using more than just the three variables used in the analyses for the present study.

The lack of normed assessments was also an issue in this study. While the format of the MAV had been previously used as the Expressive Vocabulary Scale, it had not been used with general academic vocabulary words. Additionally, neither the VLT nor the PPVT were designed to measure the construct of general academic vocabulary words. Thus, robust findings across measures were not expected; indeed, the current study shows the need for the development of academic vocabulary assessments. Despite the small sample size and the lack of normed assessments, conservative data analyses still showed growth on the target words, suggesting the promising nature of the intervention.

Implications and future research directions

In relation to the adolescent literacy literature, the current study provides further evidence for the principles of successful adolescent literacy programs (see Biancarosa & Snow, 2004, for an overview) and also makes a novel contribution to the literature. For example, this was the first time research-based strategies were

implemented in an experimental study with an adolescent EL student sample and academic vocabulary words. Furthermore, this effort was apparently successful. Indeed, the within-subject design revealed that participants showed greater growth in academic vocabulary knowledge during the intervention rather than the control period. Thus, this study strongly suggests that vocabulary instruction strategies previously determined to be effective with younger and NS populations are also effective with adolescent EL students and academic vocabulary. In addition, this study suggests that EL students with intermediate to advanced English proficiency who are not making gains as rapidly as their peers in the day-to-day school context may benefit the most from similar interventions. Furthermore, the after-school setting used to deliver intervention illustrates an example of how the supplemental support advocated by Hakuta et al. (2000) might be effectively implemented.

This study highlights two important areas for future research. First, the relationship between reading comprehension of academic texts and academic vocabulary knowledge deserves focused research attention, particularly for middle school EL students. In regards to Language Workshop, future research should include reading comprehension measures as well as the integration of reading comprehension instruction into the intervention itself. The second area of future research is the assessment of academic vocabulary, and academic English in general. Pressley, Disney, and Anderson (2007) assert that future research on vocabulary acquisition includes careful analyses of degrees of word learning, and that assessments need to detect differences in degrees of word knowledge. In the larger picture of academic English development, Bailey (2007) questions the efficacy of current ELD assessments in measuring academic language. Bailey and Butler (2007) advocate for the development of academic English language assessments that can accurately predict EL students' readiness to maintain the pace of mainstream classrooms. Such assessments should include both general and content-specific academic vocabulary, as well as the other components of academic English.

The current study, and the body of related research, emphasizes the importance of, and urgency for, supporting adolescent EL students. Research on early elementary EL students has revealed the promising effects of early ELD support and intervention (D'Angiulli et al., 2004; Lipka, Siegel, & Vukovic, 2005). However, following participation in these programs, many adolescent EL students may still need support. Furthermore, EL students will continue to immigrate to English-speaking countries later in their childhood or in early adolescence when they may not be able to benefit from early intervention. Clearly, there is a great need to continue and accelerate the research and instructional efforts to support the academic literacy development of adolescent EL students.

Appendix

Appendix A Sample week from language workshop

Week 2: Academic target words	Main instructional activities
Monday:	• Picture puzzler with discussion around words
structure	• Reading from *Great Inventions* (Wood, 1995), specifically looking at the structure and significance of some of Leonardo da Vinci's inventions
issue	
significant	• Matching game—students find their groups by matching up their cards and then draw a picture of their words on flip chart paper. Students from other groups then guess what word the pictures represent
Tuesday:	• Picture puzzler with discussion around words
section	• Speed sentences—using pictures from previous day posted around the room, students start at one and write a sentence for the word. Every 2 minutes, a bell rings and the groups rotate to write a sentence for the next picture they come to. Followed by corrections of and discussion on word use
occur	
process	
	• Reading aloud from *Wayside School is Falling Down* (Sachar, 2003)
Wednesday:	• Picture puzzler with discussion around words
method	• Reading from *Great Inventions* (Wood, 1995) with attention to target words
analysis	• Pictionades Game—combination of Pictionary and Charades using target words and general vocabulary from instructional texts (This game provides opportunities to practice with concrete words)
role	
Thursday:	• Music puzzler—students listen to a selection of music excerpt with discussion about the songs using target words, i.e. Can you *interpret* what the song means? What did you like, *specifically*, about this song? Is this song *similar* to another song you like?
interpret	
specific	
similar	• Taboo Game—2 teams, pairs of students work together in their teams to get their teams to say certain words (This game provides opportunities to practice with abstract words)
	• Reading aloud from *The Bad Beginning* (Snicket, 1999)

References

Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics, 24,* 425–438.

Anderson, V., & Roit, M. (1996). Linking reading comprehension instruction to language development for language-minority students. *The Elementary School Journal, 96,* 295–309.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20,* 50–57.

Bailey, A. L. (2007). Introduction: Teaching and assessing students learning English in school. In A. L. Bailey (Ed.), *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.

Bailey, A. L., & Butler, F. A. (2007). A conceptual framework of academic English language for broad application to education. In A. L. Bailey (Ed.), *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal, 40,* 447–494.

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*, 251–271.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life*. New York: The Guilford Press.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*, 506–521.

Bessai, F. (2001). Peabody picture vocabulary test – III. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 908–911). Lincoln, NE: The Buros Institute of Mental Measurement.

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading, 9*, 43–61.

Biancarosa, G., & Snow, C. (2004). *Reading next: A vision for action and research in middle and high school literacy*. New York: Carnegie Corporation of New York.

Biemiller, A. (1999). *Language and reading success*. Newton Upper Falls, MA: Brookline Books.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology, 93*, 498–520.

Blachowicz, C. L. Z., & Fisher, P. (2000). Vocabulary instruction. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 503–523). Mahwah, NJ: Lawrence Erlbaum Associates.

Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly, 13*, 31–42.

Bruna, K. R., Vann, R., & Escudero, M. P. (2007). What's language go to do with it? A case study of academic language instruction in a high school "English Learner Science" class. *Journal of English for Academic Purposes, 6*, 36–54.

Calderon, M., August, D., Slavin, R., Duran, D., Madden, N., & Cheung, A. (2005). Bringing words to life in classrooms with English-Language Learners. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 115–136). Mahwah, NJ: Lawrence Erlbaum Associates.

California Department of Education. (2007). *California English Language Development Test (CELDT)*. San Francisco, CA: California Department of Education.

California State Board of Education Standards and Frameworks. (2004). Retrieved March 1, 2005, from http://www.cde.ca.gov/be/st/.

Carlo, M. S., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., et al. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 40*, 188–215.

Carlo, M. S., August, D., & Snow, C. E. (2005). Sustained vocabulary-learning strategy instruction for English-language learners. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 137–153). Mahwah, NJ: Lawrence Erlbaum Associates.

Chiappe, P., Siegel, L. S., & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading, 6*(4), 369–400.

Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A Rationale for Pedagogy*. New York: Cambridge University Press.

Corson, D. (1997). The learning and use of academic English words. *Language Learning, 47*, 671–718.

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

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing, 16*, 129–147.

Coxhead, A., & Nation, I. S. P. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes*. New York: Cambridge University Press.

CTB/McGraw-Hill. (2005). *California English Language Development Test*. Monterey, CA: CTB/McGraw-Hill.

Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada. A reassessment. *Applied Linguistics, 2*, 132–149.

Cunningham, J. W., & Moore, D. W. (1993). The contribution of understanding academic vocabulary to answering comprehension questions. *Journal of Reading Behavior, 25*, 171–180.

Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*, 82–88.

D'Angiulli, A., Siegel, L. S., & Maggi, S. (2004). Literacy instruction, SES, and word-reading achievement in English-language learners and children with English as a first language: A longitudinal study. *Learning Disabilities Research & Practice, 19*, 202–213.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*, 78–103.

Dunn, L., & Dunn, M. (1997). *Peabody picture vocabulary test – III.* Circle Pines, MN: American Guidance Services.

Geva, E., Yaghoub-Zadeh, Z., & Schuster, B. (2000). Understanding individual differences in word recognition skills of ESL children. *Annals of Dyslexia, 50*, 123–154.

Graves, M. F. (1986). Vocabulary learning and instruction. *Review of Research in Education, 13*, 49–89.

Guerrero, M. D. (2004). Acquiring academic English in one year: An unlikely proposition for English language learners. *Urban Education, 39*, 172–199.

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara, CA: The University of California Linguistic Minority Research Institute.

Houk, F. A. (2005). *Supporting English language learners: A guide for teachers and administrators.* Portsmouth, NH: Heinemann.

Institute of Education Sciences. (2007). *NAEP data explorer.* Washington, DC: National Center for Educational Statistics.

Jimenez, R. T., Garcia, G. E., & Pearson, P. D. (1996). The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly, 31*, 90–112.

Kame'enui, E. J., Dixon, R. C., & Carnine, D. W. (1987). Issues in the design of vocabulary instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum Associates.

Klinger, J. K., & Vaughn, S. (2004). Strategies for struggling second-language readers. In T. L. Jetton & J. A. Dole (Eds.), *Adolescent literacy research and practice*. New York: The Guilford Press.

Levy, D. H. (2003). *Stars and planets*. New York: Barnes & Noble Books.

Lipka, O., Siegel, L. S., & Vukovic, R. (2005). The literacy skills of English language learners in Canada. *Learning Disabilities Research & Practice, 20*, 39–49.

Margosein, C. M., Pascarella, E. T., & Pflaum, S. W. (1982). The effects of instruction using semantic mapping on vocabulary and comprehension. *Journal of Early Adolescence, 2*, 185–194.

McKeown, M. G., & Beck, I. L. (2004). Direct and rich vocabulary instruction. In J. F. Baumann & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 13–27). New York: The Guilford Press.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior, 15*, 3–18.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20*, 522–535.

Mohan, B., & Slater, T. (2006). Examining the theory/practice relation in a high school science register: A functional linguistic perspective. *Journal of English for Academic Purposes, 5*, 302–316.

Nagy, W. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Nagy, W., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 269–284). Mahwah, NJ: Lawrence Erlbaum Associates.

Nash, H., & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary knowledge: A controlled evaluation of the definition and context methods. *International Journal of Language & Communication Disorders, 41*, 335–354.

Nation, I. S. P., & Kyongho, H. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System, 23*, 35–41.

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.* Washington, DC: National Institute of Child Health and Human Development.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*, 554–566.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy*. New York: Cambridge University Press.

Pressley, M., Disney, L., & Anderson, K. (2007). Landmark vocabulary instructional research and the vocabulary instructional research that makes sense now. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 205–248). New York: The Guilford Press.

Ramirez, S. Z. (1986). The effects of Suggestopedia in teaching English vocabulary to Spanish-dominant Chicano third graders. *Elementary School Journal, 86*, 325–333.

Read, J. (2000). *Assessing vocabulary*. New York: Cambridge University Press.

Sachar, L. (2003). *Wayside school is falling down*. New York: Harper Trophy.

Scarcella, R. C. (2003). *Accelerating academic English: A focus on English language learners*. Regents of University of California.

Schleppegrell, M. J., & Go, A. L. (2007). Analyzing the writing of English learners: A functional approach. *Language Arts, 84*, 529–538.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing, 18*, 55–88.

Scott, J. A., Jamieson-Noel, D., & Asselin, M. (2003). Vocabulary instruction throughout the day in twenty-three Canadian upper-elementary classrooms. *The Elementary School Journal, 103*, 269–312.

Scott, J. A., & Nagy, W. (1997). Understanding the definitions of unfamiliar verbs. *Reading Research Quarterly, 32*, 184–200.

Shanahan, T., & Beck, I. L. (2006). Effective literacy teaching for English-language learners. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners*. Mahwah, NJ: Lawrence Erlbaum Associates.

Snicket, L. (1999). *The bad beginning*. New York: HarperCollins Publishers.

Snow, C. (2006). Cross-cutting themes and future research directions. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners* (pp. 631–651). Mahwah, NJ: Lawrence Erlbaum Associates.

Snow, C., & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York: Carnegie Corporation of New York.

Spycher, P. (2007). Academic writing of adolescent English learners: Learning to use "although". *Journal of Second Language Writing, 16*, 238–254.

Stahl, S., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72–110.

Stahl, S., & Nagy, W. (2006). *Teaching word meanings*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stahl, S., Richek, M. A., & Vandevier, R. J. (1990). *Learning meaning vocabulary through listening: A sixth-grade replication*. Paper presented at the National Reading Conference, Miami, FL.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.

Wesche, M. B., & Paribakht, T. S. (2000). Reading-based exercises in second language vocabulary learning: An introspective study. *The Modern Language Journal, 84*, 196–213.

West, M. (1953). *A general service list of English words*. London: Longman.

Wood, R. (1995). *Great inventions*. New York: Barnes & Noble Books.

Zwiers, J. (2006). Integrating academic language, thinking, and content: Learning scaffolds for non-native speakers in the middle grades. *Journal of English for Academic Purposes, 5*, 317–332.