Enhancing social studies vocabulary and comprehension for 7th grade English language learners: Findings from two Experimental studies

Sharon Vaughn, Leticia R. Martinez, Colleen K. Reutebuch, Coleen D. Carlson,
Sylvia L. Thompson and David J. Franci

Background/context. Adolescent ELLs who lack academic English language knowledge and who demonstrate low literacy levels are at risk for academic failure in content area classes (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006). Research regarding evidenced-based instruction for ELL adolescents in content learning is lacking. Our goal was to address the need for research in this area by examining the effects of an enhanced social studies instruction designed specifically for students who are ELLs that would benefit all students.

Purpose / objective / research question / focus of study. We identified instructional practices associated with improved outcomes for ELLs: (1) research-based vocabulary and concept instruction, (2) the use of media to build comprehension and concept knowledge, (3) the use of graphic organizers, and (4) structured peer-pairings. The purpose of our two studies was to examine the efficacy of incorporating instructional practices associated with improved outcomes into middle-school social studies instruction as a means of enhancing vocabulary knowledge and comprehension for English language learners. Our primary research question was: How does a multi-component instructional routine developed to enhance effective outcomes for English language learners and provided by classroom social studies teachers influence students' outcomes in vocabulary and comprehension?

Setting. Two middle schools from two districts in central Texas with large numbers of ELs.

Population / Participants / Subjects. Two different non-overlapping samples of classes of 7th grade students (N=381 and N=507) were randomly assigned at the classroom (i.e., section) level to a social studies intervention or to business as usual comparison groups. The number of sections assigned to treatment was seven and nine in experiments one and two, respectively.

Participants and Setting: Experiment 1

Schools. Participants were drawn from two middle schools in the same central Texas school district. Both schools were considered to have a substantial number of English language learners who were designated by the school as "Limited English Proficient (LEP)". At least 65% of the population at both schools was Latino with 11.45% of the students at one school designated as LEP and 13.80% of students at the other school as LEP. The proportion of students who qualified for the free or reduced-price lunch program ranged from 70 to 82%.

Teachers. The four participant teachers provided seventh-grade Texas History instruction to all the students in this study. Of the four teachers (two female, two male), two were first-year teachers who were certified to teach Social Studies (4-8) in the state of Texas. One of the male teachers had eight years of experience teaching social studies in secondary school settings. Additionally, one of the female teachers was Texas certified as a Generalist (4-8) and self-contained Bilingual/ESL teacher (1-8) and had six years of teaching experience. Classes were randomized to treatment and control conditions. Participating teachers, with support from research staff, implemented treatment conditions in intervention classes and continued with their typical instruction in comparison classes.

Participants and Setting: Experiment 2

Students. In the year after the completion of experiment 1, two middle schools from two districts in central Texas with large numbers of ELLs participated in experiment 2. Only one of

the schools had also participated in experiment 1. That school's LEP student population grew from 14% in year one to 20% in year two. The second school was new to the study and had 51% Latino students and 14% with a LEP status. The percentage of students who were eligible for free or reduced-price lunch was 68% at one school and 85% at the other school.

Teachers. During experiment 2, four teachers were identified by the principals as teaching social studies and participated in the study. All four teachers were males and certified to teach Social Studies (4-8) in the state of Texas. Two of the teachers were second-year teachers, one was a first-year teacher, and one had three years of teaching experience.

Intervention / Program / Practice. The treatment intervention was comprised of: a) overview and vocabulary instruction, b) the use of brief videos and purposeful discussion to build concepts, c) the use of graphic organizers and other writing activities to build comprehension and vocabulary through writing, and d) structure paired grouping. Students in the treatment classes received the intervention during their regularly scheduled seventh-grade social studies class. The intervention was implemented for 50 minutes a day, five days a week for approximately nine to twelve weeks. The number of lessons was the same across teachers and studies but interruptions in school schedule extended the number of weeks it took to complete the intervention. The researcher-designed lessons were used by teachers and included all the aforementioned intervention components Lesson plans identified the core subject matter and the "big ideas" that the students needed to learn in their social studies course as well as guided the teachers on the use of specific instructional practices to convey the subject matter. These practices were designed to enhance students' understanding of social studies content and of expository text by giving all their students opportunities to learn and use the vocabulary, concepts, big ideas, and issues associated with social studies. The lesson plans were not meant to be a script for teachers, but rather a guide for how to build in the study's strategies and materials. The unit lessons were designed around one or two central ideas that served as organizing concepts to help the teacher focus the events and ideas in each unit. Every lesson was organized similarly to encourage the teacher to develop a routine for the intervention.

Research Design. Two experimental studies in two successive school years with non-overlapping samples were conducted. Classes were randomly assigned to treatment and control conditions. The advantage to this design is that teachers were the same for both of our conditions and students in both the treatment and business as usual conditions covered the same material over the same period of time using the same textbook providing students in each condition equal access to learning content and key vocabulary.

Data Collection and Analysis. Prior to the intervention and after its completion all students were assessed with a researcher-developed Content-based measure (CBM). The measure was designed to cover students' understanding of the content taught during a nine to twelve week period and was meant to serve as an indicator of growth in social studies learning. It resembled traditional assessments of content area classes in that it consisted of vocabulary matching items and comprehension questions. The items were developed based on content in textbook and weekly quizzes. Students in both treatment and controls covered this same content. The vocabulary section had 20 items that included definitions that had to be matched with vocabulary terms used within the context of a sentence that contained social studies information. For example, the definition, to officially give up power or territory, had to be matched to the target

word *cede* in the sentence Mexico agreed to *cede* much of its northern territory to the United States. The second part of the assessment included 10 questions asking students to identify and explain the big ideas of the social studies units taught during the instruction. For example, one comprehension item required students to explain two ways in which slaves' human rights were violated. The content represented in these big ideas was part of the instructional materials and state standards and thus the content was part of the instructional materials for both treatment and control students.

Analyses of pre-and post-vocabulary and comprehension performance were examined separately for each study. The first step of the analyses examined differences in pre-test scores as a function of group (treatment or control). The second step examined group differences in post-intervention performance as a function of treatment group controlling for pre-test measures of the outcome variable. All analyses were conducted using three-level, hierarchical linear models in HLM 6.06 (Raudenbush, Bryk, & Congdon, 2008). For all HLM analyses, we report results for fixed effects of treatment based on robust standard errors. The three-level models included variability due to students within section, sections within teachers, and teachers. In all models, treatment was entered at level-2 (i.e., at the section level). We also tested for heterogeneity of regression between Treatment and Control sections, and examined models that allowed for pre-test regressions that varied by teacher, as well as the possibility that treatment effects differed across teachers. However, due to the small number of teachers in the study, we focus here on results from models where pre-test regressions and treatment effects were constrained to be fixed across teachers. Thus, in all reported models, random effects due to teachers were limited to effects on the intercept, i.e., the average value across all sections, both treatment and control, for that teacher.

Findings / Results.

Results for Study 1

ELL and non-ELL students differed at the pre-test, although these differences were comparable for Treatment and Control sections. More importantly, pre-test scores were not different between Treatment and Control sections. A three-level HLM analysis of pre-test scores showed no differences between Treatment and Control sections for either Comprehension ($t_{(13)}$ = -0.970, p = .350) or Vocabulary ($t_{(13)}$ = 0.552, p = .590). Analysis of post-test scores using a three-level analysis of covariance revealed statistically significant differences between students in Treatment and Control sections for both Comprehension and Vocabulary. For comprehension there was a significant effect for treatment condition indicating that students in treatment sections were performing at significantly higher levels than students in control sections post-intervention ($t_{(13)}$ = 14.31, $p \leq .001$). The estimated difference in comprehension scores between treatment and control sections was 1.57 with a standard error of 0.109. Similar results were found for vocabulary, although the absolute difference between treatment and control groups was found to be somewhat larger (Treatment Effect Estimate = 2.53, s.e. = 0.629, $t_{(13)}$ = 4.026, p = .002). For both vocabulary and comprehension, treatment effects did not interact with student status as ELL or non-ELL indicating that ELL and non-ELL students benefitted equally from participation in treatment sections. Translating the above differences into effect sizes shows that the effects of the intervention were large for both vocabulary and comprehension. For comprehension, we estimated the effect size to be g = 1.12. For vocabulary, we estimated g = 0.53.

Results for Study 2

Analysis of pre-test performance scores indicated no significant, differences between treatment and control sections for either vocabulary ($t_{(15)} = 1.450$, $p = .168$) or comprehension ($t_{(15)} = 0.934$, $p = .366$. Similar to study 1, students performed more poorly on the comprehension measure at both time points than on the vocabulary assessment. However, one must be cautious in interpreting these differences between domains as the assessments were not developed in a manner that would guarantee that they are equivalent in difficulty. Results from the three-level HLM analysis with student pre-test scores as a covariate yielded results highly similar to those from study 1. As in study 1, performance at the pre-test significantly predicted student performance at the post-test for both vocabulary ($\beta = 0.515$, s.e. = 0.107, $t_{(363)} = 4.82$, $p < .001$) and comprehension ($\beta = 0.596$, s.e. = 0.030, $t_{(361)} = 19.46$, $p < .001$). More importantly, students in treatment sections again outperformed those in control sections on both the comprehension and vocabulary measures (Comprehension: $\beta = 1.09$, s.e. = 0.403, $t_{(15)} = 2.71$, $p = .016$; Vocabulary: $\beta = 1.94$, s.e. = 0.550, $t_{(15)} = 3.53$, $p = .003$). As in study 1, treatment effects did not interact with students' ELL status for either outcome, indicating that both ELL and non-ELL students benefited equally from being assigned to sections that were randomly assigned to the treatment condition. Expressing the treatment – control differences as effect sizes using *g* as in study 1, we find that effect sizes are overall somewhat smaller, but again are classified as large or moderate, and are somewhat larger for comprehension than for vocabulary. In the case of vocabulary, there was some suggestion that pre-test regressions might differ across teachers and that treatment effects might differ across teachers. However, because significance tests for random effects may be misleading when the number of sampling units is small, in this case $n_{teachers} = 4$, we have focused above on the average treatment difference between treatment and control sections, averaged across teachers from the three-level HLM model that constrained the regression effect for the pre-test to be the same for all teachers.

Conclusions. Although this intervention was developed to address the instructional and language needs of ELLs, the students who were not limited English proficient in the intervention classes also benefited. Students who were limited English proficient outperformed their counterparts in the comparison condition on both the vocabulary and comprehension measures. When both the target group (English language learners) and their classmates benefit from an intervention or practice, it meets the criteria for universal design. This finding is particularly relevant for teachers who have both ELLs and non-LEP students in their classrooms and who may be concerned about the possible detrimental effect for other students of instruction that targets ELLs. If effective instructional practices for ELLs also benefit non-ELLs, teachers have a strong rationale for implementing the instructional practice. Furthermore, ELLs in the comparison condition made the least gains and lagged behind all other groups on both the vocabulary and content comprehension measures, providing further support for interventions such as the one in this study to alter the course for ELLs.

Appendices

Appendix A. References.

Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions (Under cooperative agreement grant S283B050034 for U.S. Department of Education).* Portsmouth, NH: RMC Corporation, Center on Instruction, Center on Instruction. Available online at http://www.centeroninstruction.org/files/ELL1-interventions.pdf.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2008). HLM 6 (Version 6.06). Lincolnwood, IL: Scientific Software International.