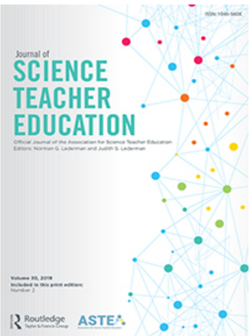

Routledge
Taylor & Francis Group

Journal of Science Teacher Education

ISSN: 1046-560X (Print) 1573-1847 (Online) Journal homepage: https://www.tandfonline.com/loi/uste20

Examining the Effects of a Professional Development Initiative on English Learning and Economically Disadvantaged Adolescents' Scores on a High-Stakes Science Test

Julie K. Jackson, Margarita Huerta, Tiberio Garza & Rose Narvaez

To cite this article: Julie K. Jackson, Margarita Huerta, Tiberio Garza & Rose Narvaez (2019) Examining the Effects of a Professional Development Initiative on English Learning and Economically Disadvantaged Adolescents' Scores on a High-Stakes Science Test, Journal of Science Teacher Education, 30:2, 122-143, DOI: 10.1080/1046560X.2018.1538299

To link to this article: https://doi.org/10.1080/1046560X.2018.1538299

JOURNAL OF SCIENCE TEACHER EDUCATION
2019, VOL. 30, NO. 2, 122–143
https://doi.org/10.1080/1046560X.2018.1538299

Examining the Effects of a Professional Development Initiative on English Learning and Economically Disadvantaged Adolescents' Scores on a High-Stakes Science Test

Julie K. Jackson[a], Margarita Huerta[b], Tiberio Garza[c], and Rose Narvaez[d]

[a]Department of Curriculum and Instruction, Texas State University, San Marcos, Texas, USA; [b]Department of Educational and Clinical Studies, University of Nevada, Las Vegas, Las Vegas, Nevada, USA; [c]Educational Psychology and Higher Education, University of Nevada, Las Vegas, Las Vegas, Nevada, USA; [d]Edgewood Independent School District (Retired), San Antonio, Texas, USA

ABSTRACT:
This study addressed the pressing need to help adolescent English learning and economically disadvantaged student populations succeed in science and close achievement gaps. It reports the results of a two-year professional development initiative implemented at a high-poverty suburban high school in South Texas. We examined how the combination of purposeful planning, innovative academic vocabulary instruction, Saturday school, structured tutoring, and a timely focused review impacted the science achievement of EL and economically disadvantaged high-school students, at-risk of academic failure, on a high-stakes state exam. Difference-in-proportions tests were used to determine if students at the Intervention Campus showed positive achievement gains on a high-stakes state test. Then we analyzed data between schools by comparing high-stakes exam passing rates between the Intervention Campus and a Comparison Campus. This study found statistically significant results which suggest that the initiative was beneficial to the students within the Intervention Campus and showed promising effects when compared against a matched Comparison Campus.

KEYWORDS
At-risk students; English learners; economically disadvantaged; high-stakes tests; science education

Science education should be for all students and is important for developing individuals able to succeed in scientifically and technologically driven societies. Students in the process of learning English, or English learners (ELs), and students from economically disadvantaged backgrounds consistently score well below their peers on national and international measures of science achievement, including high-stakes tests (McFarland et al., 2017; Organisation for Economic Co-operation and Development, 2014), and often struggle to develop proficiency in academic English language alongside academic content (August & Shanahan, 2006). These students are considered most at risk for academic failure.

Language-infused interventions in the science classroom have been effective at helping increase the achievement scores of ELs and economically disadvantaged students (e.g., Ardasheva & Tretter, 2017; August, Artzi, & Barr, 2016; August, Branum-Martin,

CONTACT Julie K. Jackson jj32@txstate.edu Department of Curriculum and Instruction, College of Education, Texas State University, 601 University Drive, San Marcos, TX 78666, USA.
Supplemental data for this article can be accessed here.

Cardenas-Hagan, & Francis, 2009; Cuevas, Lee, Hart, & Deaktor, 2005; Echevarria, Richards-Tutor, Canges, & Francis, 2011; Lara-Alecio et al., 2012; Lee, Deaktor, Hart, Cuevas, & Enders, 2005; Lynch, Kuipers, Pyke, & Szesze, 2005; Maerten-Rivera, Ahn, Lanier, Diaz, & Lee, 2016; Shaw, Lyon, Stoddart, Mosqueda, & Menon, 2014). These interventions mention a variety of strategies to help ELs and economically disadvantaged students in their science achievement, such as vocabulary-building strategies, science notebooks, home–family literacy connections, and intense teacher preparation to consider the cultural and linguistic needs of students. Few of these studies have focused on adolescent groups, creating a concern for evidence-based strategies for promoting adolescent literacy skills (August et al., 2016). Yet researchers have noted how students who do well in the elementary grades will likely struggle as they enter the middle school and high school grades because of the high complexity of academic language and content (Nation, Cocksey, Taylor, & Bishop, 2010).

Academic language and ELs

Academic language can be defined as language that is different from the everyday language used in social or nonacademic settings. Academic language is language used in academic settings, including school settings (Cummins, 1981). The academic language of science, for example, is made up of distinctive features, including vocabulary and patterns of discourse different from everyday language use (Gee, 2004; Lemke, 1990). Everyday words such as *volume* can take on different meanings in the academic context of science, including the loudness of sound or the amount of space taken up by an object. Though arguably challenging for all students (Gee, 2004), the features of science language "can and do present significant comprehension challenges to adolescent students, especially struggling readers and English-language learners who do not have sufficient experience with academic texts in content areas such as science" (Fang, 2006, pp. 505–506).

Theory and policy consistently emphasize the importance of building ELs' academic language to facilitate their academic achievement and long-term societal success (Cummins, 1981; Quinn, Lee, & Valdés, 2012; Scarcella, 2003). Scarcella (2003), a second language theorist, stated, "Learning academic English is probably one of the surest, most reliable ways of attaining socio-economic success in the United States today. Learners cannot function in school settings effectively without it" (p. 3). Researchers involved in policy related to science standards also acknowledge the critical role that language plays in science learning. They note that teachers should act as "supporters of the language learning that occurs in a content-rich and discourse-rich classroom environment" (Quinn et al., 2012, p. 1). Clearly, researchers must continue to consider the critical role that language plays in ELs' learning in the content areas.

Academic language and economically disadvantaged students

These linguistic and academic challenges are also present for students enrolled in high-poverty schools (Kohlhaas, Lin, & Chu, 2010; Lesaux & Kieffer, 2010; Muller, Stage, & Kinzie, 2001). Low socioeconomic status has been found to be negatively related to students' achievement in science (Kohlhaas et al., 2010; Muller et al., 2001). That is, the lower the socioeconomic status of a student, the less likely it is for the student to achieve at or above set academic standards. In addition, researchers have noted uneven linguistic development patterns for English speakers enrolled in high-poverty schools (Lesaux & Kieffer, 2010). Interventions assisting the simultaneous language and academic development of students should also consider focusing on economically disadvantaged students in content area classrooms.

Academic language and science interventions

Science interventions that have included samples of EL and economically disadvantaged adolescent students across the United States have reported positive gains on student achievement measures (Ardasheva & Tretter, 2017; August et al., 2009; Echevarria et al., 2011; Fang & Wei, 2010; Lara-Alecio et al., 2012; Lynch et al., 2005; Tong et al., 2014). The studies and their findings are synthesized here and informed the current work.

Multifaceted interventions

Middle school studies situated in Texas public schools used interventions with similar components, such as careful alignment of the curriculum to state standards by teachers and researchers, use of the 5-E (Engage, Explore, Explain, Elaborate, and Evaluate) science inquiry learning cycle (Bybee et al., 2006), and ongoing professional development (August et al., 2009; Lara-Alecio et al., 2012; Tong et al., 2014). August et al.'s (2009) 9-week study reported statistically significant gains from pretest to posttest on researcher-developed science tests and vocabulary and reading comprehension subtests from a commercial assessment. Note that Lara-Alecio et al.'s (2012) and Tong et al.'s (2014) yearlong quasi-experimental study reported significant and positive intervention effects in favor of treatment students on measures of district-wide curriculum-based science tests. The researchers also reported statistically significant gains on state standardized reading tests. Both of these studies also considered culturally relevant practices, such as use of the home language, home–school literacy connections, and enhanced instruction relevant to diverse students' learning needs.

Interventions focused on existing structures or curricula

Other studies in middle school settings analyzed whether existing lesson plan sequences or highly rated curricula had an impact on students' learning (Echevarria et al., 2011; Lynch et al., 2005). Echeverria's 8-week quasi-experimental study conducted in California public schools compared treatment groups using the Sheltered Instruction Observation Protocol (SIOP) model lesson plan sequence in science and control groups that did not use this model in science. The researchers reported nonsignificant gains between seventh-grade treatment and control ELs on science unit and language assessments. Though the students' teachers received professional development training on the SIOP model, the researchers attributed the nonsignificant findings to the short duration of their training and of the intervention. Somewhat similarly, Lynch et al. (2005), in a 6- to 8-week quasi-experimental study with eighth-grade public school students in the Washington, DC, area did not find statistically significant effects for the EL students in their sample on science unit assessments created by the researchers. They did, however, find statistically significant gains in their measures for non-EL students. It is problematic that the study did not mention professional development or the specific use of language supports, and the researchers stated that the literacy demands of the curriculum were likely “too great” for EL students (Lynch et al., 2005, p. 942). Other researchers have also noted that Lynch et al.’s (2005) study did not include curricular adjustments for students’ language or cultural backgrounds (August et al., 2009).

Interventions focused on specific language domains

Last, researchers in middle school and high school settings have focused on building students’ specific language domains, such as reading skills or academic vocabulary within science instruction (Ardasheva & Tretter, 2017; Fang & Wei, 2010). Fang and Wei’s (2010) year-long quasi-experimental intervention with sixth-grade students considered the effects of an inquiry-based science curriculum integrating explicit reading strategy instruction. These researchers included ongoing professional development for teachers on reading strategy instruction, a home science reading program, and a curriculum that included teacher-directed inquiry activities and field trips. The researchers reported statistically significant gains for the experimental group on a curriculum-referenced science test, in their academic year science grades based on class grades, and on a standardized reading test. Ardasheva and Tretter’s (2017) 12- to 14-week-long pretest/posttest intervention with ninth- and 10th-grade high school refugee ELs included specific vocabulary instruction. The researchers reported statistically significant pre-to-post gains in student vocabulary retention using researcher-created weekly vocabulary quizzes. Note that they did not include science achievement results.

Summary and considerations

In summary, successful science studies with samples of ELs and economically disadvantaged students, including at-risk students, shared the following key elements: (a) a curriculum created by researchers and educators and aligned with standards, (b) ongoing professional development, (c) culturally relevant practices, (d) literacy strategies embedded within science teaching practices, and (e) time (i.e., longer interventions tended to yield better results). These elements were considered in the present study’s intervention design (see “Professional Development Initiative”).

It is also important to consider that researchers reporting science gains with samples of at-risk adolescent learners report statistically significant gains for more proximal and not distal assessments. *Proximal assessments* can be defined as assessments closer in proximity to the taught curriculum, such as unit tests; *distal assessments* can be defined as assessments that test more general concepts and are further removed from the taught curriculum, such as standardized and often high-stakes tests (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). It can be argued that proximal assessments are more sensitive at detecting changes in student performance. The reality, however, is that achievement gaps for at-risk students exist in distal assessments and must be closed (Kieffer, Lesaux, Rivera, & Francis, 2009). For these reasons, we were especially interested in considering the effects of our intervention on distal high-stakes state exams.

Research purpose and questions

The pressing need to help adolescent EL and economically disadvantaged student populations succeed in science and close achievement gaps prompted the following research question: Does the combination of purposeful planning, innovative academic vocabulary instruction, Saturday school, structured tutoring, and a timely focused review impact the science achievement of English language learners and economically disadvantaged high school students on high-stakes science measures? We developed a professional development initiative (described below) to test this question and implemented it in a high-poverty high school with a large percentage of at-risk students (hereafter referred to as "the intervention campus").

We specifically asked the following:

(1) Do the standardized achievement passing rates of economically disadvantaged and EL students at the intervention campus show a statistically significant difference after students receive 1 year of the professional development initiative intervention? If so, what is the magnitude of the difference?
(2) Do the standardized achievement passing rates of economically disadvantaged and EL students at the intervention campus show a statistically significant difference after students receive 2 years of the professional development initiative intervention? If so, what is the magnitude of the difference?
(3) Are there statistically significant differences between the standardized achievement passing rates of economically disadvantaged and EL students at the intervention campus and a comparison campus? If so, what are the differences in the magnitudes?

Professional development initiative

As noted previously, the professional development initiative was modeled on theory, research, policy, and practice noted to be effective for all students, including ELs.

Purposeful planning

Standards-based instruction has been called for and conceptualized by national science education reform efforts (American Association for the Advancement of Science, 1993; National Research Council, 1996; NGSS Lead States, 2013). States, districts, and teachers operationalize reform efforts when they use national standards to frame curriculum, instruction, and assessments. States and districts interpret national standards for local use, but teachers are the critical piece in standards-based instruction. Teachers are "front line providers of instruction," and the decisions they make during planning and instruction determine the rigor and relevance of the academic content they teach (Mangiante, 2018, p. 208). Therefore, it is essential that teachers understand and use standards when they plan lessons, deliver instruction, and create assessments.

Lesson planning precedes instruction and requires teachers to consider the standards-based content that they will teach, how they will implement instruction, and how learning will be assessed. Based on the 1996 National Science Education Standards, the Texas Essential Knowledge and Skills (TEKS) are the standards used in Texas schools. The intervention campus science department members were given copies of the Texas high school biology TEKS and instructed to use this primary source document to plan lessons, align instruction, and develop assessments. Alignment of instruction and assessments with the state high school biology standards underpinned all professional development activities. Planning meetings provided teachers with opportunities to discuss content, select instructional strategies, and review academic vocabulary while heeding state standards and district guidelines. Figure 1 includes each element of the professional development initiative and the sequence in which it was implemented and organizes supporting details. Readers interested in additional information about purposeful planning are referred to Jackson and Durham (2016).

Vocabulary selection

Selecting vocabulary to teach is challenging and subjective. Furthermore, little information is available to guide teachers in the complex task of selecting academic vocabulary for lessons (Graves et al., 2014). To ensure that the vocabulary selection process was research based, focused, and systematic, teachers used the Vocabulary Planning Document (see Figure S1), which reflected the four components of a comprehensive vocabulary program: Teach a few well-selected words, teach word-learning strategies, foster word consciousness, and provide rich and varied language experiences (Graves, 2006). Teachers read the TEKS associated with the units or lessons they were planning and completed the Vocabulary Planning Document during department and team planning meetings. Vocabulary included on this document underpinned the selection of all instructional materials and framed interactive word walls.

Science department members used completed Vocabulary Planning Documents to ensure that lesson materials were aligned with the TEKS (i.e., aligned with standards). If lesson materials did not contain vocabulary included on the Vocabulary Planning Document, teachers located or created materials to support teaching the concept. If available lesson materials contained vocabulary not on the Vocabulary Planning Document, teachers discussed the merits of including the vocabulary in their lessons. If the vocabulary was aligned with the science standard, the word was added to the Vocabulary Planning Document and included in lessons. If it was not aligned, all lesson materials associated with it were discarded.

This vocabulary-driven planning process focused department and team meetings and ensured that all lessons, teaching materials, and assessments included academic vocabulary that was aligned with standards with fidelity. It also supported a robust and informed review of instructional materials because it provided a mechanism for testing the alignment of commercially available and/or traditionally used lesson materials with state science standards. After identifying essential academic vocabulary, teachers planned the structure of their interactive word walls and developed rich and varied language experiences to support academic vocabulary during instruction. Readers interested in step-by-step instructions that explain how to use the Vocabulary Planning Document to identify essential academic vocabulary and examples of completed Vocabulary Planning Documents are referred to Jackson and Durham (2016) and Jackson (2018).

First Element	Second Element	Third Element	Fourth Element
Purposeful Planning	Innovative Vocabulary Instruction Interactive Word Walls	Saturday School Structured Tutoring	Focused Review
• Department and team planning meetings, • Reviewed state science standards, • Reviewed released STAAR EOC exam questions, • Used the *Vocabulary Planning Document* to identify essential academic vocabulary, • Identified or created standards aligned lessons, • Planned the structure of interactive word walls, • Created chapter tests and checkpoint assessments, • Strong administrative support.	• Highlighted connections between science investigations, in class activities, and science concepts, • Connected science concepts and essential academic vocabulary, • Included artifacts from inquiry investigations and in class activities, • Built by students during the "Explain" phase of 5-E lessons, • Used by students to frame oral and written discourse, • Strong administrative support.	• Guided by disaggregated assessment data, • Strategically targeted content gaps, • Highlighted connections between science activities and science concepts, • Connected science concepts and essential academic vocabulary, • Reinforced understanding, • Food always available, • Raffle prizes available, • Strong administrative support, • Strong community support.	• Ten school days before STAAR EOC exam, • Replaced daily science instruction, • Guided by disaggregated assessment data, • Explicit instruction and review of essential academic vocabulary, • Strategically targeted content gaps, • Reinforced understanding, • Opportunities for students to transfer learning to released exam questions, • Strong administrative support.

Figure 1. Elements and sequence of the professional development initiative. STAAR EOC = State of Texas Assessments of Academic Readiness End-of-Course; 5-E = Engage, Explore, Explain, Elaborate, and Evaluate.

Previous exam analysis

The science department used planning time to study released biology State of Texas Assessments of Academic Readiness End-of-Course exams (STAAR EOCs). The teachers aligned state standards with released exam questions in order to better understand how EOC exam writers and the Texas Education Agency (TEA) interpreted content. They noted the use of visuals on the exam and discussed how they could use visuals to enhance instruction. Because the biology STAAR EOC contains multiple-choice questions, they carefully reviewed the structure of released multiple-choice questions. They studied question stems (problems) and reviewed answer choices, including the correct or best answer and distractors. Distractors are wrong answers designed to be reasonable choices and are usually linked to common misconceptions. Teachers used this information to guide the construction of chapter tests and checkpoint assessments and to teach students how to interpret and answer biology STAAR EOC multiple-choice questions. Data analysis, state standards, and the linguistic needs of students drove the planning process and informed all instructional decisions.

Innovative vocabulary instruction: Interactive word walls

Word walls serve as visual scaffolds and are a common classroom tool used to display vocabulary that students have learned in class. The structure of word walls varies, but they are usually static collections of words or word banks created by teachers that students reference to learn or review academic vocabulary. Research found that student achievement improved when word walls were conceptually organized, included student-generated materials, and linked academic vocabulary and visual cues (Jackson & Ash, 2012). To support vocabulary development in science, we replaced traditional teacher-generated word walls with word walls that were created and used by students.

Interactive word walls support academic language development during science instruction because they strategically target academic vocabulary, visually display connections between inquiry science activities and academic vocabulary, and are student generated. Interactive word walls may resemble graphic organizers or data tables. They highlight connections between concepts and artifacts (realia) from inquiry-based science activities while connecting scientific concepts and academic vocabulary. Interactive word walls usually include a visual representation of specific vocabulary words and labels. Definitions are optional. Interactive word walls are an effective scaffold that supports students as they develop an understanding of key academic vocabulary (Jackson & Ash, 2012).

Teachers plan the structure of interactive word walls, select the academic vocabulary, and organize the sequence in which the word walls are built. Students complete interactive word walls in class during the “Explain” section of a 5-E lesson. The 5-E lesson plan model (Bybee et al., 2006) is designed to structure and support scientific inquiry. Lesson plan components are Engage, Explore, Explain, Elaborate, and Evaluate. Building interactive word walls allows students to have multiple encounters with new and familiar academic vocabulary in contexts as they actively process a word’s meaning and connect academic vocabulary to inquiry activities. Photos of interactive word walls and descriptions of how teachers involve students in the construction of interactive word walls are included in

Jackson (2018), Jackson and Durham (2016), Jackson et al. (2017), and Jackson and Narvaez (2013).

Saturday school and structured tutoring

The intervention campus Saturday school and structured tutoring programs were established to provide remediation and tutoring for students who performed poorly on a benchmark assessment, chapter test, and/or checkpoint assessment. The benchmark assessment was a released biology STAAR EOC. Chapter tests and checkpoint assessments were designed to be rigorous evaluations. Created by biology teachers, they included reformatted Advanced Placement (AP) biology exam questions, and in an attempt to mirror the structure of the biology STAAR EOC they also included multiple-choice questions with illustrations, diagrams, data tables, charts, and graphs.

In combination, these assessments provided immediate feedback regarding student performance on subsets of state science standards. Intervention campus science teachers and the campus administrator in charge of the science department used disaggregated assessment data to guide remediation and tutoring objectives and to identify students who needed to attend Saturday school and/or before- or after-school tutoring. This data-driven process ensured that no student was left behind and that Saturday school and tutoring lessons strategically targeted content gaps. Identified students were given a letter notifying them that Saturday school attendance and/or tutoring was required. A copy of the letter was mailed to their parent or guardian. The letter was signed by the principal and explained that passing the biology STAAR EOC was a high school graduation requirement; described the school's Improvement Required (IR) status; and detailed how the rating impacted students, the school, and the community. If invited students did not attend tutoring or Saturday school, the administrator in charge of the science department was notified by e-mail. Then the administrator spoke directly with the student and called his or her parent or guardian.

The administrator informed the parent/guardian and the student that passing the biology STAAR EOC was a high school graduation requirement. When attendance obstacles were present (the need to work, transportation challenges, or the need to care for family members), the administrator worked with the parents and students to create workable solutions. If students incurred a second absence, a parent conference was scheduled.

Saturday school was implemented as an academic intervention in 2013–2014 and 2014–2015. It was held twice a month before the district benchmark. After the benchmark, it was held weekly and continued until the Saturday before the biology STAAR EOC. Saturday school and tutorial activities provided opportunities for students to experience content from a new perspective via examples and materials not used during original instruction. Saturday school began at 8:30 a.m. and ended at 12:30 p.m. and was divided into two blocks.

The first block was designed for students who failed the district benchmark. These students rotated through three different remediation lessons that focused on content gaps exposed on the benchmark. These lessons included hands-on inquiry science activities, highlighted connections between the science activities and scientific concepts, and helped students connect scientific concepts and academic vocabulary. The majority of these students were ELs.

The second block was composed of students who had passed the district benchmark but could benefit from additional instruction/review. The goal was to make sure that these students did not become complacent or regress. The teachers working with these students developed tutorial lessons based on AP biology lessons and labs. Tutorial groups were limited to 10 students, and student benchmark and chapter test scores determined which tutorials the students attended. All sessions concluded with a quiz that included biology STAAR EOC–style questions.

Originally only students who failed the benchmark tests were invited to attend Saturday school and attendance was low. In an effort to rally local support for Saturday school and tutoring, community stakeholders were informed about the intervention campus's IR rating and information about IR ratings and the need to improve STAAR EOC scores was disseminated in English and Spanish by a free Spanish newspaper, housing project information centers, parent–teacher association websites, and the local library. Community support was immediate. The high school principal provided free breakfast tacos and orange juice, a local grocery store donated free snacks that were available throughout the morning, and community businesses donated raffle prizes. Saturday school teachers distributed raffle tickets that the students used to enter weekly drawings. Prizes included movie tickets, iTunes gift cards, and McDonald's gift cards. Saturday school quickly became very popular and attendance increased to well over 120 students each Saturday. Saturday school attendance was tracked, and scores on Saturday school quizzes were recorded and used to select topics for future intervention lessons.

Focused review

The science department and the administrator in charge of science instruction met a month before the biology STAAR EOC date to plan an intense focused review. The review was designed to address identified gaps in content knowledge, review key academic vocabulary, and reinforce understanding. Benchmark assessments, chapter tests, checkpoint assessments, and Saturday school quiz results were revisited in order to identify concepts and academic vocabulary that students continued to find challenging. The science department used this information to select content and academic vocabulary that would be included in review activities. The review replaced daily science instruction and occurred 2 weeks (10 school days) before the biology STAAR EOC. Review lessons included explicit academic vocabulary instruction and review, contextual application of academic vocabulary, and an opportunity for students to transfer learning to released biology STAAR EOC questions.

Table 1. Intervention campus enrollment history and ethnic distribution.

School year	All students	African American (%)	Hispanic (%)	White (%)	Native American (%)	Asian (%)	Pacific Islander (%)	Economically disadvantaged (%)	Limited English proficient (%)[a]	At risk (%)
2014–2015	1,244	1.2	98.1	0.4	0.0	0.2	0.0	91.0	9.7	83.5
2013–2014	1,270	0.9	98.2	0.6	0.0	0.2	0.0	97.8	10.5	82.5
2012–2013	1,264	0.7	98.1	0.7	0.1	0.2	0.0	88.4	8.1	76.3

[a]The Texas Education Agency refers to students not fluent in English as *limited English proficient*, but in practice *English learner* is more common.

Timeline

The intervention campus biology team (three teachers) and the campus administrator in charge of science instruction met before the beginning of the school year (2013–2014 and 2014–2015) to review the instructional intervention plan that would be used to guide science instruction. The plan included purposeful planning, innovative academic vocabulary instruction, Saturday school, structured tutoring, and a timely focused review. To support the intervention and provide a channel for clear and timely communication, the administrator in charge of science instruction attended all science department meetings, science department planning days, and biology team planning sessions. In addition, the administrator regularly observed science lessons and provided detailed feedback that teachers used to improve instruction.

Science department meetings were held twice a month to provide teachers with important information and to allow them to interact with the administrator in charge of science instruction. The science department also had a monthly planning day. Department planning days were used to plan upcoming lessons, review state science standards, and provide ongoing professional development to support innovative academic vocabulary instruction. The biology teachers shared a common planning period and met daily to plan lessons, create innovative academic vocabulary activities related to interactive word walls, discuss upcoming investigations, create quizzes, and look at student achievement data from the previous week and/or Saturday school. At the end of each 9-week grading period, the science department joined the biology team for 1 day of planning. This allowed all teachers in the science department to experience the biology team's focused, intense planning structure. Department members also helped the biology teachers organize the 10-day biology STAAR EOC focused review.

Method

Context

A Grade 9–12 public high school, the intervention campus is part of a small school district located in a suburban setting in South Texas. Bordered by two large military bases and serving 11,700 students, this district has a predominantly Hispanic economically disadvantaged student population. Built in 1969, the high school has historically served economically disadvantaged Hispanic students. Table 1 contains data on enrollment history and ethnic distribution for the intervention campus collected by TEA across a 3-year period.

Measures

As noted, our research project was driven by concerns about the low science achievement of adolescent ELs and economically disadvantaged students on high-stakes tests. The high

Table 2. 2014 campus comparison.

Campus	Grade span	Number of students	Economically disadvantaged (%)	English leaners (%)	Mobility rate
Intervention	9–12	1,270	97.8	10.5	24.9
Comparison	9–12	1,028	91.8	10.2	21.4

school intervention campus in which the present study's professional development initiative took place was rated IR on the 2012–2013 school year Texas Academic Performance Report because of low scores on STAAR EOCs. IR indicates unacceptable performance and is assigned to campuses and districts that do not meet established TEA targets on required indices for which data are available. Required indices include STAAR results, graduation rates, graduation plan rates, and college and career readiness. The biology STAAR EOC is a distal high-stakes state exam administered to first-time test takers in the spring of their freshman year of high school. If students fail their first attempt, multiple retakes across multiple years are available. The biology STAAR EOC includes 54 multiple-choice questions measuring five reporting categories. Passing rates in the State of Texas for 2012–2015 were Unsatisfactory (scores below 37%, considered not passing), Satisfactory (scores between 37% and 82%, considered passing), and Advanced (scores 83% or better, considered passing). The biology STAAR EOC has been reported to have an external validity correlation measure of .70 to .79 (TEA, 2012). Readers interested in viewing state-released sample test items can do so on the TEA website (https://tea.texas.gov/student.assessment/staar/).

STAAR passing rates and graduation rates are two of four indices used to determine campus and district accountability ratings. In addition, passing the biology STAAR EOC is a high school graduation requirement. As a result, improving the biology STAAR EOC passing rates was a high priority for the campus and district. For these reasons we developed and implemented a professional development initiative, based on theory and research, with the intent of helping improve students' biology STAAR EOC passing rates at the Satisfactory level and with the intent of seeing changes at the Advanced level as well.

Data collection

Using publicly available state databases, we collected spring administration (only first-time test takers) biology STAAR EOC passing rates for the intervention campus from 2013 to 2015. We then collected spring administration (only first-time test takers) biology STAAR EOC passing rates for a comparison campus. Comparison campuses are schools assigned by TEA to a unique group of 40 other public schools (from anywhere in Texas) matching the other schools on the following characteristics: campus type, campus size, percentage of economically disadvantaged students, percentage of students with limited English proficiency, and mobility rates. From this group of 40 schools, we matched a school from the same region in Texas as the intervention campus (South Texas) based on state-reported school characteristics. This additional level of matching from within the same state region was intended to provide us with the best match for comparison based on the percentage of economically disadvantaged students (at least 90%) and the percentage of ELs (at least 10%). Table 2 presents demographic data for the intervention and comparison campuses for the first year of comparison (i.e., 2014; see "Analysis").

Analysis

Publicly available state data report the proportion of students by level of accomplishment. We therefore used difference in proportions tests to determine whether students showed positive achievement gains at the intervention campus as a result of the treatment.

Table 3. One-year (2013–2014) comparison of Satisfactory and Advanced passing rates within the intervention campus after receipt of the professional development initiative intervention.

Category	2013 baseline	2014 Intervention Year 1	z	d	95% confidence interval	
					Lower	Upper
Satisfactory						
Overall	57%	81%	−7.111***	0.644	0.463	0.825
Economically disadvantaged	56%	81%	−7.376***	0.667	0.485	0.848
EL	37%	63%	−7.125***	0.587	0.424	0.750
At risk	44%	75%	−8.654***	0.739	0.568	0.910
Advanced						
Overall	1%	4%	−2.632**	0.781	0.153	1.410
Economically disadvantaged	1%	5%	−3.2116*	0.910	0.294	1.526
EL	0%	0%				
At risk	1%	1%				

Note. 2013 $N = 375$, 2014 $N = 376$. d = standard mean difference; EL = English learner.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Difference in proportions tests utilize both percentages and sample sizes for two groups and calculate the probability level associated with the difference in proportions evidenced between the two groups. Based on the results of this test, the difference in proportions, or percentages, between two groups of respondents can be determined to be either statistically significant or nonsignificant. In addition, we calculated effect sizes (i.e., d = the standard mean difference) to determine the magnitude of the changes, if any. We specifically used a two-proportion z test to determine whether the difference between the two proportions was significant.

In addition, we used Cohen's (1969) scale to gauge the magnitude of the effects, with 0.2 = small, 0.5 = medium, and 0.8 = large. It is important to report effect sizes, such as Cohen's d, because they allow researchers to compare results across studies and to evaluate the replicability of new results (Thompson, 2008). Effect sizes also provide a sense of the magnitude and importance of a study's findings (American Psychological Association, 2010; Grissom & Kim, 2005).

We first analyzed data within the school by comparing 2013 passing rates (the baseline year before the intervention was implemented) and 2014 and 2015 passing rates (the years during which the intervention was implemented). We compared the Satisfactory and Advanced passing rates separately in order to detect any changes among the highest performers. We then analyzed data between schools by comparing test passing rates between the intervention campus and the comparison campus for the years during which the intervention was implemented (2014 and 2015).

Because we were especially interested in EL and economically disadvantaged students, we analyzed these demographics. According to TEA (2015), *ELs* (or *limited English proficient students*) are defined as students who use a language other than English as the primary language in their home and whose English language proficiency is determined to be limited, and *economically disadvantaged students* are defined as those students eligible for free or reduced-price meals under the National School Lunch and Child Nutrition Program. We also analyzed overall school passing rates to gauge how the intervention helped all students. Last, we considered the at-risk category in the data because many EL and economically disadvantaged students fall into this category, which means that they are at risk for not graduating from high school (TEA, 2015), a factor pertinent to the students in our study. In schools with high populations of ELs, students who fall into one of these

Table 4. Two-year (2013–2015) comparison of Satisfactory and Advanced passing rates within the intervention campus before and after receipt of the professional development initiative intervention.

Category	2013 baseline	2015 Intervention Year 2	z	d	95% confidence interval	
					Lower	Upper
Satisfactory						
Overall	57%	89%	−8.367***	0.997	0.746	1.248
Economically disadvantaged	56%	89%	−8.582***	1.020	0.768	1.271
EL	37%	64%	−6.516***	0.611	0.424	0.797
At risk	44%	86%	−10.343***	1.134	0.902	1.365
Advanced						
Overall	1%	8%	−4.463*	1.187	0.569	1.805
Economically disadvantaged	1%	8%	−4.463*	1.187	0.569	1.805
EL	0%	0%				
At risk	1%	3%	−1.823	0.617	−0.079	1.313

Note. 2013 $N = 375$, 2015 $N = 237$. d = standard mean difference; EL = English learner.
* $p < .05$. *** $p < .001$.

categories often fall into others. Each category is nonetheless distinct and important to pursue in terms of specific investigation versus assuming that all categories are the same (e.g., some ELs may not necessarily be at risk or economically disadvantaged even if the majority of ELs fall into these classifications).

Results

Research Question 1

In order to answer Research Question 1 (Do the standardized achievement passing rates of economically disadvantaged and EL students at the intervention campus show a statistically significant difference after students receive 1 year of the professional development initiative intervention? If so, what is the magnitude of the difference?), we conducted difference in proportions tests between the intervention campus's 2013 end-of-year biology STAAR EOC (preintervention baseline) and 2014 end-of-year biology STAAR EOC (1 year postintervention) passing rate exam data for Satisfactory and Advanced, reported in Table 3.

As can be noted in Table 3, there were substantial Satisfactory percentage passing rate increases from 2013 to 2014 for all student categories: 57% to 81% overall, 56% to 81% for economically disadvantaged, 37% to 63% for EL, and 44% to 75% for at risk. Statistically significant differences with effect sizes of medium to large magnitude were noted for all student categories at the Satisfactory level: overall ($z = -7.111$, $p < .001$, $d = 0.644$), economically disadvantaged ($z = -7.376$, $p < .001$, $d = 0.667$), EL ($z = -7.125$, $p < .001$, $d = 0.587$), and at risk ($z = -8.654$, $p < .001$, $d = 0.739$). Advanced percentage passing rates increased from 2013 to 2014 for the overall (1% to 4%) and economically disadvantaged (1% to 5%) categories but not for the EL (0% to 0%) or at-risk (1% to 1%) student categories. Statistically significant differences at the Advanced level were noted for the overall ($z = -2.632$, $p < .01$) and economically disadvantaged ($z = -3.2116$, $p < .05$) categories, and the effect sizes were of large magnitude ($d = 0.781$ and $d = 0.910$, respectively). No statistically significant differences were calculated for the EL and at-

Table 5. 2013 baseline comparison of Satisfactory and Advanced passing rates between the intervention and comparison campuses.

Category	Intervention baseline	Comparison	*z*
Satisfactory			
Overall	57%	68%	−2.864**
Economically disadvantaged	56%	68%	−3.097**
EL	37%	57%	−5.054***
At risk	44%	62%	−4.532***
Advanced			
Overall	1%	3%	−1.870
Economically disadvantaged	1%	3%	−1.870
EL	0%	2%	−2.750*
At risk	1%	1%	0.000

Note. Intervention $N = 375$, comparison $N = 274$. EL = English learner.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6. Comparison of Satisfactory and Advanced passing rates between the intervention and comparison campuses for Year 1 of the intervention (2014).

					95% confidence interval	
Category	2014 intervention	2014 comparison	*z*	*d*	Lower	Upper
Satisfactory						
Overall	81%	78%	0.983	0.102	−0.101	0.304
Economically disadvantaged	81%	77%	1.300	0.102	−0.101	0.304
EL	63%	38%	6.604***	0.563	0.394	0.732
At risk	75%	68%	2.053*	0.190	0.008	0.372
Advanced						
Overall	4%	3%	0.715	0.164	−0.288	0.616
Economically disadvantaged	4%	3%	0.715	0.164	−0.288	0.616
EL	0%	0%	0.000			
At risk	1%	0%	1.808			

Note. Intervention $N = 376$, comparison $N = 325$. d = standard mean difference; EL = English learner.
* $p < .05$. *** $p < .001$.

risk categories because there was no percentage change in their scores at the Advanced level.

Research Question 2

In order to answer Research Question 2 (Do the standardized achievement passing rates of economically disadvantaged and EL students at the intervention campus show a statistically significant difference after students receive 2 years of the professional development initiative intervention? If so, what is the magnitude of the difference?), we conducted difference in proportions tests between the intervention campus's 2013 end-of-year biology STAAR EOC (preintervention baseline) and 2015 end-of-year biology STAAR EOC (2 years postintervention) passing rate exam data for the Satisfactory and Advanced levels, reported in Table 4.

As can be noted in Table 4, there were substantial Satisfactory percentage passing rate increases from 2013 to 2015 for all student categories: 57% to 89% overall, 56% to 89% for economically disadvantaged, 37% to 64% for EL, and 44% to 86% for at risk. Statistically significant differences with effect sizes of large magnitude (except for EL, which had an effect size of medium magnitude) were noted for all student categories at the Satisfactory level: overall ($z = -8.367$, $p < .001$, $d = 0.997$), economically disadvantaged ($z = -8.582$,

Table 7. Comparison of Satisfactory and Advanced passing rates between the intervention and comparison campuses for Year 2 of the intervention (2015).

					95% confidence interval	
Category	2015 intervention	2015 comparison	*z*	*d*	Lower	Upper
Satisfactory						
Overall	89%	84%	1.686	0.239	−0.040	0.517
Economically disadvantaged	89%	83%	1.989*	0.279	−0.002	0.555
EL	64%	55%	2.131*	0.207	0.016	0.397
At risk	86%	78%	2.397*	0.303	0.053	0.553
Advanced						
Overall	8%	3%	2.640*	0.570	0.131	1.009
Economically disadvantaged	8%	2%	3.351*	0.800	0.295	1.303
EL	0%	0%	0.000			
At risk	2%	0%	2.533*			

Note. Intervention N = 237, comparison N = 318. d = standard mean difference; EL = English learner.
* p < .05.

$p < .001$, $d = 1.020$), EL ($z = -6.516$, $p < .001$, $d = 0.611$), and at risk ($z = -10.343$, $p < .001$, $d = 1.134$). In addition, there were Advanced percentage passing rate increases for students overall (1% to 8%) and in the economically disadvantaged (1% to 8%) and at-risk (1% to 3%) categories but not the EL category (0% to 0%). Statistically significant differences with effect sizes of large magnitude at the Advanced level were noted for the overall ($z = -4.463$, $p < .05$, $d = 1.187$) and economically disadvantaged ($z = -4.463$, $p < .05$, $d = 1.187$) student categories. No statistical differences were calculated or noted for students in the EL and at-risk categories, though the at-risk category had an effect size of medium magnitude ($d = 0.617$) at the Advanced level.

Research Question 3

In order to answer Research Question 3 (Are there statistically significant differences between the standardized achievement passing rates of economically disadvantaged and EL students at the intervention campus and a comparison campus? If so, what are the differences in the magnitudes?), we first analyzed the 2013 data for the two schools to ensure that the baseline did not have any form of unfair advantage for the intervention over the control. Table 5 shows a comparison of the 2013 data for the intervention and comparison campuses.

As can be noted in Table 5, the comparison campus had slightly higher percentage scores on all measures than the intervention campus at baseline (e.g., 68% vs. 57% overall, 57% vs. 37% for EL, respectively) except for the at-risk category at the Advanced level (1% and 1%, respectively). In addition, statistically significant differences in favor of the comparison campus were noted for all categories at the Satisfactory level ($p < .01$ and $p < .001$) and specifically for the EL category at the Advanced level ($z = -2.750$, $p < .05$). The intervention campus therefore did not have an advantage over the control campus at baseline.

We then conducted difference in proportions tests between the intervention campus's 2014 and 2015 (postintervention) and the comparison campus's 2014 and 2015 end-of-year biology STAAR EOC passing rate exam data for Satisfactory and Advanced. Table 6 presents the comparison for the 2014 data, after 1 year of intervention at the intervention campus.

As can be noted in Table 6, the intervention campus had higher percentages than the comparison campus for all categories (e.g., Satisfactory overall, 81% vs. 78%; Satisfactory

EL, 63% vs. 38%; Advanced overall, 4% vs. 3%; Advanced at risk, 1% vs. 0%, respectively) except the EL category at the Advanced level, for which scores were equal (0% and 0%, respectively). Statistically significant differences were noted only for the EL ($z = 6.604$, $p < .001$) and at-risk ($z = 2.053$, $p < .05$) categories at the Satisfactory level; the effect sizes were medium ($d = 0.563$) to small ($d = 0.190$), respectively.

Table 7 shows the comparison for the 2015 data, after 2 years of intervention at the intervention campus. As can be noted in Table 7, the intervention campus showed higher percentages for all categories (e.g., Satisfactory overall, 89% vs. 84%; Satisfactory EL, 64% vs. 55%; Advanced overall, 8% vs. 3%; Advanced at risk, 2% vs. 0%, respectively) except EL at the Advanced level, for which scores were equal (0% and 0%, respectively). Statistically significant differences were noted for all categories in favor of the intervention campus (all at $p < .05$) except for the Satisfactory level overall and the EL category at the Advanced level. Effect sizes were small overall, ranging from .239 to .303 on the low end; however, it is worth noting that the effect size for the overall category at the Advanced level was of medium size ($d = 0.570$) and the effect size for the economically disadvantaged category at the Advanced level was large ($d = 0.800$).

Conclusion

In Texas, high-stakes, distal end-of-course exams are high school graduation gatekeepers. ELs, economically disadvantaged students, and at-risk students have well-documented achievement gaps in academic language (Cummins, 1981; Fang, 2006; Quinn et al., 2012) that create unique challenges on high-stakes assessments (Kieffer et al., 2009). The purpose of this study was to design a professional development initiative, based on theory and empirical evidence, to help EL and economically disadvantaged high school students who are at risk for academic failure to pass high-stakes assessments. Scarcella (2003) proposed that academic English is required for students to be successful in school settings. Gee (2004) and Lemke (1990) suggested that success in science requires students to understand the unique text structures and vocabulary associated with scientific discourse. The professional development initiative included a combination of elements that addressed these points. Purposeful planning identified essential standards-aligned academic vocabulary and ensured that lessons and assessments also included standards-aligned academic vocabulary. Innovative academic vocabulary instruction was student centered and connected academic vocabulary with inquiry investigations and in-class activities. Saturday school, structured tutoring, and a focused review provided additional exposure to academic vocabulary and strategically targeted content gaps.

Teachers at the intervention campus received 2 years of ongoing professional development training to help their students simultaneously learn the academic language and content of science. Our results indicate that overall the professional development initiative was beneficial to the students within the intervention campus and showed promising effects compared to a matched comparison campus. Major findings are summarized below, with discussion based on the previous literature, a consideration of the limitations of the study, and recommendations for future research and practice.

Within-school findings

After 1 year of implementation, the professional development initiative benefited student passing scores on the high-stakes state standardized biology STAAR EOC. Note that statistically significant results and medium effect sizes were detected for the Satisfactory passing rate of students overall, students who were economically disadvantaged, and students who were ELs. In the Advanced passing rate category (i.e., the highest passing rate), statistically significant and large effect sizes were noted for students overall and for students who were economically disadvantaged, although no changes were noted at this level for students in the EL or at-risk categories.

Findings were similar after 2 years of the intervention within the school, but the effect sizes were large. In addition, although no statistical significance was noted for the Advanced passing rate for ELs and at-risk students after 2 years, there was a change in the passing rate for the at-risk category with a medium effect size. These changes indicate that time was an important factor in increasing the magnitude of the effect of the intervention. As previous researchers in the field of ELs and science have noted, teachers need time to learn and refine the strategies and skills gained through ongoing professional development (Adamson, Santau, & Lee, 2013; Santau, Secada, Maerten-Rivera, Cone, & Lee, 2010). The lack of change in the EL scores at the highest Advanced rate informs us that ELs may need even more time to develop the nuanced academic language of science needed to pass with highest distinction given that it takes ELs up to 7 years to master academic language (Collier & Thomas, 1989), when they are exposed to academic language at all (Scarcella, 2003).

Overall, the findings within the school across 2 years align with previous research that found that interventions fusing science and literacy increase students' outcome measures pre- and postintervention with samples of elementary EL and economically diverse, at-risk students (e.g., Lee et al., 2005; Shaw et al., 2014). Our findings, however, contribute to much-needed research in the area of at-risk adolescent student science outcomes (August et al., 2016; Nation et al., 2010).

Between-school findings

After 1 year of the professional development implementation, statistically significant differences were noted for the EL and at-risk student Satisfactory passing rates. Effect sizes were small in all categories except for the EL category, with a medium effect size. After 2 years of implementation, statistically significant results were noted for more subcategories of students, including economically disadvantaged, EL, and at-risk students at the Satisfactory and Advanced levels (with ELs not changing, again, at the Advanced level). These findings reinforce the importance of time as an important factor in noting effects for at-risk students in professional development intervention studies (Adamson et al., 2013; Santau et al., 2010). Although it was promising to see statistical significance at the Advanced level for economically disadvantaged and at-risk students after 2 years of the intervention, it would have been more promising to see changes for ELs at this level as well. Again, it is likely that time is a factor needed for ELs to show changes at this level (Collier & Thomas, 1989).

Our findings align with previous studies that have implemented interventions fusing science and literacy strategies to help increase scores for at-risk, economically disadvantaged, and EL students and compared findings across schools (e.g., Ardasheva & Tretter, 2017; August et al., 2016; Lara-Alecio et al., 2012; Lee et al., 2005; Maerten-Rivera et al., 2016; Shaw et al., 2014). However, our study is unique in that it found statistically significant results for high-stakes standardized science assessment scores.

Limitations

Though every attempt was made to consider longitudinal change in the pre/post comparison and carefully match schools for added comparison, causality cannot be inferred because the study lacked a randomized design. What can be inferred, however, is that implementing theory- and research-based professional development in a highly at-risk high school had promising effects on student scores on a high-stakes standardized science assessment. If anything, this study is a first step in affirming the benefits of language-infused standards-based science interventions with a focus on evidence-based professional development for at-risk students' achievement on standardized science assessments.

Recommendations for future research and practice

Certainly, future research should consider randomized treatment and control intervention designs with analyses of secondary high-stakes science assessments. As noted previously, studies to date with quasi-experimental and pre/post designs considering adolescent at-risk populations have found positive treatment effects on proximal (i.e., curriculum-based) measures of student science achievement (Ardasheva & Tretter, 2017; August et al., 2009; Fang & Wei, 2010; Lara-Alecio et al., 2012; Tong et al., 2014). To date, however, none of these studies have found treatment effects on more distal (i.e., high-stakes) science standardized achievement scores for at-risk students at the secondary level. For this reason, the present study adds much-needed research to the field and prompts future researchers to consider the types (and length) of professional development initiatives needed in order to help economically disadvantaged, EL, and at-risk adolescent students achieve in science.

Given these results, researchers, science methods instructors, and practitioners should continue to critically consider, implement, and examine the types of pedagogical practices and professional development that promote the achievement of secondary at-risk students at all levels, including distal high-stakes assessments. The professional development initiative described in this study is a good place to begin. We recommend that teachers use a system that supports standards-based instruction and experience using standards to identify essential academic vocabulary (i.e., purposeful planning). We support inquiry instruction partnered with literacy-infused, student-centered lessons (i.e., innovative vocabulary instruction) and providing teachers with space and time to review and reinforce learning (i.e., Saturday school/structured tutoring and focused review).

Though researchers, instructors, and practitioners should always adapt their curriculum and pedagogy to their specific contexts, this initiative provides a solid roadmap based on theory and research for creating the types of pedagogical practices that afford all adolescent students the opportunity to learn and succeed in science.

References

Adamson, K., Santau, A., & Lee, O. (2013). The impact of professional development on elementary teachers' strategies for teaching science with diverse student groups in urban elementary schools. *Journal of Science Teacher Education*, *24*(3), 553–571. doi:10.1007/s10972-012-9306-z

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York, NY: Oxford Press.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Ardasheva, Y., & Tretter, T. R. (2017). Developing science-specific, technical vocabulary of high school newcomer English learners. *International Journal of Bilingual Education and Bilingualism*, *20*(3), 252–271. doi:10.1080/13670050.2015.1042356

August, D., Artzi, L., & Barr, C. (2016). Helping ELLs meet standards in English language arts and science: An intervention focused on academic vocabulary. *Reading & Writing Quarterly*, *32*(4), 373–396. doi:10.1080/10573569.2015.1039738

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness*, *2*(4), 345–376. doi:10.1080/19345740903217623

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc. doi:10.2167/le129b.0

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E Instructional Model: Origins, effectiveness, and applications. BSCS Executive Summary*. Colorado Springs, CO: Biological Sciences Curriculum Study (BSCS). Retrieved from https://bscs.org/bscs-5e-instructional-model

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. San Diego, CA: Academic Press.

Collier, V. P., & Thomas, W. P. (1989). How quickly can immigrants become proficient in school English? *Journal of Educational Issues of Language Minority Students*, *5*, 26–38.

Cuevas, P., Lee, O., Hart, J., & Deaktor, R. (2005). Improving science inquiry with elementary students of diverse backgrounds. *Journal of Research in Science Teaching*, *42*(3), 337–357. doi:10.1002/tea.20053

Cummins, J. (1981). Empirical and theoretical underpinnings of bilingual education. *The Journal of Education*, *163*(1), 16–29. doi:10.1177/002205748116300104

Echevarria, J., Richards-Tutor, C., Canges, R., & Francis, D. (2011). Using the SIOP model to promote the acquisition of language and science concepts with English learners. *Bilingual Research Journal*, *34*(3), 334–351. doi:10.1080/15235882.2011.623600

Fang, Z. (2006). The language demands of science reading in middle school. *International Journal of Science Education*, *28*(5), 491–520. doi:10.1080/09500690500339092

Fang, Z., & Wei, Y. (2010). Improving middle school students' science literacy through reading infusion. *Journal of Educational Research*, *103*(4), 262–273. doi:10.1080/00220670903383051

Gee, J. P. (2004). Language in the science classroom: Academic social languages as the heart of school-based literacy. In *Establishing scientific classroom discourse communities* (pp. 19–37). Lawrence Erlbaum Associates. doi:10.4324/9781410611734

Graves, M. (2006). *The vocabulary book: Learning and instruction*. New York, NY: Teachers College Press.

Graves, M., Baumann, J., Blachowicz, C., Manyok, P., Bates, A., Cieply, C., & Von Gunten, H. (2014). Words, words, everywhere, but which ones do we teach?. *The Reading Teacher*, *67*(5), 333–346. doi:10.1002/trtr.1228

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Erlbaum.

Jackson, J. (2018). Build an interactive word wall: Connecting scientific concepts with academic vocabulary. *The Science Teacher*, *85*(1), 42–46.

Jackson, J., & Ash, G. (2012). Science achievement for all: Improving science performance and closing achievement gaps. *Journal of Science Teacher Education*, *23*(7), 723–744. doi: 10.1007/s10972-011-9238-z

Jackson, J., & Durham, A. (2016). Put your walls to work: Planning and using interactive word walls to support science and reading instruction. *Science and Children*, *54*(3), 78–84.

Jackson, J., & Narvaez, R. (2013). Interactive word walls: Create a tool to increase science vocabulary in five easy steps. *Science and Children*, *51*(1), 42–49.

Jackson, J., Wise, E., Zurbuchen, K., & Gardner, N. (2017). Interactive word walls: Visual scaffolds that transform vocabulary instruction. *Science Scope*, *40*(9), 72–79.

Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research*, *79*(3), 1168–1201. doi:10.3102/0034654309332490

Kohlhaas, K., Lin, -H.-H., & Chu, K.-L. (2010). Disaggregated outcomes of gender, ethnicity, and poverty on fifth grade science performance. *Research in Middle Level Education Online*, *33*(7), 1–12. doi:10.1080/19404476.2010.11462070

Lara-Alecio, R., Tong, F., Irby, B. J., Guerrero, C., Huerta, M., & Fan, Y. (2012). The effect of an instructional intervention on middle school English learners' science and English reading achievement. *Journal of Research in Science Teaching*, *49*(8), 987–1011. doi:10.1002/tea.21031

Lee, O., Deaktor, R. A., Hart, J. E., Cuevas, P., & Enders, C. (2005). An instructional intervention's impact on the science and literacy achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching*, *42*(8), 857–887. doi:10.1002/tea.20071

Lemke, J. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex Publishing Corporation.

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, *47*(3), 596–632. doi:10.3102/0002831209355469

Lynch, S., Kuipers, J., Pyke, C., & Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse students: Results from a planning grant. *Journal of Research in Science Teaching*, *42*(8), 912–946. doi:10.1002/tea.20080

Maerten-Rivera, J., Ahn, S., Lanier, K., Diaz, J., & Lee, O. (2016). Effect of a multiyear intervention on science achievement of all students including English language learners. *The Elementary School Journal*, *116*(4), 600–624. doi:10.1086/686250

Mangiante, E. S. (2018). Planning for reform-based science: Case studies of two urban elementary teachers. *Research in Science Education*, *48*(1), 207–232. doi:10.1007/s11165-016-9566-2

McFarland, J., Hussar, W., DeBrey, C., Snyder, T., Wang, X., Wilkinson-Flicker, S., & Hinz, S. (2017). *The condition of education 2017* (NCES 2017-144). Washington, DC: U.S. Department of Education. doi:10.1037/e492172006-019

Muller, P. A., Stage, F. K., & Kinzie, J. (2001). Science achievement growth trajectories: Understanding factors related to gender and racial-ethnic differences in precollege science achievement. *American Educational Research Journal*, *38*(4), 981–1012. doi:10.3102/00028312038004981

Nation, K., Cocksey, J., Taylor, J. S., & Bishop, D. V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry*, *51*(9), 1031–1039. doi: 10.1111/j.1469-7610.2010.02254.x

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press. Retrieved from www.nextgenscience.org/

Organisation for Economic Co-operation and Development. (2014). *Education at a glance 2014: OECD indicators*. Paris, France: Author. doi:10.1787/eag-2014-en

Quinn, H., Lee, O., & Valdés, G. (2012). *Language demands and opportunities in relation to Next Generation Science Standards for English language learners: What teachers need to know*. Stanford, CA: Stanford University, Understanding Language Initiative at Stanford University (ell.stanford.edu).

Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, *39*(5), 369–393. doi:10.1002/tea.10027

Santau, A. O., Secada, W., Maerten-Rivera, J., Cone, N., & Lee, O. (2010). US urban elementary teachers' knowledge and practices in teaching science to English language learners: Results from the first year of a professional development intervention. *International Journal of Science Education*, *32*(15), 2007–2032. doi:10.1080/09500690903280588

Scarcella, R. (2003). *Academic English: A conceptual framework* (Technical Report 2003-1). Berkeley, CA: University of California Linguistic Minority Research Institute. (ERIC Document Reproduction Service No. ED531333). Retrieved from http://escholarship.org/uc/item/6pd082d4

Shaw, J. M., Lyon, E. G., Stoddart, T., Mosqueda, E., & Menon, P. (2014). Improving science and literacy learning for English language learners: Evidence from a pre-service teacher preparation intervention. *Journal of Science Teacher Education*, *25*(5), 621–643. doi:10.1007/s10972-013-9376-6

Texas Education Agency [TEA]. (2012). *STARR ECO external validity studies*. Retrieved from https://tea.texas.gov/staar/vldstd.aspx

Texas Education Agency [TEA]. (2015). *Glossary of terms, 2007-08: Division of research and analysis*. Retrieved from https://rptsvr1.tea.texas.gov/acctres/gloss0708.html

Thompson, B. (2008). *Foundations of behavioral statistics: An insight-based approach*. New York, NY: Guilford Press.

Tong, F., Irby, B. J., Lara-Alecio, R., Guerrero, C., Fan, Y., & Huerta, M. (2014). A randomized study of a literacy-integrated science intervention for low-socio-economic status middle school students: Findings from first-year implementation. *International Journal of Science Education*, *36*(12), 2083–2109. doi:10.1080/09500693.2014.883107