Spring 2019

Using Culturally Responsive Visual Aids to Increase English Learners' Speaking Assessment Scores

Eric A. Martin
California State University, Monterey Bay

Using Culturally Responsive Visual Aids to Increase English Learners' Speaking Assessment Scores

Eric A. Martin

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Arts in Education

California State University, Monterey Bay

May 2019

Using Culturally Responsive Visual Aids to Increase

English Learners' Speaking Assessment Scores

Eric A. Martin

APPROVED BY THE GRADUATE ADVISORY COMMITTEE

_____

Kerrie Chitwood, Ph.D.
Advisor and Program Coordinator, Master of Arts in Education

_____

Dennis Kombe, Ph.D.
Advisor, Master of Arts in Education

Approved by the Dean of Graduate Studies 05/20/2019

Kris Roney, Ph.D. Associate Vice President
Academic Programs and Dean of Undergraduate & Graduate Studies

Abstract

The number of California English Language Learners (ELLs) continues to rise. While the population has increased, assessment scores measuring English speaking proficiencies have not. Culturally Responsive Visual Aids (CRVA) have been used to increase ELLs' achievement on speaking assessments by bridging two bodies of research: culturally responsive pedagogy and visual aid scaffolding. This study used a two-group non-equivalent pre-post quasi-experimental design to compare ELLs' scores on speaking assessments. The control group (n=16) received weekly speaking practice and the treatment group (n=14) received daily, direct instruction on how to increase scores on speaking assessments using CRVA. Independent sample t-tests were completed to examine the difference in scores between the two groups. The results suggest that using CRVA for daily speaking practice increased scores on speaking assessments given to ELLs; however, these scores were not statistically significant. Future research should examine whether CRVA or daily direct instruction for speaking assessment has a larger impact on ELL achievement.

Keywords: Culturally Responsive Visual Aids, Speaking Assessment, English Language Learners, Culturally Responsive Pedagogy, Culturally Relevant Assessment

Acknowledgements

I would like to thank my wife, Alanna, for sustaining me for the last ten years. Her encouragement and patience made this possible. Thank you to my beautiful daughter, Camille Amelia, for making this process much more bearable with her unending and effervescent love, joy, and support. I would like to thank my parents, Craig and Ruth Martin, for their support. In so many ways, they made this accomplishment possible. A special thank you to all of my in-laws for their emotional support and babysitting expertise for the last year: Ed and Dorothy Sinclair, Stacie and Matt Silva, Theresa and Trevor Porter, and Ellen and Dave Johnson. Thank you to my close friends Brad, Matt, and Chris for always having my back. Thank you to Erin Gemar from South County Cal-SOAP for inspiring me to pursue a career as an educator. Thank you to my team at San Benito High School. Lastly, thank you to Dr. Kerrie Chitwood, Dr. Erin Ramirez, and Dr. Dennis Kombe for their professionalism and encouragement through the last year.

Table of Contents

Using Culturally Responsive Visual Aids to Increase English Learners' Speaking Assessment Scores

Literature Review

English Language Learners (ELLs) are one of the lowest performing groups of students in the United States (Wolf, Herman, & Dietel, 2010). Students are classified as ELLs based on the results of the English Language Proficiency Assessments for California (ELPAC), and demonstrate emerging in proficiencies in the English language in the domains of reading, writing, speaking, and listening (CA Department of Education, 2018a). The ELPAC is a required state test given to students who are learning English as a second language to check for English language proficiency (CA Department of Education, 2018b). According to the California Department of Education (2018c), more than one million ELLs are enrolled in California schools, many of whom are struggling to meet grade-level standards. For example, in 2017, high school ELLs took the ELPAC and over 50% of the ninth and tenth graders assessed were categorized as level one (i.e., minimally developed), the lowest performance proficiency (CA Department of Education, 2018a). According to the California Department of Education (2018) the four levels include level one (i.e., minimally developed), level two (i.e., somewhat developed), level three (i.e., moderately developed), and level four (i.e., well developed). Level one ELLs might know some English words and phrases, but need significant assistance using English to communicate and acquire new knowledge at school difficult to succeed in general education settings (CA Department of Education, 2018a).

Out of the four domains measured (i.e., reading, writing, speaking, and listening), students performed the lowest on the speaking assessments (CA Department of Education, 2018a). Some researchers have attributed low levels of language acquisition and performance to

the affective filter hypothesis, the lack of scaffolds, and pushing students too far beyond current proficiencies (Krashen, 1982; Swain, 1993; Vygotsky, 1978). Researchers have also linked low performance on language assessments to culturally irrelevant testing materials (Prosser & Solano-Flores, 2010; Solano-Flores, 2008). While many students are able to communicate in spoken English, the fact that so many are earning the minimally developed descriptor on assessments is a cause for concern. There is a clear need for research on ways to raise ELLs' performance on speaking assessments. To understand how educators might be able to help ELLs, the processes of language acquisition and barriers to language acquisition must be understood.

English Language Learners and Language Acquisition Theory

The processes for acquiring a second language have been widely researched, and in some cases, even disputed. Krashen (1982), for example, introduced the theory of second language acquisition, which includes the Input Hypothesis. According to the Input Hypothesis, ELLs acquire language when challenged with new language just beyond the current level of proficiency. For some ELLs, everything in English is language beyond the current level of proficiency because they are learning the basics along with academic language. This can be difficult situation for a high school student who is new to the country and the language altogether. Language acquisition, according to this hypothesis, is input based. Students are exposed to new academic language through various modalities with the expectation that they will acquire it, however; there is no output expectation (i.e., writing and speaking). For example, a high school teacher may use direct instruction to introduce and teach content-specific or general academic vocabulary, or may present students with a challenging reading task. The issue is that the direct instruction is not directly followed by student language production. Krashen (1982) suggested that if students are receiving higher-level language instruction (i.e., input), then ELLs

will begin to acquire and internalize higher level language. This theory of language acquisition is supported by Vygotsky's (1978) theory known as the Zone of Proximal Development.

Language acquisition and the zone of proximal development. Vygotsky's (1978) Zone of Proximal Development (ZPD) suggests that students can acquire language and knowledge when pushed just out of their comfort zones if supports are provided. The ZPD is the gap between what the learner can and cannot do independently. Learning a second language is difficult to do without help; therefore, supports need to be provided to help ELLs succeed. Vygotsky asserts that the ZPD is the optimal place for learning to take place; however, educators are often guilty of pushing students beyond the ZPD (Vygotsky, 1978).

Language acquisition and affective filter. If challenged beyond the ZPD, ELLs can experience fear, low self-confidence, anxiety, and embarrassment when learning a new language (Krashen, 1982). Additionally, the language used in assessments can cause negative emotions, which make it difficult to perform (Solano-Flores, 2008). These negative emotions can become a barrier to acquiring new language. This theory is known as the Affective Filter Hypothesis (Krashen, 1982). The Affective Filter Hypothesis suggests that language acquisition has psychological constructs, which can inhibit or facilitate language acquisition. When the affective filter is high, students experience stress, low self-efficacy, and may be reluctant to speak out in classrooms. If students are reluctant to speak out in the classroom, then they are not practicing enough to perform well on speaking assessments.

As Swain (1993) points out, when educators fail to provide ELLs with adequate practice in speaking in the classroom, they inhibit ELL's language acquisition altogether. Moreover, it is impossible to measure the acquisition of English as a second language without giving students multiple opportunities to actually produce the second language. Furthermore, Swain (1993)

argues that students need to be given multiple exposures to content and opportunities to produce language, verbally and in writing, in order to internalize new language. This is known as the Output Hypothesis (Swain, 1993).

Language acquisition and output hypothesis. Swain (1993) developed the Output Hypothesis, which suggests that language cannot be acquired only by input (i.e., reading and listening), but by the process of oral production and output (i.e., writing and speaking). Multiple exposures to speaking tasks are critical in the acquisition of language. Krashen (1998) refuted this theory stating that being forced to speak out loud does not lead to language acquisition. Swain (1993) was not suggesting that parroting the teacher or instructor would lead to acquisition, but that output is tantamount to input for language acquisition. While Krashen disagreed, other theorists (i.e., Bruner, 1978; Sinclair, 1987; Wright, 1989) support Swain's hypothesis that output is key for language acquisition, but with a new element: *scaffolds*.

Scaffolding language acquisition. According to Bruner (1978), scaffolding refers to a variety of temporary supports (i.e., graphic organizers, videos, and photos) in the classroom used to aid students in accomplishing tasks, reaching goals, and acquiring new knowledge. All learners, including ELLs, need to struggle to some extent, in the gap between prior knowledge and the knowledge to be acquired (i.e., ZPD). Scaffolding tools can be used to aid their struggles to acquire new knowledge and new language. It is important to note that scaffolding tools can be used for input (i.e., reading and listening), but can, and must, also be used to support output (i.e., speaking and writing), especially for ELLs. When students are accorded multiple opportunities to produce language, more language is acquired (Swain, 1993). If these multiple attempts are aided by temporary supports, then students will acquire new language. It is important to scaffold input and output for language acquisition but research needs to be done on the effect that scaffolds

have on the assessment of ELLs' language acquisition. While there are various types of scaffolds, one type of scaffold which has been used with ELLs are visual aids.

Visual Aid Scaffolds for Language Acquisition and Assessment

Research has shown how visual aids can be used to increase language acquisition and production for ELLs (Nobuyoshi & Ellis, 1993; Sinclair, 1987; Wright, 1989). Sinclair (1987) was the first to introduce visual aids as an educational scaffolding tool. Visual aids can be, but are not limited to, a visual representation or an image, which is painted, drawn, or photographed. When visual aids are used, they can lead to higher levels of student performance (Sinclair, 1987). For example, Nobuyoshi and Ellis (1993) conducted a quantitative study of ELLs in Tokyo using pictures (i.e., visual aids) for language production. Through research, they found that students acquired more language when visual aids were used and when students were pushed to produce verbally. Moreover, results indicated that oral production is key to language acquisition; therefore, all scaffolding provided for language acquisition in the classroom should be accompanied by visual aids (Nobuyoshi & Ellis, 1993).

Furthermore, Wright (1989) also found that the use of visual aids made students more motivated to take part in speaking activities. Students who are motivated to speak in class may increase in the quantity of output, but not in the quality of spoken language. More recently, educators have found that the use of relevant high-quality photos and illustrations can support students in linking the content to actual language production (Echevarria, Vogt, & Short, 2013). In addition, the use of quality visuals paired with a written picture description has been shown to be a useful scaffold for ELLs. For example, researchers Lavalle and Briesmaster (2017) conducted a mixed methods study with ELLs in Chile. They found that the use of picture descriptions enhanced oral communication skills amongst language learners aged thirteen and

fourteen. Specifically, students increased scores on pronunciation, vocabulary, and grammar (Lavalle & Briesmaster, 2017). If students increase their scores on pronunciation, vocabulary, and grammar, it is likely they will earn higher scores on standardized spoken language assessments. Furthermore, Philominraj, Jeyabalan, and Vidal-Silva (2017) found that ELLs were encouraged by visual learning and concluded that visuals are essential to language acquisition and output.

Although Wright's (1989) research shows that students may be more motivated to contribute in the classroom, there is little research on the impact visual aid scaffolding tools have on the quality of language produced. Similarly, additional studies (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Wright, 1989) indicate that students might be able to acquire new language with the use of visual aids; however, it does not speak to the language quality. More importantly, the research does not show how visual aid scaffolding tools help ELLs perform better on speaking assessments. Visual aid scaffolds are important in the acquisition of a second language; yet , researchers have recently theorized that cultural relevance might be one of the missing components in the assessment of ELLs (Prosser & Solano-Flores, 2010; Solano-Flores, 2008).

Culturally Responsive Pedagogy for Language Acquisition

ELLs struggle with second language assessments because much of the content used in the assessments is culturally irrelevant (Solano-Flores, 2008). Solano-Flores and Nelson-Barber (2001) argued that culturally relevant assessments reflect a multi-cultural perspective by including the values, beliefs, experiences and prior knowledge of the students being assessed. Culturally irrelevant tests are as much a test on specific content as they are on the second language, and therefore, lack validity in the assessment of language acquisition (Basterra,

Trumbull, & Solano-Flores, 2011). In California, there are over one million ELLs with various ethnic and cultural backgrounds (CA Department of Education, 2018c). To help ELLs better understand the assessments they are given, the content must be relevant to their diverse lives and experiences.

Culturally Responsive Pedagogy (CRP) is a relatively new concept and theory, which suggests that the use of CRP increases achievement for students whose backgrounds are different from the instructor. Cochran-Smith, Davis, and Fries (2003) define CRP as a pedagogical method that recognizes and incorporates students' cultural background in the classroom as a means of delivering content. According to Gay (2000), ELLs and other emerging students greatly benefit when teachers use the unique cultural perspectives and experiences of diverse students to teach new concepts and skills. If educators can connect to the prior knowledge of ELLs in the classroom, student self-efficacy, or their belief that they can succeed, will increase. ELLs' cultures are assets in the classroom and should be used by the teacher in connecting prior knowledge to new knowledge (Trueba, Moll, Diaz, & Diaz, 1984).

It stands to reason that if students are more familiar with the content, and in some cases, if ELLs are the *content experts* in the room, they may be more motivated and more equipped to contribute to high quality output during classroom speaking activities. If teachers provide students with multiple exposures to visual aids and multiple opportunities to engage in speaking activities while in the classroom, they may be better prepared for assessments (Wright, 1989). Moreover, if the assessments are culturally relevant, performance will increase (Basterra et al., 2011). CRP might seem like the latest educational fad; however, culturally relevant instruction is not just a trend. CRP is a lasting and critical crux in closing the achievement gap for ELLs (Cochran-Smith et al., 2003).

Culturally Responsive Visual Aids

To best support ELLs, oral language skills must be directly tethered to cultural inclusivity through visual scaffolding. For example, Samson and Collins (2012) argued that there is a clear need to emphasize the connection between oral language skills and culturally inclusive teaching methods. Furthermore, research shows that visual aid scaffolding tools help ELLs produce oral language (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Wright, 1989). Research shows that CRP helps ELLs connect with content (Cochran-Smith et al., 2003; Gay, 2000; Trueba et al., 1984). These two bodies of research have not been combined to examine if culturally responsive scaffolding can be used to increase scores on assessments given to ELLs. The research suggests that the intersection between culturally responsive pedagogy and visual aid scaffolding tools used in assessments will increase the quality of language production during speaking assessments given to English Language Learners (Basterra et al., 2011; Cochran-Smith et al., 2003; Gay, 2000).

Method

The purpose of this study was to examine whether the intersection between culturally responsive pedagogy (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984) and visual aids (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Sinclair, 1987; Wright, 1989) increased high quality language production during speaking assessments given to ELLs. Specifically, this study explored whether the regular exposure to Culturally Responsive Visual Aids (CRVA) and the direct instruction on the use of CRVA increased the quality of language produced during speaking assessments (i.e., ELPAC) given to high school ELLs. A CRVA is defined by the researcher as any visual representation of culturally relevant content (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984), including

but not limited to photographs, videos, paintings, drawings, or slideshows, used as a scaffolding tool (Bruner, 1978) to aid instruction and student assessment.

Research Question

Does regular exposure and direct instruction on the use of culturally responsive visual aids (CRVA) increase the quality of speaking performance of high school ELLs during speaking tasks as measured by the ELPAC Speaking Performance Task?

Hypothesis

Based on the research, the hypothesis for this study was that the regular exposure and direct instruction on the use of CRVA would increase the quality of speaking performance of ELLs on speaking assessments. Visual aids assist in the production of oral language from ELLs (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Sinclair, 1987; Wright, 1989). CRP assists in language acquisition and production for ELLs (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984). Therefore, the research suggested that combining visual aids and CRP would increase the quality of speaking performance, as measured by the ELPAC Speaking Performance Task (ELPAC, 2018).

Research Design

This study used a two-group non-equivalent pre-post quasi-experimental design to assess the impact of CRVA on ELLs' performance on speaking assessments. Using a treatment group (i.e., the group that gets the intervention) and a control group (i.e., the group that gets regular instruction) aided in determining the effectiveness of the intervention. Both groups took a pretest before the intervention began and a posttest after the intervention. During the intervention, the treatment group received instruction with the addition of the intervention (i.e., the daily exposure and direct instruction on how to use CRVA for speaking assessments); whereas, the control

group received normal instruction. The control group had normal exposure (i.e., weekly) to the CRVA, however, did not receive any direct instruction or feedback on how to use the stimuli.

Independent variable. The independent variable in this study was regular exposure (i.e., daily) and direct instruction on the use of CRVA. For the purposes of this study, CRVA was defined by the researcher as any visual representation of culturally relevant content (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984), including but not limited to: photographs, videos, paintings, drawings, or slideshows, used as scaffolds (Bruner, 1978) to aid instruction. Visual aids assist in the production of oral language from ELLs (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Sinclair, 1987; Wright, 1989). Visual aids include but are not limited to: a visual representation of an image, drawing, painting, photograph, and video. (Sinclair, 1987). Similarly, CRP assists in language acquisition for ELLs (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984). Therefore, the intervention sought to explore the intersection of these two theories, and the impact of CRVA on ELLs' performance on speaking assessments.

Dependent variable. The dependent variable in this study was the quality of speaking performance of ELLs on a speaking assessment. Students were asked by the researcher to speak out loud and describe a visual aid (i.e., CRVA). This variable was operationalized using the Speaking Performance rubric (see Appendix A) used for the ELPAC (ELPAC, 2018). Specifically, the students completed the task type called "Talk about a Scene" (see Appendix B) in which students verbally described a CRVA (ELPAC, 2018).

Setting & Participants

The setting for this study was a Central California High School with over 3,000 students. Of the 3,000 students, 72% are Hispanic/Latino, 53% qualified for Free and Reduced Lunch, and

12% were ELLs (Education Data Partnership, 2018). In this study, a purposeful convenience sample was used to identify a treatment group and a control group, referred to as Block 2, and Block 1, respectively. The total sample consisted of thirty ELLs enrolled in the researcher's two English Language Development classes; therefore, the sample was a convenience sample. The sample was purposeful because it matched the target population (i.e., high school ELLs). Block 1 was chosen as the control group because the researcher taught this group before teaching the Block 2 group during the school day. Block 2 is the treatment group.

Treatment group. Fourteen ELLs were enrolled in the treatment group. Prior to the intervention, student ELPAC speaking proficiencies were as follows: three students were designated as ELPAC level one (i.e., moderately developed), and eleven students were designated level two (i.e., somewhat developed). Student speaking proficiencies were measured in 2018 by the ELPAC Performance Level Descriptors (CA Department of Education, 2018a) and these scores were used to place students in the aforementioned levels. There were five ninth graders, two tenth graders, four eleventh graders, and three twelfth graders. Five participants in the treatment group were female and nine were male. All participants in the treatment group spoke Spanish as their first language. All participants in this group made up the entirety of Block 2 at the researcher's high school.

Control group. Sixteen ELLs were enrolled in the control group. Prior to the intervention, student ELPAC speaking proficiencies were as follows: five students were designated as ELPAC level one (i.e., moderately developed), six students were designated level two (i.e., somewhat developed), four students were designated level three (i.e., moderately developed), and one student had not been previously assessed. Student speaking proficiencies were measured in 2018 by the ELPAC Performance Level Descriptors (CA Department of

Education, 2018a). There were ten ninth graders, one tenth grader, two eleventh graders, and three twelfth graders. Nine participants in the control group were female and seven were male. All participants in the control group spoke Spanish as their first language. All participants in this group made up the entirety of Block 1 at the researcher's high school.

Measures

The measure used in the study was the ELPAC Speaking Performance rubric (ELPAC, 2018) in which students were asked to verbally describe a visual depicting a scene (ELPAC, 2018). The students looked at the scene (i.e., CRVA) and answered six questions about the scene using a single word, a short phrase, or a longer response. Each student, one by one, completed the task with the researcher or the second administrator, during a single class period. This speaking task occurred as a pretest and as a posttest. Each individual assessment occurred during class time and took no longer than five minutes. Based on the responses to the questions, students earned a score of zero (response was not relevant), one (response was limited or partially relevant), or two (response was relevant) as measured by the ELPAC rubric's score criterion (ELPAC, 2018). This rubric was created by the California Department of Education and was last updated in April 2018.

Validity. The rubric (see Appendix A) used was created by a variety of experts in the field and was published by the California Department of Education. This rubric was chosen by the researcher to ensure validity. Student performance was measured by two different assessors (i.e., the researcher and second administrator) who have been trained on the use of the ELPAC Speaking Performance Task rubric to ensure the accuracy of the data (ELPAC, 2018). The assessment was also created by experts in the field and published by the state of California, adding to the validity of the assessment.

Reliability. The ELPAC rubric is a California standardized assessment rubric created by a variety of experts in the field so it has internal reliability (ELPAC, 2018). To ensure reliability of data, two assessors utilized inter-rater reliability to maintain accuracy during the assessment period. Prior to the pretest, the two assessors calibrated and normalized the score criterion by which the scores were given. The pretest was administered by the two assessors (i.e., the researcher and second administrator). The second administrator scored 20% of each group, along with the researcher. The two assessors of the assessment achieved at least 80% reliability, meaning, the scores matched at least 80% of the time between the two assessors to be deemed reliable. This process was replicated for the posttest.

Intervention

The intervention included the daily direct instruction on the use of CRVA as a means of increasing scores on the ELPAC Speaking Performance Task (ELPAC, 2018). The students in the treatment group received five weeks of practice using CRVA during daily speaking practice assessments. At the start of each class period, students in the treatment group (i.e., Block 2) looked at a scene (i.e., CRVA) displayed on the classroom televisions and were tasked with answering six questions about the scene (see Appendix B). The visuals used were chosen by the researcher because each photo was considered culturally relevant to the participating students. Prior to student responses, the researcher pointed out key details in the scene, provided verbal and written sentence frames, and modeled exemplar responses and non-examples. Students worked with partners to practice speaking out loud and describing the scene (i.e., CRVA) to one another in small groups. After two minutes, the researcher asked the students to share out to the whole group. After each group completed the task and shared out to the entire class, the researcher gave instant verbal feedback on strengths and areas of growth and practice as it

pertained to the Performance Task rubric, which was written on the classroom whiteboard. Instant verbal feedback included a score, as measured by the ELPAC rubric (ELPAC, 2018).

Procedures

The intervention procedure started with a pretest given to the control and treatment groups, separately, on a one-to-one student-to-teacher basis. The pretest and posttest were measured by the ELPAC Speaking Performance Task rubric (ELPAC, 2018). The pretest and the posttest included the CRVA. Following the pretest, the intervention started. Through the intervention period, the treatment group received daily exposure to a CRVA, and received daily direct instruction on how to use the CRVA to perform well on the ELPAC Speaking Performance Task (ELPAC, 2018). The control group received regular instruction, during which students answered weekly questions about a CRVA, but did not receive direct instruction on how to use them. The intervention period lasted five weeks, and then the posttest was administered to both groups. Both groups were formatively assessed through the duration of the intervention by the two independent assessors to monitor progress; however, only the pretest and posttest were used to determine if student scores increased because of the intervention.

Data collection. Data was collected before and after the intervention. The ELPAC Speaking Performance Task was administered to students using a 1:1 ratio for both the pretest and the posttest. The first assessor collected data for 80% of students, and the second assessor collected data for 20% of the students. Student scores were generated using the ELPAC Speaking Performance Task rubric, and were scored on a 1:1 basis (ELPAC, 2018).

Fidelity. To ensure fidelity, the second assessor made classroom observations throughout the intervention to ensure that the treatment was being administered as described (i.e., daily exposure and direct instruction). The second assessor used a fidelity checklist (see Appendix C).

The assessor ensured that the control group was receiving instruction as normal (i.e., weekly exposure to CRVA without direct instruction). The observations happened one time per week, for five weeks.

Ethical Considerations

Respect for persons, beneficence, and justice were considered through the duration of the study. The treatment did not involve removing anything, only adding enrichment to the normal instruction. No information or tools were withheld from the control group; however, they did not receive the treatment. All students participating in this study benefitted because following the intervention, the ELPAC was fully administered at the high school, and each of the students participating in the study had to take the assessment. All student information (i.e., names, demographics, and scores) remained anonymous and confidential.

Validity threats. The ELPAC rubric was the number one way to ensure that biases were not a threat to the validity and outcome of the study. The rubric was used by two independent assessors to decrease bias and maintain fidelity in the study. Inter-rater reliability was used to ensure that all students are tested the same to maintain validity in the study. It is important to note that the subjects in this study continued to attend their other classes, so, it is possible that the other courses attended might have had an impact on the subjects' performance on the assessments. However, this is the only situation in which they received daily exposure and direct instruction on the CRVA in connection with the ELPAC assessment (ELPAC, 2018).

Quantitative Data Analysis

All data was entered into the Statistical Package for the Social Sciences ® (SPSS®) for Windows, version 24.0.0 (IBM SPSS, 2016). No names of identifying information were included in the data analysis. Before analyses was conducted all data was cleaned to ensure no outliers

were present (Dimitrov, 2012). After cleaning the data, Independent samples t-tests (control and treatment groups) and dependent samples t-tests (pretest and posttest) were conducted to determine the significant difference in the quality of speaking performance between the two means scores on the ELPAC Speaking Performance rubric (ELPAC, 2018). Further, before interpreting the analytical output, Levene's Homogeneity of Variance was examined to see if the assumption of equivalence was violated (Levene, 1960). If Levene's Homogeneity of Variance was not violated (i.e., the variances were equal across groups), data was interpreted for the assumption of equivalence; however, if the variances were not equal across groups the corrected output was used for interpretation.

Results

Two independent samples t-tests were conducted on the whole sample ($n = 30$) for both the pre and post assessment scores. Results for the pre-test were: Levene's Homogeneity of Variance was not violated ($p > .05$), meaning the variance between groups was not statistically different and no correction was needed, and the t-test showed non-significant differences between the mean scores on the pre-tests between the two groups $t\ (28) = 1.115$, $p>.05$. This means that the groups were similar because there was no significant difference between the means on the pre-test for either group (see Table 1). Results for the posttest were: Levene's Homogeneity of Variance was not violated ($p >.05$), meaning the variance between groups was not statistically different and no correction was needed and the t-test showed non-significant differences between the mean scores on the posttests between the two groups $t\ (28) = -1.282$, $p>.05$. This means that the groups were comparable because the means of both groups were similar (see Table 1).

Table 1

Results of Independent Samples T-Tests

	Mean	SD
Pre Test		
Treatment	6.07	2.369
Control	7.00	2.191
Post Test		
Treatment	10.71	1.490
Control	10.00	1.549

Note. SD = Standard Deviation.

After determining the differences between pre and post assessment scores between groups, two paired t-tests were run for both groups (i.e., treatment and control) to determine if participants' mean scores from pre to post were significantly different within each group (See Table 2). Results for each group were as follows: treatment group, t (13) = -8.006, p< .001; control group, t (15) = -5.477, p< .001. This indicates that the treatment and control groups showed a statistically significant difference in mean scores from the pre-test to the post-test. Additionally, the negative t-values for each group indicate an increase in scores from pre to post assessment (See Table 2). The control group increased by 3.00 and the treatment group increased by 4.64; indicating the treatment group had a greater average increase than the control group. However, these gains were not statistically significant as shown by the results of the paired t-tests.

Table 2

Results of Paired T-Tests

	Mean	SD
Treatment Group*		
Pre	6.07	2.369
Post	10.71	1.490
Control Group*		
Pre	7.00	2.191
Post	10.00	1.549

Note. SD = Standard Deviation. * = $p < .001$.

Discussion

According to the California Department of Education (2018c), there are more than one million ELLs enrolled in California schools. Many of these students are struggling to meet grade-level standards and many of them are emerging in language proficiencies in reading, writing, speaking, and listening (CA Department of Education, 2018a). The aim of this research was to raise ELLs' scores on speaking. Researchers have suggested that culturally relevant testing materials may increase assessment scores for ELLs in all four domains (Basterra et al., 2011; Prosser & Solano-Flores, 2010; Solano-Flores, 2008). Other researchers have suggested that visual aids can assist the language production of ELLs (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Sinclair, 1987; Wright, 1989). Lastly, others have suggested that CRP can assist in language acquisition and production for ELLs (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984). While these bodies of research have been effective independent of one another, they have not been combined to examine their collective effectiveness on the assessment of ELLs.

This study aimed to combine these bodies of research in order to help ELLs achieve greater results on speaking assessments. This study included thirty ELLs enrolled in the researcher's classes. Sixteen high school ELLs (i.e., control group) received regular instruction, during which these students answered weekly questions about a CRVA. Another fourteen high school ELLs (i.e., treatment group) received daily exposure to a CRVA, with the addition of daily direct instruction on how to use the CRVA to perform well on the ELPAC Speaking Performance Task (ELPAC, 2018). The hypothesis for this study was that the regular exposure and direct instruction on the use of CRVA would increase ELLs' scores on speaking assessments. The measure used was the ELPAC Speaking Performance Task and corresponding rubric.

The results of the intervention in Table 2 indicate that the assessment scores of the treatment group increased from pre to post assessment. The treatment group increased their mean score from 6.07 to 10.71 (out of 12 possible points). These results are consistent with the literature regarding culturally relevant testing materials (Basterra et al., 2011; Prosser & Solano-Flores, 2010; Solano-Flores, 2008), visual aids (Echevarria et al., 2013; Nobuyoshi & Ellis, 1993; Sinclair, 1987; Wright, 1989), and CRP (Cochran-Smith et al., 2003; Gay, 2000; Samson & Collins, 2012; Trueba et al., 1984). These gains, however positive, do not conclusively show that the intervention was the sole reason, or even one of the reasons, for the increase in the assessment scores. Interestingly, the control group also had an increase in their mean scores. The control group increased their mean score from 7.00 to 10.00 (out of 12 possible points). While only the treatment group received the intervention and showed more growth, the results in Table 2 show that both groups had comparable mean scores on the posttest suggesting that both groups improved similarly. While the treatment group showed more growth than the control group, there

was no statistically significant difference between the two groups. The control group received quality teaching and benefited from other strategies which may have led to an increase in scores. The treatment group received the intervention which may have led to an increase in scores. While both methods of teaching were effective, these results also suggest that perhaps there were other factors that may have contributed to the higher assessment scores other than the planned and executed intervention.

Limitations and Future Studies

There were many limitations to this study. The greatest limitations of the study were regarding sample size and type of sample. Overall, the total sample consisted of thirty ELLs enrolled in the researcher's two English Language Development classes. The sample size was not nearly large enough to generalize the results to a larger population of ELLs. Additionally, the sample type was also problematic as the researcher utilized a convenience sample consisting of students enrolled in the researcher's classes. Future iterations of this study and research should utilize a much larger sample size, and a sample that can better generalized to the population as a whole.

Furthermore, some of the students in the sample were enrolled in more than one support class with the researcher. While the students in the control group did not receive the same intervention as the treatment group, some of the students in both groups received additional daily interventions as a school directive, before, during, and after the intervention period. These additional interventions included the daily use of language acquisition programs Duolingo and Rosetta Stone. These language acquisition programs likely contributed to the overall increase in speaking assessment scores by virtue of design. Both of these programs require students to practice speaking in order to progress through the daily learning modules. Future research should

utilize a sample that is participating in only one intervention at a time in order to determine the effect that the intervention may have. Specifically, the treatment group should not be enrolled in multiple interventions. Perhaps a second treatment group can be established to determine whether Duolingo and Rosetta Stone are more effective than CRVA for raising speaking assessment scores.

Lastly, all of the students in the sample were enrolled in five or six other classes throughout the entire intervention period. These students were enrolled in general education classes including Biology, English, History, Physical Education, and Art, all of which are taught in the target language. Research supports that input can lead to language acquisition (Krashen, 1998; Swain, 1993); therefore, it stands to reason that if students are spending six or seven hours a day listening to teachers speak English and reading texts in English, they are more than likely to acquire the target language. This study was five weeks so the intervention results may simply reflect time spent immersed in the target language. Furthermore, the teachers of the other general education classes may have been using their own interventions to help this population of students increase their English proficiencies altogether. Future research should examine whether the daily practice in speaking tasks, immersion in general education classes, or the CRVA are more effective in the acquisition and production of language by high school ELLs.

References

Basterra, M. D., Trumbull, E., & Flores, G. S. (2011). *Cultural validity in assessment: Addressing linguistic and cultural diversity*. New York: Routledge.

Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R., J. Jarvelle, and W. J. M. Levelt (eds.) *The child's concept of language*. New York, NY: Springer-Verlag.

California Department of Education. (2018a). *California English Language Development Test* [Data file]. https://dq.cde.ca.gov/dataquest/CELDT/results.aspx?year=2016-2017&level=state&assessment=1&subgroup=1&entity=

California Department of Education. (2018b). *English Learners FAQ - CalEdFacts.* https://www.cde.ca.gov/sp/el/er/elfaqs.asp#idplace

California Department of Education. (2018c). *Facts about English Learners in California - CalEdFacts.* https://www.cde.ca.gov/ds/sd/cb/cefelfacts.asp

Cochran-Smith, M., Davis, D., & Fries, K. (2003). Multicultural teacher education: Research, practice, and policy. In J. Banks & C. Banks (Eds.), *Handbook of research on multicultural education.* New York, NY: MacMillan.

Dimitrov, D. M. (2012). *Statistical methods for validation of assessment scale data in counseling and related fields*. Alexandria, VA: American Counseling Association.

Echevarria, J., Vogt, M., & Short, D. J. (2013). *Making content comprehensible for English learners: The SIOP® model: International edition.* Boston, MA: Pearson.

Education Data Partnership. (2018). Retrieved from https://www.ed-data.org

ELPAC Practice Tests. (2018). https://www.elpac.org/resources/practicetests/

Gay, G. (2000). *Culturally responsive teaching: Theory, research and practice.* New York, NY: Teachers College Press.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Oxford, England: Pergamon.

Krashen, S.D. (1998). Comprehensible output. *System, 26*(2), 175-182. doi.org/10.1016/s0346-251x(98)00002-5

Lavalle, P. I. & Briesmaster, M. (2017). The study of the use of picture descriptions in enhancing communication skills among the 8th grade subjects – learners of English as a foreign language. *i.e.: inquiry in education*, *9 (*1*),* 6-10.

Levene, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics, 1*, 278-292.

Nobuyoshi, J. & Ellis R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*(3), 203-210. doi:10.1093/elt/47.3.203

Philominraj, A., Jeyabalan, D., & Vidal-Silva, C. (2017). Visual learning: A learner centered approach to enhance English language teaching. *English Language Teaching, 10*(3), 54-61. doi:10.5539/elt.v10n3p54

Prosser, R. R. & Solano-Flores, G. (2010). *Including English language learners in the process of test development: A study on instrument linguistic adaptation for cognitive validity.* Paper presented at the Annual Conference of National Council of Measurement in Education, Denver, Colorado, April 29-May 3.

Samson, J. F., & Collins, B. A. (2012). *Preparing all teachers to meet the needs of English*

language learners: Applying research to policy and practice for teacher effectiveness. Washington, DC: Center for American Progress.

Sinclair, J. (1987). *Collins COBUILD, Collins Birmingham University International Language Database: English language dictionary.* London, England: Collins.

Solano-Flores, G. (2008). Who is given tests in what language by whom, when, and where? The need for probabilistic views of language in the testing of English language learners. *Educational Researcher*, 37(4), 189–199. https://doi.org/10.3102/0013189X08319569

Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching, 38(5),* 553-573.

Swain, M. (1993, October). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*(1), 158-164. https://doi.org/10.3138/cmlr.50.1.158

Trueba, H. T., Moll, L. C., Diaz, S., & Diaz, R. (1984). *Final report: Improving the functional writing of bilingual secondary students.* Washington, DC: National Institute of Education. (ERIC Document Reproduction Service No. ED240862).

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wolf, M. K., Herman, J. L., & Dietel, R. (2010). Improving the validity of English Language Learner assessment systems (CRESST Policy Brief No. 10—Full Report). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing. https://doi.org/10.1037/e685352011- 001

Wright, A. (1989). *Pictures for language learning.* New York, NY: Cambridge University Press.

Appendix A

ELPAC Speaking Performance Rubric

SPEAKING Talk About a Scene

In this task type, the student looks at a picture of a familiar scene in a school context. The student then answers six questions about the scene with a single word, a short phrase, or a longer response.

Aligned 2012 ELD Standards: PI.A.1, PII.B.3, PII.B.4, PII.B.5[5]

Score 0	Score 1	Score 2
• Response is not relevant. • Response contains no English. • No response, "I don't know," or is completely unintelligible.	• Response is limited or partially relevant. • Errors in grammar, pronunciation, or intonation impede meaning.	• Response is relevant. • Errors in grammar, pronunciation, or intonation do not impede meaning.

Appendix B

Talk About a Scene Example Task

Instructions: Look at the four visual aids below. Choose one and answer the six questions using a single word, short phrase, or a longer response.

CRVA 1	CRVA 2
	
CRVA 3	**CRVA 4**
	

SAY: Look at pictures above and choose one. I am going to ask you some questions about it.

Questions:

1. What is the character doing in this scene?
2. How would you describe the location of this scene?
3. Using a complete sentence, describe what is happening in the scene.
4. Using complete sentences, describe what one of the characters is wearing.
5. Books, movies, and television shows have titles. What would you title this scene?
6. What is one thing you want to know about this scene?

Appendix C

Fidelity Checklist

Observation 1

Treatment Group		
Week of Observation	Date Observed	Signature

Observed Behaviors	Initials
Teacher uses Culturally Responsive Visual Aids (CRVA) for direct instruction.	
Teacher provides sentence frames to students for speaking tasks.	
Teacher gives verbal feedback to using ELPAC Speaking Performance rubric.	
Students are working in groups to practice speaking tasks using CRVA.	
Notes:	

Observation 2

Control Group		
Week of Observation	Date Observed	Signature

Observed Behaviors	Initials
Teacher is engaging the students in speaking tasks.	
Students are working in groups to practice speaking tasks.	
Teacher gives verbal feedback to students about their speaking tasks.	
Notes:	

Observation 3

Treatment Group		
Week of Observation	Date Observed	Signature

Observed Behaviors	Initials
Teacher uses Culturally Responsive Visual Aids (CRVA) for direct instruction.	
Teacher provides sentence frames to students for speaking tasks.	
Teacher gives verbal feedback to using ELPAC Speaking Performance rubric.	
Students are working in groups to practice speaking tasks using CRVA.	
Notes:	

Observation 4

Control Group		
Week of Observation	Date Observed	Signature

Observed Behaviors	Initials
Teacher is engaging the students in speaking tasks.	
Students are working in groups to practice speaking tasks.	
Teacher gives verbal feedback to students about their speaking tasks.	
Notes:	

Observation 5

Treatment Group		
Week of Observation	Date Observed	Signature

Observed Behaviors	Initials
Teacher uses Culturally Responsive Visual Aids (CRVA) for direct instruction.	
Teacher provides sentence frames to students for speaking tasks.	
Teacher gives verbal feedback to using ELPAC Speaking Performance rubric.	
Students are working in groups to practice speaking tasks using CRVA.	
Notes:	

Observation 6

Control Group		
Week of Observation	Date Observed	Signature

Observed Behaviors	Initials
Teacher is engaging the students in speaking tasks.	
Students are working in groups to practice speaking tasks.	
Teacher gives verbal feedback to students about their speaking tasks.	
Notes:	