ENGLISH LANGUAGE AND LITERACY ACQUISITION VALIDATION: ELLA-V

TEXAS A&M UNIVERSITY
IN PARTNERSHIP WITH
SAM HOUSTON STATE UNIVERSITY, JOHNS HOPKINS UNIVERSITY, 25 TEXAS URBAN, SUBURBAN, SMALL TOWN, AND RURAL SCHOOL DISTRICTS
AND
TEXAS EDUCATION SERVICE CENTERS IV AND VI

A Validation Grant Proposal under the
Investing In Innovation Grant Program (CFDA 84.411B)

Table of Contents
Introduction- Absolute Priority and Competitive Preference Priorities 1
A. Quality of Project Design 1
A.1. Goals, Objectives, Strategies, Action Activities 3
A.2. Planning for Ongoing Work 7
A.3. Extent to Which Cost are Reasonable 8
A.4. Cost for Scaling 8
B. Significance 9
B.1. Exceptional Approach 9
B.2. Up-To-Date Knowledge 9
B.3. Importance and Magnitude of the Effect to be Obtained 17
C. Quality of Management Plan/Personnel 21
C.1. Adequacy of the Management Plan to Achieve the Objectives 22
C.2. Qualifications of the Project Director & Key Project Personnel 23
C.3. Capacity to Bring the Proposed Project to Scale 25
D. Quality of the Project Evaluation 25
D.1. Methods of Evaluation 25
D. 2. High-Quality Implementation: Progress Toward Outcomes 28
D.3. Key Elements of Project to Replicate in Other Settings 32
D.4. Sufficient Resources to Carry Out the Evaluation Effectively 32
References (Found in Appendix J.10)

Note: The Narrative that follows contains 35 pages total.
ENGLISH LANGUAGE AND LITERACY ACQUISITION VALIDATION: ELLA-V

Texas A&M University, in collaboration with Sam Houston State University and Johns Hopkins University (Evaluation Group), and in consortium with 25 LEAs, two Education Region Service Centers (Texas SEAs), submits English Language and Literacy Acquisition-Validation (ELLA-V) which meets one absolute priority and two competitive priorities. ELLA-V is based on moderate evidence from a prior IES funded RCT, Project ELLA. Absolute Priority 1, Improving the Effectiveness of Teachers, is met by increasing the number of teachers who are effective as measured by data that include, in significant part, student achievement or student growth from state-wide reading English language proficiency assessments and other student measures that are rigorous and comparable across schools, as well as a measure for teacher pedagogical effectiveness in English as second language (ESL) classrooms. Competitive Priority (CP) 6 is met by improving the quality of early learning programs (K-3) for high-need students by validating innovative practices and strategies via previously developed interventions that address the limited English proficient (LEP)/English language learners (ELL) students’ cognitive readiness for core academic subjects in English with a standards-aligned curriculum. Competitive Priority 8 is met by including LEP/ELLS as the sample from participating districts with the validation of practices and strategies designed to improve ELLs’ academic outcomes.

A. QUALITY OF THE PROJECT DESIGN

The ELLA-V project is designed to validate, via a randomized control trial (RCT), the intervention components of the original Project ELLA (hence called ELLA, while ELLA-V denotes the current validation project), across urban, suburban, small town, and rural Texas sites.

1 In all of our research work, we have referred to limited English proficient (LEP) students as English language learners (ELL) or English learners (EL); therefore, for the remainder of the proposal, we will respectfully request the reviewers’ indulgence as we refer to LEP with the equivalent of ELL.
ELLA, a prior IES-funded, longitudinal RCT (R305P030032) yielded moderate evidence from a 4-year, K-3 intervention in English language and literacy for ELLs. Findings in ELLA were based on combined intervention components; however, which of those combined intervention components by grade level was most sensitive to native Spanish-speaking ELL student growth on English language acquisition at each grade, K-3, was not determined.

Thus, ELLA-V is to validate the ELLA interventions at each of four grades, K-3, to determine the degree of impact that each intervention independently has on ELLs’ English oral language, reading, and/or science (Gr. 3). ELLA-V will be conducted: (a) as an RCT design, K-3, (b) in randomly assigned schools with students served in English as second language (ESL) classrooms (schools = unit of analysis), (c) in schools serving majority native Spanish-speaking ELLs in urban, suburban, small town, and rural districts, and (d) with treatment (T) ESL classes compared to typical practice control/comparison (C) ESL classes. Secondly, ELLA-V will examine the impact of high quality professional development (PD) provided for T teachers through virtual PD using Blackboard, Tegrity, Citrix GotoMeeting, and ThereNow classroom cameras versus typical PD provided by the districts to the C group teachers on (1) pedagogical skills measured by the Transitional Bilingual Observation Protocol (TBOP) from the Four Dimensional Bilingual Pedagogical Theory (Lara-Alecio & Parker, 1994; Bruce, Lara-Alecio, Parker, Hasbrouck, Weaver, & Irby, 1997), (2) overall observed T and C lesson effectiveness measured by the Teacher Observation Record (TOR- used for fidelity of T), and (3) ELL academics measured by state/ national English lang./reading tests K-3 and a science test, Gr. 3.

Texas LEA letters (Appendix G) demonstrate the extensive connections, contacts, and pre-planning that has been undertaken to gain buy-in from partner LEAs to conduct ELLA-V. In addition to the 25 districts specifically named as partners from urban, suburban, small town, and
rural Texas with up 587 total schools, there is the involvement of two Education Region Service Centers (ESC) IV and VI that have access to their consortium of 105 districts. Additionally, we currently have private partner matches committed for a total of $656,460 plus another verbal $325,000 and we will secure other partners needed for the private sector match (Appendix C).

A1. GOALS, STRATEGIES, ACTIONS ALIGNED WITH PRIORITIES

Two goals address ELLA-V and can be placed in the required Excel Management Sheet.

GOAL 1. Aligned with Competitive Priorities 6 and 8, Goal 1 is to determine (a) the extent to which each of the ELLA intervention components when implemented independently over 28-wks impact K-3 ELL s’ growth in English language, reading, and science and (b) differences between T and C group ELLs in the same areas, beginning with Gr. 3 and working backward each year through K so as not to contaminate any grade level by teacher exposure or student exposure to any of the intervention components. Objectives, strategies, and actions follow.

Objective 1.0. Determine the extent to which 28 weeks of the ELLA intervention components of Santillana Intensive English (SIE), and Story reTelling and higher order thinking for English Language and Literacy Acquisition (STELLA) combined with Academic Oral Language (AOL) implemented with two T groups (T-1: SIE, T-2: STELLA/AOL) impact K- ELLs’ English oral language and literacy skills.

Strategy 1.1. Implement ELLA-V in the Fall 2016-Spring 2017 in 75 randomly assigned schools (T and C) in 150 K classrooms for ELLs during ESL time. Activity 1.1.1. Assess ELLs in T and C schools pre-intervention using the native language Spanish Test de Vocabulario en Imagenes Peabody (TVIP) to ascertain equivalence of groups. Activity 1.1.2. Pre/Posttest ELLs in T and C Kinder classes on specified measures.

---

2 Excel Management Sheet is required to be completed with the Officer upon funding.
Objective 2.0. Determine the extent to which the implementation of 28 wks of ELLA interventions of SEI (first 14 wks) and Early Interventions in Reading (EIR), Level I (second 14 wks), and STELLA/AOLS in two treatment (T) groups (T-1: SIE/EIR, T-2: STELLA/AOLS) impact Gr. 1 ELLs’ English oral language and reading.

Strategy 2.1. Implement ELLA-V in the Fall 2015 and Spring 2016 in 75 randomly assigned schools (T and C) in 150 Gr.1 classrooms for ELLs during ESL time. Activity 2.1.1. Pre/ Posttest ELLs in T and C Gr. 1 classes on specified measures


Strategy 3.1. Implement ELLA-V in the Fall 2014 and Spring 2015 in 75 randomly assigned schools (T and C) in 150 Gr. 2 classrooms for ELLs during ESL time. Activity 3.1.1. Pre/ Posttest ELLs in T and C Gr. 2 classes on specified measures

Objective 4.0. To determine the extent to which 28 weeks of components of Content Reading Integrating Science for English Language and Literacy Acquisition (CRISELLA) and STELLA/AOWLS in two T groups (T-1: CRISELLA, T-2: STELLA/AOWLS impact Grade 3 ELLs’ English oral language, reading, writing, and science achievement.

Strategy 4.1. Implement ELLA-V in the first (Fall 2013) and second (Spring 2014) years in 75 randomly assigned schools (T and C) in 150 Gr. 3 classrooms for ELLs during ESL time.

Activity 4.1.1. Pre/ Posttest ELLs in T and C Gr. 3 classes on specified measures
GOAL 2. Aligned with the Absolute Priority 1, Goal 2 is to determine the impact of one of the original ELLA intervention components, PD of teachers, in ELLA-V as was in ELLA--bimonthly, but in an altered condition as virtual PD.

Objective 5.0. To examine the impact of high quality PD provided for T teachers compared to routine school district PD for C teachers on (a) pedagogical skills measured by the TBOP, (b) lesson effectiveness measured by TOR, (c) ELLs’ achievement measured by STAAR, TELPAS, SAT-10 Science, TOPA2, GotoLearn, DIBELS, and WMLS-R, and (d) PD via qualitative data.

Strategy 5.1. Observe 150 T and C teachers in the ELLA-V three times per year with TBOP

Activity 5.1.1. Using the ThereNow technology, Evaluation Group (EG) observe using TBOP (Appendix J.1) three times (one initial, mid, and ending 28 wks) in T and C classrooms. Activity 5.1.2. Using the ThereNow technology, coordinators observe T and C using the TOR for fidelity three times (one initial, mid, and ending 28 wk period) -also used for fidelity of treatment

Strategy 5.2. Implement and test a virtual observation (TOR) and mentor feedback system by the coordinator/mentors (C/M) using ThereNow technology. Activity 5.2.1. Train on the ThereNow technology (the C/M, the T and C teachers, the PI's). Activity 5.2.2. C/M take field notes during TOR and mentoring time

Strategy 5.3. EG holds focus groups with T teachers. Activity 5.3.1. Gathered from focus groups, survey, and field notes of T teachers, data will be analyzed qualitatively for satisfaction, usage, effectiveness, preference, and similarities and differences of virtual observations and mentoring by EG

Strategy 5.4. Analyze the T (PD) and C teachers’ (no PD) students’ achievement measured by STAAR (Gr. 3), TELPAS (K-3), DIBELS (K-3), SAT-10 Science (Gr. 3), and WMLS-R (K-3).

Activity 5.4.1. Pre/Posttest ELLs in T and C classes at K-3 on specified measures noted. Activity 5.4.2. Evaluation Group trains all testers and monitors their testing
Strategy 5.5. Provide direct virtual PD and mentoring by Lead Coordinators and Coordinators/Mentors bimonthly (14 consecutive training sessions) to ELLA-V T teachers via Blackboard and Citrix GotoMeeting. Activity 5.5.1. Use original ELLA PD model and modify and adapt it virtually. Activity 5.5.2. PIs provide training to all coordinators on mentoring and coaching. Activity 5.5.3. Teachers are mentored in the use of reflection within portfolio development. Activity 5.5.4. Data from reflections in portfolios will be gathered in TK20 (portfolio/data management) and analyzed by EG

Objective 6.0. Determine sustainability of the PD

Strategy 6.1. Foster a collaborative support mechanism, a community of practice dyad (two treatment teachers per campus and a blog established for each treatment type for communication among all teachers in specific treatments). Activity 6.1.1. Train the treatment teachers in community of practice and in the use of the blog used via Blackboard, Tegrity or Citrix GotoMeeting (Mentors/Graduate Students monitor blog). Activity 6.1.2. Survey the teachers in an open-ended survey regarding their community of practice and analyze blog comments and postings (conducted by EG)

Strategy 6.2. Treatment teachers share ELLA-V skills with ESL teachers ending each grade level Activity 6.2.1. Develop webinars that can be viewed by other ESL teachers

Strategy 6.3. Train principals on intervention components and observation instruments

Activity 6.3.1. Provide virtual PD for principals in an initial meeting (with two follow-up PDs in the 5 years) about how to use the observation instrument in ESL classrooms, the intervention components, and the pedagogical premises of ELLA-V (after the full intervention time period is finished, the 25 C school principals will be trained on the same components as the T group principals in one webinar-25 principals). Activity 6.3.2. Gather T school principals’ (50)
perceptions via open-ended questionnaire related to perceived intervention impact on students and teachers via a survey – gathered by EG.

**Strategy 6.4.** Follow up with T teachers per grade level in the following year (and in the case of K, the final year of the project—will occur during the final 3 mos. of the project). **Activity 6.4.1.** T teachers complete surveys the year after the intervention implementation to determine the degree of continued implementation of the intervention

**Strategy 6.5.** Develop ELLA-Virsky as an outgrowth for dissemination and scaling.

**Activity 6.5.1.** Provide via the TAMU Language Diversity Network (LDN; ldn.tamu.edu; Appendix J.11) to Education Service Centers (ESC) IV and VI with the developed and edited webinars for virtual PD (the ELLA-V becomes at that point—ELLA-V (ELLA-Virsky (virtual university) housing multiple webinars for free for the 18,309 bilingual/ESL teachers across Texas). **Activity 6.5.2.** Edit PD webinars for public uploading to the LDN and ESCs. **Activity 6.5.3.** Assess the number of hits on the ELLA-V PD from the LDN and ESCs

**A2. PLANNING FOR ONGOING WORK OF THE APPLICANT AND PARTNER**

The PIs have experience in disseminating information from the original IES RCT as well as from other funded projects. There have been 11 refereed ELLA publications, nine dissertations, and six in progress. The ELLA PIs have experience in facilitating processes for teacher reflection, mentoring, and improvement (one member is author of the Reflection Cycle from *The Principal Portfolio* & *The Career Advancement Portfolio* and editor for the *Mentoring and Tutoring Journal*, Taylor & Francis) and teacher observation (one member is author of TBOP specifically for ESL). An Advisory Board (AB, 16 total) will be utilized with 4 university science and language faculty, 3 superintendents, 3 science coordinators, and 3 multilingual directors, 1 SEDL rep (SEDL is a national research, development, and dissemination laboratory), and 2 Educational Service Center Region (ESC) VI and IV personnel. The AB will meet annually with the EG and the PIs to
monitor the progress of the grant and aid in dissemination with one face-to-face meeting and one virtual meeting. PIs will develop and present plans to AB as to how ELLA-V will be a part of ongoing research/development work such as: dissemination of the results in national and state journals, in the media; and via PDs with the ESCs (statewide) and SEDL (nationally), and expand partners to promote ELLA-V in a broader context.

**A3. EXTENT TO WHICH THE COSTS ARE REASONABLE & A4. COSTS FOR SCALING**

ELLA-V directly will impact 14,400 Texas ELLs over the 5-year period (150 teachers each year x 24 students per classroom x 4 yrs in the project) with 600 teachers involved, and 75 principals on the randomly assigned urban, suburban, small town, and rural schools during ESL time\textsuperscript{3} with treatment of control classrooms (Total direct impact: 15,075). Approximately 40% percent of the total budget will directly support PD and mentoring of the teachers (Goal 2), with the remainder pointed toward the intervention, testing, and evaluation of ELLA-V Goals (Goal 1). The total grant budget is $14,827,438.00 over 5 years, and inclusive of the required 10% match totals $16,310,181.00. The cost requested of the sponsor for ELLA-V is $983.57 per participant ($14,827,438/15,075), and the overall cost inclusive of the private partnership is $1082 per participant ($16,310,181/15,075). To scale the cost is: 100,000=108,200,000; 250,000=$270,500,000; 500,000=$541,000,000. ELLA-V has further-reaching impacts related to costs via the PD webinars recorded from the grant per grade level. ESC IV, Houston area, has approximately 66,894 K-3 ELLs (Texas Education Agency-TEA, 2012) in K-3, and ESC VI, northern greater Houston suburban, small town, and rural area has approximately 5,764 ELLs

\textsuperscript{3} There are two general types of programs for ELLS that the ELLA intervention had success within the ESL timeframe: (a) Structured English immersion- SEI or (b) transitional bilingual education- TBE. Since the intervention produced moderate evidence (with a median effect size of .25) in either of the programs, the intervention can be used in either program type during ESL time.
(TEA, 2012) in K-3; thus, ELLA-V has potential to expand to a total of over 72,000 ELLs. The webinars will be provided to the ESCs for use with not only their ESL consortium teachers, but also teachers in the other 18 Texas ESCs. ELLA-V has the potential to reach over 516 ESL teachers in ESC VI and 3855 in ESC IV. In Texas, in K-3, there are 257,787 ELLs (17% of the K-3 population; TEA, 2012) and approximately 10,740 ELL teachers. When ESCs disseminate the PD to their consortium schools, accounting for ELLs teachers, indirect impact increases by 77,034 (+ 15,075 participants) with anticipated costs of $16,310,181/92,109 or $177 per student and teacher. The estimated cost of implementation of the ELLA-V interventions per grade level over K-3 is approximately $4,428 per classroom for materials/supplies/PD (24 students/1 teacher). There will be no overhead costs as is associated with the validation grant itself in conducting the research component/evaluation/instruments/implementation. Regular state assessments would suffice to check progress of students and the observation instrument, included in PD costs is sufficient to monitor the teacher’s pedagogy.

**B. SIGNIFICANCE - B.1. EXCEPTIONAL APPROACH & B.2. UP-TO-DATE KNOWLEDGE**

Recent demographics indicate that ELLs comprise 21% of the national enrollment in public elementary and secondary schools, with 79% of those students being Spanish speakers (National Center for Education Statistics [NCES], 2010). In Texas, over 830,796 students (an increase of almost 60,000 students served in 2009-2010) were served in ELL programs in 2010-2011, accounting for 16.9% of the school population (TEA, 2012). These numbers are significant when achievement is compared to that of non-ELLs. For example, at Grade 8, national data show that only 3% of ELLs achieved at or above proficient level on the 2009 National Assessment of Educational Progress (NAEP) reading assessment, in comparison to a 35% of native English-speaking students, or a statistically significant 43 points difference (NCES, 2010). Similarly, the
percentage of ELLs at or above proficient at Grade 8 was 5% in math and 2% in science, as compared to 36% in math and 32% in science among English speakers (NCES, 2010). Such an achievement gap was also reflected in Texas’ standardized science assessment in which Grade 5 ELLs scored lower than any other subgroup including Special Education, Economically Disadvantaged, and At-Risk (TEA, 2010). Science is embedded in ELLA-V for ESL instruction.

Related to ELLA-V’s approach to improving oral language development inclusive of vocabulary usage, Snow, Burns, and Griffin (1998) reported that children from non-English-speaking families at a low socio-economic status (SES) are more likely to enter school with a lower level of English proficiency than their monolingual middle-class counterparts. This was supported by Kieffer (2008) who explored the differences in English reading development trajectories among native English speakers and ELLs with limited/fluent initial English proficiency at school entry in Kinder and found that ELLs with low initial levels consistently performed below their fluent ES peers in English reading achievement, after adjusting for demographics, SES, and school poverty status. Furthermore, Kieffer underlined the pressing instructional need to accommodate ELLs in order to reduce the risk of reading difficulties. We contend that there is lack of a strong research base on which educational programs can be built that are conducive to ELLs’ academic success; it is difficult to form a base without a thorough understanding about interventions validated in rigorous experimental or quasi-experimental designs which address the nature of English language development (Tong et al., 2008). ELLA-V aids in addressing this need.

Regrettably, most research on effective approaches of English acquisition for ELLs is based on what has been found to be effective in native language (L1) literacy acquisition for monolingual speakers (August et al., 2005; Lopez, 2004). Furthermore, there has not been
sufficient research aimed at understanding how to improve the quality of literacy instruction for ELLs (August & Hakuta, 1997; August & Shanahan, 2006). Among the limited number of experimental studies that have been conducted, Vaughn et al. (2006) designed a 7-month English intervention Grade 1 struggling ELLs in reading in both Spanish and English. ELLs in the direct instruction intervention demonstrated significantly higher ability in phonemic awareness, basic reading, and reading comprehension than did non-intervention ELLs, with no observed difference in oral language ability. The same intervention was replicated by Vaughn et al. (2006) for another sample of ELLs with similar characteristics as those in the earlier study, and positive effects of the intervention were identified. However, since both ELL samples had reading problems in both English and Spanish in both studies, the results are to be considered with caution when applying them to non-struggling ELLs in one or both languages. ELLA-V anticipates, as was in ELLA, the majority of participating ELLs to have adequate L1 proficiency.

ELLA identified the effectiveness of direct, scaffolded, and explicit English instruction during a separate ESL block (e.g., Tong et al., 2008; Tong et al., 2010). Participants in the studies were identified as ELLs at the beginning of K and received the ELLA intervention over time. For example, Tong et al. (2008) documented oral English development over a 2-year period (K-1) and concluded that experimental students experienced a faster growth in expressive and receptive oral English skills. In addition, ELLs in experimental bilingual classes demonstrated similar rates of acquisition when compared to ELLS receiving the same ESL intervention in English-only classes. Similar findings were reported by Tong et al. (2010) in that differences were in favor of bilingual over English immersion students in English phonological awareness, oral language skills, letter word knowledge, and reading proficiency after 3 years of placement for both groups in an intervention (K-2). English literacy development can be
supported by using research-based programs that build on oral language, incorporate systematic phonics instruction, and vocabulary development (Slavin & Chueng, 2004; Thompson, 2006). ELLA-V validates standards-aligned structured curricular intervention at K-3 and provides the PD for developing ESL teacher effectiveness. 

_The ELLA Framework_ (Figure 1), representing the overall original intervention, will have its independent intervention components validated. As will be discussed later in the proposal, the RCT yielded moderate evidence (also see Appendix D). Noted in Figure 1, increased ELL achievement occurred in the English language and literacy intervention when (a) curriculum was standards-aligned and inclusive of research-based ESL strategies of academic oral and written language tested for validation in with oral language practice that was conducted daily with academic language content from science, (b) there was increased overall time for structured English with science as the area of content for integration in Grades 1-3 [Tier II], and (c) struggling ELLS [Tier III] were provided increased time with structured tutorials(The way we increased time during ESL time was by including science as the support content area). Time, an altered condition from the ELLA, will be applied within the 45-minute state-required ESL time and the interventions will be covered at the Tier II level only, (d) oral language was developed in K and Gr. 1 (1.5 years) in which students are purposefully speech-engaged and then students move to the next levels of needed instruction indicated in e and f as follows, (e) Direct Instruction in reading occurred in Gr. 1 and 2 (1.5 years), (f) content area reading in science in which structured and expository text was used in Gr. 3, and (g) consistent teacher PD with self-assessment, reflection, monitoring, and mentoring.
**INTERVENTION COMPONENTS** Just as was in ELLA, ELLA-V will include components within two levels: Level 1- Teachers’ PD and Level 2- Student Achievement. There were *three tiers* in Level 2 in the original ELLA that included (a) Tier I of general standards-aligned curriculum for all students, (b) Tier II of the ELLA standards-aligned curriculum for all ELLs, and (c) Tier III which was for the struggling reader ELLs. *Tiers I and III intervention components will not be validated* in ELLA-V; rather, ELLA-V will focus the validation with all ELLs at Tier II. **LEVEL 1: PD for ELLA-V.** Bimonthly PD will be provided to T teachers (6 hrs per mo) Teachers will be paid for PD time. PD is outlined and scripted, along with video clips and materials with the following activities: (a) review and practice upcoming lessons, (b) reflect on and discuss student learning, (c) assess pedagogical progress as a teacher in the intervention, (d) ESL instructional strategies (Appendix J.2), (e) vocabulary building and fluency, (f) oral-language development and the importance of planned student talk with less teacher talk, (g) literacy development, (h) reading comprehension, (i) language of instruction clarifications, (j) language of instruction, language content, communication modes, and activity structures, and (k) reflective practice via portfolio development. PIs will ensure the virtual
training is operational, will train the personnel in the use of the virtual observation, feedback, and mentoring, will review and edit the ELLA-V training manual, and will observe six PDs per year for fidelity. Just as ELLA, all teachers in ELLA-V will be recognized as highly qualified by Texas Public Law 107 110, Section 9101 (23)(B) (i). No other requirement will be made on the teachers as this would be a routine practice and a state mandate for all teachers of ELLs. **LEVEL II Student Instruction for ELLA-V.** Level II for ELLA-V is comprised of two tiers. **Tier I** is what all ELLs receive through their regular courses in the school’s specified program language of instruction. In all districts in Texas, curriculum is aligned to state standards for each content area and English language proficiency (ELPS) with alignment to the state exam, STAAR, beginning Gr.3. **Tier II** is of interest in the ELLA-V intervention components, which is the structured alignment and direct English intervention, delivered to treatment ELLs during the state-required 45-minute ESL block in K-3. As an altered condition, Tier II treatments will be implemented at individual grade levels and will not be dependent upon previous grade’s treatment effects in order test the independent impact of each component; therefore, the initiation of the ELLA-V begins at Gr. 3 with backward implementation to K. Instruction includes the following components by grade; components unique to each grade are shared with number of treatment groups noted.

**Year 4-5: Grade K. Treatment 1 Group in K.** Santillana Intensive English (SEI; Ventriglia & Gonzalez, 2000). SEI (Sample in Appendix J.3.), implemented for 28 wks for 45 minutes of ESL time, provides a series of scripted Texas ELPS standards-aligned lessons based on effective reading practices in phonemic awareness, phonics, vocabulary development, reading fluency, and reading comprehension and is a systematic approach to language instruction (engage, explore, teach, practice, apply, relate, and extend). **Treatment 2 Group in K:** Story-reTelling and...
higher-order thinking for English Literacy and Language Acquisition ([STELLA] Irby, Lara-Alecio, Mathes, Rodriguez, & Quiros, 2004; Quiros, Lara-Alecio, Tong, & Irby, 2011).

STELLA, to be implemented for 28 wks. during 20 min. daily in ESL time, has authentic children’s literature, narrative and expository with one book per wk., and used with scripts (example in Appendix J.4.) developed to increase oral language and to implement Bloom’s Taxonomy with leveled questions and vocabulary words aligned to ELPS, and pre-selected ESL strategies. Combined with STELLA is Academic Oral Language in Science AOLS (Appendix J. 5), aligned to state ELPS and science standards, implemented for 28 wks, 10 min daily in ESL time, develops students’ oral language.

Year 3-4: Grade 1. Treatment 1 Group Grade 1: SEI, implemented for 14 wks in Gr. 1 follows the same pattern and standards as in Gr. K. ELLs have SEI the first semester, and with Early Interventions in Reading I (EIR-1; Mathes, Torgesen, Menchetti, Wahl, & Grek, 2004), aligned to ELPS and reading standards, second semester 14 wks, 45 min daily in ESL time, was adapted by Mathes for ELLA and is taught in whole group instruction and aligned with five central strands: Phonemic Awareness, Letter-Sound Correspondences, Word Recognition and Spelling, Fluency, and Comprehension Strategies. Treatment 2 Group Grade 1: STELLA implemented for 28 wks, 20 min daily in ESL, follows the same format as in Gr. K. Combined with STELLA is Academic Oral Language in Science AOLS, implemented for 28 wks, 10 min daily, is used in the same manner as in K.

Year 2-3: Grade 2. Treatment 1 Group Grade 2: Early Intervention in Reading II (Mathes, et al, 2004). EIR-II, ELPS and reading state standards aligned, is implemented 28 wks, 45 min daily in ESL. Treatment 2 Group Grade 2. STELLA, implemented 28 wks, 35 min during ESL,
as it was in Gr. K and 1. Combined with STELLA is Academic Oral and Written Language in Science AOWLS, implemented 28 wks for 10 min daily (writing is added).

Year 1-2: Grade 3. Treatment 1 Group Grade 3. **Content Reading Integrating Science for English Language and Literacy Acquisition (CRISELLA)**, Irby, Lara-Alecio, Mathes, Rodriguez, & Quiros, 2007) (Appendix J.6). CRISELLA, implemented 28 wks, 45 min daily in ESL, is a structured, standards-aligned curriculum with state and new national science standards (altered condition as the prior CRISELLA was aligned with previous national science standards) with a set of instructional skills for science-embedded English language development. As in ELLA, we will use the 2006 Edition of Pearson-Scott Foresman’s Gr. 3 science text due to the scaffolded expository reading passages that assisted ELLs to develop science academic language and expository reading skills; it is not a selected science adoption for any of the districts and will not interfere with the regular science program (*Please note--The text is not the basis* of the instruction.). CRISELLA includes pre-reading skills, vocabulary building activities, partner reading, graphic organizers, hands-on inquiry activities, cooperative grouping, scaffolded and leveled questions, vocabulary extensions, fluency practice, and direct teaching of reading skills. The 5E Model (originated with Bybee, 1987), Engage, Explore, Explain, Elaborate, Evaluate, is employed for comprehension and retention. **Treatment 2 Group Grade 3.** STELLA - implemented 28 wks for 35 min at Gr. 3. AOWLS continues as a 10 min. component.

**COMPARISON CONDITION** For ELLA, the primary question was whether the overall intervention treatment was more effective for English development than what is generally available and utilized in schools for ELLs. The C group received what is called *business-as-usual*, the typical practice that a school implements. Business-as-usual generally refers to situations in which the standard or frequent practice across the district or region is a relatively
undefined education treatment. In C classrooms in ELLA-V, we expect it to be similar to what it was in ELLA – basically, a specific 45-minute separate ESL block with no support from the research team K-3 and with no consistent standards-aligned content embedded. As in ELLA, we expect typical practice ESL to be aligned to state ELPS, but also to exhibit major variation across classrooms in terms of instructional practices with ESL time block not tightly controlled. As in ELLA, we expect some teachers in the C class to speak Spanish with the students during ESL. ELLs in C classes will be taught by Texas certified ESL teachers recognized as highly qualified.

**B.3. IMPORTANCE AND MAGNITUDE OF THE EFFECT TO BE OBTAINED**

No studies were found that determine if the altered condition of virtual observations and feedback and online PD impact ESL teachers in a variety of types of schools. The ThereNow classroom observation and feedback technology should facilitate this validation study and could impact observations of future teachers and practicing teachers of ELLs, and the use of Blackboard, Tegrity videoing, and webinars via Citrix GotoMeeting should have an effect on implementation and on thousands of teachers after the intervention as webinars are released.

Three programs used in the original ELLA are in wide use. The Santillana Intensive English (SIE) commercial program is in wide use across the country as evidenced by state adoptions of the program for ELLs (adopted in Arizona, California, Georgia, Idaho, Louisiana, New Jersey, New Mexico, Oregon, and Utah; see https://santillanausa.com/ adoptions). The replication of SIE as an intervention component, because there have been no rigorous, RCTs inclusive of this product to judge its worth in terms of its independent intervention impact on ELLs’ English oral language development, in particular. Such validation testing has import to the states that have adopted it and to those that have not.
The Early Interventions in Reading (EIR), Levels I and II, commercial products of SRA McGraw-Hill, are widely used across the country and have had a few studies conducted on this product with ELLs (Vaughn, Linan-Thompson, Mathes, Duradola, & Cardenas-Hagan, 2007). However, there have been no studies that have analyzed it via an RCT, the way in which the EIR was adapted for use in whole-group instruction with ELLs as opposed to small group intervention with struggling readers. EIR is noted by the Institute of Education Sciences (2008) as an evidence-based intervention for students at risk of failing to learn to read. Three initial, basic studies have been conducted using EIR (in addition to Project ELLA); however, of the three initial studies conducted, only one (Vaughn et al.), included ELLs and only at Gr. 1. No studies have been conducted that record the impact of large-group instruction using EIR as an individual component with ELLs and with a half-year for Grade 1 in conjunction with an oral language development intervention for the first half of the academic year and fully implemented for Gr. 2. It is important to have information for practitioners who may wish to use such a program in large group instruction with ELLs to understand the implications of such usage and how it might be implemented if found to significantly impact ELLs’ language/literacy/reading.

For K, we used the Lakeshore cards from Daily Oral Language and pulled cards related to science for Academic Oral Language in Science (AOLS). Ten minute lessons were created. We then developed our own cards which we called AOLS/AOWLS (previously described and these lessons are not commercialized). No RCTs exist that include a consistent daily oral language component that is based in science. Other intervention components (STELLA, AOLS, AOWLS, CRISSELLA) have not been commercialized or validated independently. If found effective, there will be eight validated components that can be implemented with ELLs across K-3.
**Moderate Evidence** To date, in addition to the overall report submitted to the IES, Project ELLA has been evaluated accumulatively at various time points (K-3) in 10 closely linked studies based on data derived from this RCT. Three of these studies were included in a most recent best evidence synthesis by Cheung and Slavin (2012, accepted by *Review of Educational Research*) on effective reading programs for elementary ELLs. The criteria for inclusion were rigorous longitudinal designs with quantitative measures, and the authors stressed a need to revisit the review of research on language of instruction for ELLs based on the new evidence from our studies among others. The major studies demonstrating moderate evidence are listed in Appendix D, and we discuss five of them (see Appendix D for full studies). In the **first study** as part of the larger experimental project ELLA, 23 schools from a large urban school district in Southeast Texas were randomly assigned to treatment or control condition totaling 54 teachers in either bilingual or English immersion classrooms, and a comparison was conducted regarding these teachers’ pedagogical practices between conditions across the school year through the use of TBOP (Lara-Alecio, et al., 2009). Treatment teachers received bimonthly PD. Chi-square test of homogeneity of proportion found that experimental teachers’ pedagogical deliveries were positively impacted by the intervention, because they devoted a significantly higher proportion of instructional time than did control teachers (ps < .01, median effect size = .18) in (a) intensive English instruction, (b) cognitive and academic English language, (c) expressive language-related communication, (d) teacher-ask/student-answer type of activity, (e) research-based ESL strategies including academic scaffolding and leveled questions, and (f) academic tasks rather than social participation tasks, all of which are considered as effective instructional practice for ELLs (i.e., Gersten & Baker, 2000; Saunders et al., 1998; Slavin & Cheung, 2005). Therefore, an early intervention with ongoing teacher PD related to structured, standards-aligned English
instruction in cognitively demanding areas can provide ELLs with better academic language development opportunities and can increase teacher effectiveness. The second ELLA study was completed at the end of Gr.1 among 11 treatment and 12 control schools to examine oral language development trajectory over a 2-year period (K-1) using latent growth modeling (Tong, et al., 2008). It was found that the combined intervention components promoted ELLs’ academic oral language acquisition throughout K and Gr.1 at a significantly faster rate than was determined among the control students (p < .05, after Benjamini-Hochberg correction method of multiple group comparison was applied). In other words, the comprehensive intervention produced a difference in accelerating ELLs’ oral English development. The magnitude (i.e., Cohen’s d with pooled standard deviation) of such statistically significant difference was .68. In the third ELLA study (Tong, et al., 2010), used a multilevel modeling approach to address the clustering effect of repeated measures for students within classrooms and classroom within schools after 3 years of intervention in English immersion classrooms within 8 treatment schools and 10 control schools. Results revealed that students receiving treatment exhibited a steeper growth and outperformed their control peers by the end of Gr. 2, reflecting a range of phonological awareness, oral language skills, and decoding and reading proficiency (ps < .05). This advantage in the intervention group can be explained by greater emphasis on phonics, phoneme discrimination, segmentation, and blending through direct instruction and structured standards-aligned lessons. Consequently, the intervention might have filled a critical gap for well-planned, implemented, direct, and monitored curricular and instructional practices that can accelerate subsequent second language acquisition so that the disadvantages in English proficiency that ELLs bring with them upon school entry can be removed. The fourth study was a 3-year (K-2) longitudinal randomized study (Tong, et al., 2008) derived from the larger project.
The main objective was to evaluate the effectiveness of ELLA in the bilingual classrooms only. Nineteen schools were randomly assigned to either T (N=10) or C schools (N=9). The initial sample size was 502 and the final sample was 262, with an attrition rate of 48%. No differential attrition rate was found between the two groups. The findings indicated that T students outperformed their control counterparts in the areas of oral language, pre-literacy skills, and reading fluency and comprehension on English measures with effect sizes ranging from +0.13 to +0.70. The median effect size across all 14 English measures was +0.23. The T students also scored significantly higher than the C students on 7 out of the 14 Spanish outcome measures with a median effect size of +0.12. The *fifth ELLA study* (Irby, et al., 2010) provides an overall evaluation of the ELLs’ performance on a state reading assessment in Gr. 3 using a 2x2 design (treatment x language of instruction), in which schools were randomly assigned to either an treatment or control. It was reported that regardless of the type of ELL program (bilingual or English immersion), students in the experimental group were able to master the necessary knowledge and skills in academic reading more readily than were their control counterparts (median effect size = .14 in English measures and .18 in Spanish measures). More importantly, it is the quality of instruction, as reflected in ELLA that made a difference in ELLs’ English reading development.

In summary, the research team has 11 publications (Appendix J.7) currently published or accepted for publication related to Project ELLA (K-3) and has presented at numerous international, national, and state conferences, including a presentation to the USDOE Officers and to IES (March 26, 2010) on the effectiveness of the holistic intervention regarding teachers’ pedagogical behavior and ELL students’ English oral and literacy development.

**C. QUALITY OF MANAGEMENT PLAN/PERSOONNEL**
C.1. ADEQUACY OF THE MANAGEMENT PLAN TO ACHIEVE THE OBJECTIVES

There are four groups involved in ELLA-V. First, there is the *Advisory Board (AB)* that ensures application in the schools for furthering the project and disseminating findings. The AB was explained earlier in the text. Second there are the PIs which serve as the *Policy and Procedural Oversight Group (PPO)* PPO is led by the primary PI and is responsible for hiring at TAMU and SHSU and for scheduling all implementation ELLA-V activities. This group does all ordering of equipment and intervention materials and trains coordinators and EG in implementation components. The PIs will monitor coordinators and mentors on implementation. The PIs will ensure that the job is completed based on the timeline and that the materials are delivered. The PIs in the ELLA-V will not be involved in data gathering and analysis. PIs will train the mentor coordinators in the mentoring of teachers and use of fidelity instruments and the EG in the observation instrument. The PIs will assist staff with dissemination once data are analyzed and reported by the EG. Third is the *Application/ Implementation Group (AIG)*; this group is made up of Lead Coordinator, Coordinator/Mentors, Post-Doc, Graduate Students, Technical Support, Dissemination Specialist, and Consultants. The AIG will deploy all interventions with teachers in the project and will utilize all technical support and consultants to implement PD and mentoring. The fourth group is the *Evaluation Group (EG)* which exists of the external evaluator group and its internal evaluators and graduate assistants. The EG will be chaired by Johns Hopkins staff and made up of their expert evaluators. This group is charged with gathering data and implementing the evaluation design described in this application.

The Management Plan in Figure 2 indicates the major milestones with objectives and strategies noted, group responsible, and the timeframe for the objective/strategy (pages 3-7).

*Figure 2. Major Milestones, Group Responsible, and Timeframe*
<table>
<thead>
<tr>
<th>Major Milestones</th>
<th>Responsible Group</th>
<th>Yr 1</th>
<th>Yr 2</th>
<th>Yr 3</th>
<th>Yr 4</th>
<th>Yr 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective 1.0</td>
<td>AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 1.1</td>
<td>PPO;AI;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Objective 2.0</td>
<td>AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 2.1</td>
<td>PPO;AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Objective 3.0</td>
<td>AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 3.1</td>
<td>PPO;AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Objective 4.0</td>
<td>AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 5.1</td>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 5.2</td>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 5.3</td>
<td>EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 5.4</td>
<td>EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 5.5</td>
<td>PPO, AIG, EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Objective 6.0</td>
<td>AB;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 6.1</td>
<td>AIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 6.2</td>
<td>AIG &amp; Teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 6.3</td>
<td>AIG &amp; Principals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 6.4</td>
<td>PPO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy 6.5</td>
<td>AIG;EG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

**Other Critical Components**

- Recruit/hire all personnel: PPO ✓
- Establish Advisory Board: PPO ✓
- Establish all subcontracts: PPO ✓
- Order all materials: PPO ✓✓✓✓✓
- Establish all training with specific vendors: PPO ✓✓✓✓✓
- Communicate with district/school administrators: PPO;AIG;EG ✓✓✓✓✓
- Establish final agreements with districts/schools: PPO ✓
- Work with i3 Project Officer on Mgt. Excel File: PPO
- All data collection/analysis/reporting: EG ✓✓✓✓✓
- Grant reporting: PPO;EG ✓✓✓✓✓
- Disseminate results: PPO;EG;AIG;AB ✓✓✓✓✓

**C.2. QUALIFICATIONS OF THE PROJECT DIRECTOR & KEY PROJECT PERSONNEL**

**Dr. Rafael Lara-Alecio,** (30% time annually) PI, Professor of Educational Psychology and Director of the bilingual/ESL Programs, at Texas A&M University, has served as PI of multiple U.S. DOE funded projects, including Field Initiated Research, IES longitudinal research (Project...
ELLA), and National Science Foundation DRK12. He has managed multimillion dollar budgets. His primary expertise is bilingual/ESL education, assessment and evaluation, and curriculum development and implementation for ELLs. He is bilingual/biliterate in English and Spanish and has authored and/or co-authored over 100 publications including refereed papers, book chapters, books, technical reports, and textbooks for ELL children grades PK-3. He has numerous research, teaching, and service honors. Dr. Beverly J. Irby, (20% time annually) PI for SHSU management of the grant, is Associate Dean for Graduate Programs, Sam Houston State University and is the sole awardee of the Texas State University System Regents’ Professor in 2009. She is bilingual and has served as PI/ Co-PI of $20,000,000+ of research/leadership development grants, including IES and NSF, related to the academic improvement of ELLs or at-risk students with disabilities or of low-income. She is author/co-author-editor of 24 books and of 100+ articles, is co-developer of a 21st Century leadership theory, The Synergistic Leadership Theory, published in the Journal of Educational Administration, is developer of the science components of the DLM Early Childhood Program for SRA McGraw-Hill. She is co-founding editor of the Advancing Women in Leadership Journal, is the editor of Mentoring and Tutoring Journal (Routledge Taylor & Francis), and is founding editor of the Graduate Research Journal. Irby holds the SHSU University Research Award. Dr. Fuhui Tong, (20% time annually), Co-PI, Assistant Professor of Educational Psychology in the Bilingual/ESL Programs at Texas A&M University, has been engaged as key personnel on multiple U.S. DOE funded projects, including Field Initiated Research and an IES longitudinal research; she is also the Co-PI for the NSF research grant. She has conducted several program evaluations on federally-funded projects; her primary expertise is bilingual/ESL assessment and evaluation, second language acquisition, and longitudinal data analyses and structural equation modeling. Cindy Guerrero, ABD, (100% time
annually) served as Lead Coordinator on two longitudinal research grants (IES-ELLA and NSF-MSSELL). She has been able to coordinate all implementation and PD elements for grants for 10 years. Her experience as a coordinator and her ability to motivate teachers and work with administrators makes her unique to this project. She is well-organized and is outcomes driven.

To distance the PIs from the project, there will be several independent external evaluators from the Johns Hopkins EG (D.4).

**C.3. CAPACITY TO BRING THE PROPOSED PROJECT TO SCALE**

The strength of ELLA-V team to build this to scale is based on the collaboration and trust with school districts in terms of the research and the researchers. It brings to the forefront research, evaluation, and accountability for ESL teachers and ELLs. During the 5 years of ELLA-V, there will be a strong PD resource developed for implementation in districts and offered free online via Education Service Centers and the LDN via *ELLA-Virsity* for replication of the model components and strategies which will be available at the state, national, and international levels. In ELLA-V, the team will provide guidance and assistance to the partner LEAs and will provide follow-up each year for continued implementation and expansion.

**D. QUALITY OF PROJECT EVALUATION--D.1. METHODS OF EVALUATION**

**Sample and Research Design.** To evaluate the effectiveness of ELLA-V, a cluster-randomized design will be employed with 75 schools from among a total of 587 urban, suburban, small town, and rural schools in Texas randomly assigned to treatment (T1=25, T2=25) or control/comparison (C=25) conditions (i.e., business-as-usual/typical practice, C = 25). Two teachers per school will be randomly selected at the specific grade level to implement the corresponding condition using propensity score analysis. Sets of 3 schools will be matched before random assignment with the following school-level demographic data: ethnicity,
socioeconomic status, previous scores on English language proficiency tests (as given by the district to determine ELL status), and regional classification (e.g., rural, small town, urban, suburban). The use of matching before random assignment can improve the precision of the estimate, hence, increase the statistical power in group-randomized studies (Raudenbush, Martinez, & Spybrook, 2007) and can eliminate the possibility of pretest differences on matching factors despite random assignment. All the measures will be administered in the same way to all participants, K-3, T and C. No school campus will house both T and C classrooms combined. Each year new teachers (classrooms) will be randomly selected to participate in the study at the appropriate grade level (e.g., Gr. 3 in Year 1, Gr. 2 in Year 2, Gr. 1 in Year 3, and K in Year 4). The integrity of such assignment will be maintained because when a school is assigned to receive T1 in Gr. 3 in Year 1, then this school will continue to receive T1 in the subsequent years for each grade. So that there will be no overlap of related impact of the implementation year by grade level, the initial replication will begin at Gr. 3 in progressing backward to K in Yr. 4-5. In this manner, there is no impact of ELLs or teachers having had the intervention in a previous grade level, with potential impact on the subsequently tested grade level. As ELLs enter and leave the schools, especially if they transfer among schools, we will keep careful track of their initial assignments and their receipt of services over time. ELLs at all T and C schools will be pre- and post-tested at the beginning and the end of each year of implementation, respectively.

**Power Analysis.** In ELLA-V, because students will be nested in schools in each grade level, we used Optimal Design (Raudenbush, Liu, Spybrook, Martinez, & Congdon, 2006) to determine the number of schools needed for the experiment. We will use hierarchical linear modeling (HLM), the appropriate multi-level analysis for clustered designs. The ability to detect a treatment effect at a certain level of power in an HLM framework depends on several factors:
intra-class correlation (ICC, ρ), the correlation between pre and posttests (r), and the average number of students in each school (n). In our power analysis, the parameters included an alpha level of .05, pre- posttest correlation of .70, a target minimum detectable effect size of .25 (although the effect sizes in our experimental research have mostly been larger than .35, with a maximum of .7, we decided to use a more conservative effect size in the Validation study), a cluster size of 40 (including 2 classes per campus), and ICC of .10, which is quite commonly found in cross-sectional studies such as the Special Education Elementary Longitudinal Study (SEELS) (Also see Hedges & Hedberg, 2007). Using these parameters, we can detect an effect size of 0.25 with a power of 0.80 if we have a school sample of 25 per condition or 75 overall, using a 2-level HLM model. As required by the TEA, each classroom will have no more than 24 students. Therefore, although the 2 classes of 48 students on each campus might not all provide consent to participate (that is why we propose a cluster size of 40 students per campus), these students will all be receiving district and state-mandated assessments, including TELPAS, benchmark tests, and STAAR in 3rd grade, and should be impacted by ELLA-V. We anticipate serving a total number of 14,400 students over the life of this project.

**Attrition and Missing Data.** The overall attrition rate of over 40% was an issue with ELLA, because it was a pure longitudinal study with no cohorts added after the initial one was established. However, ELLA-V tests independent intervention components at *each* grade K-3 and will not be a longitudinal study across years. Based on ELLA, the average overall attrition rate was 15% per year over 4 years, with a differential attrition rate of 6% between T and C groups at the student level and less than 1% attrition rate at the school level. Therefore, we do not anticipate a high attrition rate or a high differential attrition rate. Intuitively, this conclusion is reasonable because students are not likely to attrit from the study due to their T or C status.
Attrition in the ELLA was more likely related to family mobility. The EG will follow the model proposed by What Works Clearinghouse (WWC, 2011) to determine the overall and differential attrition that may bias the estimated intervention effect. EG will use an intent to treat model, following all ELLs in all schools randomly assigned at the outset to the extent possible. Finally, to handle missing data that may arise in the data collection process, EG will use NORM version 2.03 (Schafer, 1999) to impute incomplete multivariate data so as to ensure sufficient statistical power and to minimize attrition bias if there is any. **Fidelity of the Implementation of the Intervention.** TOR (Appendix J.8), developed and tested in ELLA, is an instrument used to monitor the fidelity of the intervention. TOR has internal consistency of .94. It captures the core components of intervention three times per year. Data collected via TBOP and TOR will be incorporated as teacher-level characteristics into analyses of the impact of the intervention.

**D. 2. HIGH-QUALITY IMPLEMENTATION: PROGRESS TOWARD OUTCOMES**

**Measures.** We will compare students’ achievement on the following constructs (Appendix J.9 for a list of tests) aligned with the research questions: phonemic awareness, vocabulary, oral language, and reading comprehension, as measured by both (a) state-mandated standards-aligned assessments, including STAAR (Gr. 3) and TELPAS (K-3); and (b) other rigorous standardized language and reading instruments not addressed by the State including: Test of Phonological Awareness Plus (TOPA 2+, K-3), Woodcock-Muñoz Language Survey-Revised (WMLS-R, English, K-3), GotoLearn Benchmark Tests (Language Arts, Grade 1; Reading, Grade 2 and 3; Science Grades 2 and 3), Stanford Achievement Test-10 in Science (Gr. 3), and Dynamic Indicators of Basic English Literacy Skills (DIBELS, 1-2). Among these, benchmark tests with GotoLearn are formative assessment aligned with state standards, i.e., TEKS, and are administered every 6 wks. They are embedded in instruction to provide timely feedback for
purposes of adjusting instruction to improve learning. TELPAS is a state-wide teacher rating scale aligned with state standards and ELPS to measure ELLs’ language skills in speaking, listening, reading and writing. TOPA2+ (K-3), WMLSR (K-3) and DIBELS (1-2) are standardized measure given as pre- and post-test at each appropriate grade level to all participating students. Most of these measures were used in the previous evaluation of ELLA (STAAR tests were not used because during the study the state-mandated TAKS- STAAR replaced TAKS --2011-2012). At the teacher level, we will compare instructional effectiveness between T and C teachers using TBOP, and TOR (used for fidelity of implementation in ELLA). Other qualitative instruments, ELLA-V Goal 2, will be developed and measured for validity.

**Research Questions.** In validating intervention components from ELLA, the evaluation will address the following two sets of specific questions aligned with the project goals:

**Questions to Evaluate Goal 1.** 1.a. How effective is the ELLA-V model in promoting ELL Kindergarten students’ overall academic growth in English language acquisition?

1.b. How effective is the ELLA-V model in promoting ELL first grade students’ overall academic growth in English language acquisition?

1.c. How effective is the ELLA-V model in promoting ELL second grade students’ overall academic growth in English language literacy (reading/ writing) and science achievement?

1.d. How effective is the ELLA-V model in promoting ELL third grade students’ overall academic growth in English language literacy (reading/ writing) science achievement?

1.e. Are there student or school characteristics that predict success in academic outcome for ELL students?

1.f. Do student and/or school characteristics interact with a specific condition to predict success in academic outcome for ELL students?
Questions to Evaluate Goal 2. 2.a. Is there a difference in teachers’ instructional delivery between experimental and control classrooms as measured by TBOP?
2.b. Does treatment teachers’ instructional practice improve over time as a result of PD as measured by TOR?
2.c. What is teachers’ perceived effectiveness of the virtual PD and is there a preference in the type of PD delivery between virtual and face-to-face?
2.d. How do teachers perceive the quality of mentoring feedback through observation and through delivery of PD virtually?
2.d. What is principals’ perceived effectiveness of PD, interventions, observation tools, teacher implementation, and student achievement?
2.e. How sustainable is the PD after the intervention concludes at the end of each grade level?

Data Collection. Data collection by EG will occur in Years 1 - 5. At the teacher level, classroom observations and field notes will be conducted 3 times (beginning, mid, and end) of each school year in all classrooms. Other data including principal interviews, teacher surveys, and reflections in portfolios will be collected annually. During the Yr 1 of the project, the qualitative component of the project will be refined including measures, trustworthiness, and credibility of the data. Student scores on standardized assessments such as WMLSR and Stanford Achievement Test will be collected twice a year (i.e., beginning and end of each school year). Student scores on state-mandated assessments such as TELPAS will be collected at the end of each school year. Scores on the state-mandated STAAR will be collected in the spring of Year 2 for Gr. 3.

Data Analysis. Goal 1 research questions are addressed quantitatively. Since this is a two-level design study in which students are nested within school, hierarchical linear models (HLM) will be used to analyze each of treatment effects in each grade using SAS PROC MIXED (SAS
Institute, 1997). Students are the level-1 units and schools are the level-2 units. The treatment effect is defined at level 2. The dependent variables will be post test means, and the independent variables will be pretest means and grade. For the hierarchical analysis, the level 1 model is written as: \( Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Pretest})_{ij} + r_{ij} \), which represents the posttest achievement for school i in school j regressed on the pretest and grade in school. The term \( r_{ij} \) is the level-1 residual variance that remains unexplained after accounting for the covariates. At level 2 of the model, we estimate treatment effects on the mean posttest achievement outcome in school j. The fully specified level 2 model is written as: \( \beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Mean Pretest})_j + \gamma_{02}(\text{Treatment})_j + u_{0j} \) \( \beta_{1j} = \gamma_{20} \) where the mean posttest intercept for school j, \( \beta_{0j} \) is regressed on the school-level mean pretest scores and the treatment indicator, plus a residual, \( u_{0j} \). The pretest-posttest slope, \( \beta_{1j} \), is predicted by the treatment indicator and a school-specific residual and the grade slope is treated as a fixed effect. Outcome variables with pre and post assessment results (such as WMLSR, and Stanford Achievement Test) will be analyzed annually for each grade level. In cases where there are multiple measures within the same construct, composite scores will be calculated, or Benjamini-Hochberg correction procedure will be applied to take into account multiple outcome measures in the same outcome domain tested with multiple comparison groups, following the guidelines provided by WWC (2011). For outcome variables collected only once a year that do not have a pre-test (in the case of state-mandated assessments such as TELPAS), pre-test scores from the standardized assessments will be used as covariates to control for any potential baseline difference. Specific notations: For question 1.a., considering the limited English proficiency that ELLs possess at school entry, we will administer the Test de vocabulario en imagenes Peabody (TVIP) to measure students’ vocabulary skills in Spanish, their native language and use as a covariate. For questions 2.e., and 2.f., gender will be added as student level predictor and school
demographic data level 2 predictors. **Goal 2 research questions** will be addressed quantitatively and qualitatively. To answer question 2.a. of teacher effectiveness, chi-square test of homogeneity of proportion will be employed to identify if the proportion of each category under every domain in TBOP is homogenous between T and C classrooms. To address question 2.b., a paired-sample t-test will be performed to identify the improvement on TOR, if there is any, during the year of PD among T classroom teachers. To address questions 2.c to 2.e, field notes, classroom observations, mentoring, and PD, open-ended surveys (teachers and principals), and teacher reflections/portfolios will be analyzed using constant comparative method and will be coded according to themes for identifying trends or patterns. Data, researcher, and method triangulation and low inference descriptors will address credibility (internal validity) and reliability of the data. Focus groups with T teachers at K-3 will be conducted; patterns will be drawn, description of the relationships both formal and informal will be conducted, meanings both tacit and explicit will be sought, and the ability to implement and sustain such interventions within other schools will be analyzed. Policy implications will be sought.

**D.3. KEY ELEMENTS OF PROJECT TO REPLICATE IN OTHER SETTINGS.**

The independent evaluation will be directed by the project Logic Model (Figure 3) that outlines the goals, strategies, and expected outcomes of the ELLA-Validation project. The key elements and approach of ELLA was in an urban district, but in ELLA-V, there will be other urban, suburban, small town, and rural school schools included for testing/validating the intervention components. Additionally, an altered condition will be that the PD is delivered virtually and not on site which provides evidence for implementing ELLA-V PD across schools in multiple locales. An altered condition is the observation and mentoring of teachers via virtual tools which provides evidence for ELLA-V mentoring and provides validation of observation virtually.

**D.4. SUFFICIENT RESOURCES TO CARRY OUT THE EVALUATION EFFECTIVELY**
The independent evaluation will take place at Johns Hopkins University (JHU), one of the nation’s premier research institutions. The Center for Research and Reform in Education (CRRE) at JHU will provide strong and consistent organizational support. Resources follow:

**Organizational Support and Commitment** JHU provides multiple levels of support to the CRRE. CRRE, led by Dr. Slavin, has 5 Ph.D.’s and 5 other research and support staff engaged in a wide range of research involving children from preschool through high school who are in high-poverty communities. CRRE principal investigators are full-time researchers without teaching responsibilities who are therefore able to focus on high-quality longitudinal research, including many randomized and matched field experiments. **Program Assistance** JHU offices assist in the effective and efficient operation of the Center. JHU’s Homewood Research Administration administers grants and contracts; the Controller’s Office provides accurate and timely monthly statements for budgetary control; Human Resources assists in hiring and salary administration; Library Services provides a full range of assistance in research and document procurement, and the Office of Communications assists in national dissemination. **Resources** CRRE’s office is located off campus, and therefore qualifies for low off-campus overhead rates. JHU’s facility provides office administrative and clerical services; photocopying and conference rooms; and areas for maintaining and analyzing data, inventories of Center publications, and reference publications. **Evaluation Personnel for Effective Evaluation** Alan Cheung, PhD., (20% time annually; 30% fifth year) Associate Professor, CRRE at JHU, also a research fellow at the Center for Assessment Research and Development at the Hong Kong Institute of Education, has participated in large-scale national randomized field experiments, including randomized studies of bilingual education. He is an expert on data analysis and experimental designs and has worked on evaluations in China, Northern Ireland, and the U.S. He is expert in the use of Hierarchical
Linear Analysis (HLM), which will be used in the ELLA-V, and has published over 70 journal articles, book chapters, and 30 technical reports and was a recipient of the Palmer O. Johnson Award for the best journal article published in an AERA journal in 2008. He will develop and oversee an overall plan for all aspects of the ELLA evaluation, to ensure integration and quality of the evaluation as a whole. **Steven Ross, Ph.D.** (20% time annually), currently senior research scientist and professor at the CRRE at JHU, is the author of six textbooks and over 125 journal articles. He is Editor Emeritus of the research section of the *Educational Technology Research and Development* journal. He has received many honors, including the Eminent Faculty Award and the Lillian and Morrie Moss Chair of Excellence in Urban Education at The University of Memphis. He has testified on school restructuring research before the U.S. House of Representatives Subcommittee, has been a consultant to NSF on evaluation design, and is a technical advisor or researcher on current federal and state evaluations. He will assist on data collection, analysis, and reporting of the results, quantitatively and qualitatively. **Robert E. Slavin, Ph.D.** (5% time annually), Chairman of the Success for All Foundation, Director of the JHU CRRE, and Professor at the Institute for Effective Education at the University of York (England). He has carried out many rigorous field experiments, including randomized studies of Success for All, cooperative learning, peer tutoring, bilingual education, and quasi-experimental studies of Success for All, mastery learning, individualized instruction, and other interventions and has published more than 200 articles and 20 books on these and other topics, including educational psychology and research methods textbooks. He was PI on an IES-funded randomized evaluation of the Success for All programs and will serve as an advisor to ELLA-V.

**Research associates** Cynthia Lake (30% time annually), Mary Laurenzano (40% time annually),
and Beth Comstock (30% time annually) will provide data collection and data management support.

Figure 3. ELLA-V logic model.

References—found in Appendix J.10.