EXAMINING THE IMPACT OF VIRTUAL PROFESSIONAL DEVELOPMENT AND TEACHERS' USE OF THE COOPERATIVE/COLLABORATIVE/PEER-TUTORING STRATEGIES ON ENGLISH LEARNERS' READING COMPREHENSION, ORAL READING FLUENCY, AND ORAL EXPRESSION

A Dissertation

by

SHIFANG TANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Rafael Lara-Alecio
Committee Members,	Beverly J. Irby
	R. Malatesha Joshi
	Fuhui Tong
Head of Department,	Shanna Hagan-Burke

May 2018

Major Subject: Educational Psychology

ABSTRACT

The purpose of this dissertation was to investigate the impact of a curriculum-based virtual professional development (VPD) and teachers' use of the cooperative/collaborative/peer-tutoring (CCP) strategies on English learners' (ELs') reading comprehension, oral reading fluency, and oral expression. A research synthesis was conducted to comprehensively analyze the effectiveness of the CCP strategy for on ELs' reading and oral proficiency development in the elementary settings. The findings indicated that the CCP strategy facilitated ELs' improvement in English reading and speaking and ongoing professional development improved teachers' quality of implementation of this strategy as well as ELs' achievement.

To examine how ongoing curriculum-based VPD improved the teachers' application of the CCP strategy regarding the dynamics in EL classrooms, I utilized a large data set recorded by a multidimensional classroom observation instrument from a federal funded randomized controlled trial (RCT) study, Project English Language and Literacy Acquisition Validation Study (ELLA-V). Participants consisted of 77 first grade teachers from 41 schools in the state of Texas. They were observed three times a year, which resulted in 231 videos. Via a multi-layer cross-tabulation method, the teachers' application of this strategy was micro-analyzed regarding the dimensions of Instructional Language, students' Communication Mode, and teachers' Instructional Activities. Results indicated that treatment teachers spent more instructional time in applying this strategy in higher-order thinking content instruction and provided more opportunities for ELs to practice their listening and speaking skills in English in the CCP learning activities.

To investigate the impact of VPD and teachers' use of the CCP strategy on ELs' reading and speaking proficiency, I proposed a VPD solar system that was validated by large data sets

drawn from Project ELLA-V. There were 1,198 Grade 1 ELs and 77 Grade 1 teachers from 41 schools included in the analysis. Three-level hierarchical linear modeling (HLM) approach were adopted to examine the impact of VPD and the effective dosage of the CCP strategy applied in English dense cognitive content instruction. The results indicated that VPD significantly improved ELs' oral expression with medium to large effect size. Moreover, it was also found that the quality CCP learning that encourages ELs to engage in higher-order thinking activities in English needs to exceed 10.6% of total instruction time to be considered effective.

DEDICATION

I dedicate this dissertation to these very special people.

To my parents for their unconditional love.

To my husband for his unlimited understanding and support.

To my daughter for her precious companionship and laughter.

ACKNOWLEDGEMENTS

First, I acknowledge my gratitude to Dr. Lara-Alecio, my chair, for the great opportunity, ever-lasting support, and encouragement. His passion towards research will remain with me in my future career. I also thank my committee member, Dr. Tong, for all the new ideas and insights into the analysis and results. I would like to thank my committee member Dr. Irby for her support in the conceptualization of the model and valuable feedback. I would like to thank Dr. Joshi for his expertise in the reading field.

I would like to thank Drs. Lara-Alecio, Irby, and Tong for bringing project ELLA-V to Texas A&M University. It is an incredible platform, which enables me to learn, to explore, and to grow.

I would like to thank our academic advisor, Ms. Kallina, for her support to move the process smoothly.

I would like to thank the current and former staff members in Project ELLA-V: Cindy, Laura, David, Jui-Teng, Shuqiong, and Yue. Thank you for collecting and processing the large-scale data.

Lastly, I would like to acknowledge the support of the U.S. Department of Education. Although I was not directly funded by the Department, Chapters III and IV of my dissertation included archival data from Project English Language and Literacy Acquisition Validation Study (ELLA-V) that was funded by the Department.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Lara-Alecio [advisor] and Dr. Tong of the Department of Educational Psychology, Dr. Irby of the Department of Educational Administration and Human Resource Development, and Dr. Joshi of the Department of Teaching, Learning, and Culture. All work for the dissertation was completed independently by the student.

Funding Sources

This work was made possible in part by Office of Innovation and Improvement (i3), U.S. Department of Education under Grant Number U411B120047. Its contents are solely the responsibility of the author and do not necessarily represent the official views of the Office of Innovation and Improvement (i3), U.S. Department of Education.

NOMENCLATURE

ELLA-V	English Language and Literacy Acquisition - Validation
EL	English Learner
TBOP	Transitional Bilingual Observation Protocol
HLM	Hierarchical Linear Modeling
ESSA	Every Student Succeeds Act

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Background of the Study

There has been an increasing percentage of students in each grade level being identified as English learners (ELs) in K-12 U.S. public schools (National Center for Education Statistics [NCES], 2017). According to NCES (2017), 9.4% (or 4.6 million) of the public school enrolments were identified as ELs in the school year 2014-15, as compared to 9.1% (4.3 million) in the school year 2004-2005. In the school year 2014-15, a higher percentage of public school students were identified as EL in lower grades (NCES, 2017) than in upper grades. For example, one in every six Kindergartners or first graders was an EL, while in fifth, eighth, and twelfth grade, the ratio decreased to one in ten, one in sixteen and one in twenty-five, respectively. Among these 4.6 million ELs, 77.1% of them spoke Spanish as their home language (NCES, 2017). For over a decade, ELs have been lagging behind their native English-speaking peers (NCES, 2015). According to National Assessment of Educational Progress (NAEP) data in 2015, more than two-thirds of fourth-grade ELs were classified as below-basic reading level while less than one-third of non-ELs fell in this category; such achievement gap has remained consistent since 1998 (NCES, 2015).

The Every Student Succeeds Act (ESSA) signed by President Obama on December 10, 2015, requires that all schools have to demonstrate that they are improving ELs' English language proficiency (ESSA, 2015). Under ESSA, ELs' English language proficiency and academic achievement are integrated into an accountability system as an indicator for every public school in every state (ESSA, 2015). All states are required by ESSA to provide ELs with

the efficient and high-quality language instruction or instructional programs to ensure that ELs develop proficiency in both English language and content.

Meanwhile, states and districts face the challenges to comply with the new law, especially considering that ELs are a significant share of the student population, although ELs differ on backgrounds, native languages, and educational needs. However, they all have to take the same high-stakes tests as required by ESSA. According to Cummins' (1980) analysis, ELs need two to three years to approach native-level social communication proficiency, but as long as five to seven years to reach grade level on the content subjects. It is not surprising to see that ELs underperform their English-speaking peers in reading and other subject tests.

According to Krashen's (1985) second language acquisition theory, teachers can make language and content input more comprehensible for ELs via instructional activities and strategies. Among these strategies, the cooperative/collaborative/peer-tutoring (CCP) strategies, or peer-assisted strategy was recommended by a Practice Guide provided by Institute for Education Sciences for improving ELs' English reading proficiency (Gersten, Baker, Shanahan, Linan-Thompson, Collins, & Scarcella, 2007). Researchers also recommended that teachers should make sure that content-related cooperation/collaboration/peer-tutoring among ELs should be tailored, appropriate, and meaningful for ELs under direct instruction (Genesse, Lindolm-Leary, Saunders, & Christian, 2005). In two recent meta-analyses, the CCP strategy was found to be an effective instructional strategy to improve ELs' English language proficiency in reading (Cole, 2014) and speaking (Cole, 2013).

Cummins (1979, 1981) called educators' attention to the distinctions between Basic Interpersonal Communications Skills (BICS), a students' basic conversational fluency, and Cognitive Academic Language Proficiency (CALP), a students' ability to comprehend and

express the academic concepts and ideas in the school environment. Lara-Alecio and Parker (1994) further differentiated Cummins' two levels of language content into four levels: social exchange, academic transition, light cognitive content and dense cognitive content, with the first two under BICS and the latter two under CALP. Therefore, the quality of learning can also be classified into the four levels based on Lara-Alecio and Park's (1994) criteria.

Curriculum materials can improve students' learning process (Forbes & Davis, 2010; Schmidt, McKnight, & Raizen, 1997; Usiskin, 1985). However, curriculum alone would not automatically meet ELs' educational needs. Even when the research-based curriculum has a clear advantage, teachers might not benefit fully from the curriculum because of the complexity of the curriculum or the different levels of understanding or personal experience (McNeil, 2009). As suggested by Vygotsky (1987), interacting with an "expert" can lead to trainees' cognitive development. Taylor, Getty, Kowalski, Wilson, Carlson, and Van Scotter (2015) also emphasized that practitioners in the education field need evidence from the rigorous and comprehensive research-based trials that offer year-long or longitudinal professional development (PD) with multiple features integrated that can. The curriculum-based, ongoing PD is in need to effectively enhance teachers' understanding of the philosophy and critical component of the curriculum material (Tylor et al., 2015).

Researchers (i.e., Lara-Alecio, Tong, Irby, & Mathes, 2009; Lee & Buxton, 2013; Tong, Luo, Irby, Lara-Alecio & Rivera, 2017) have agreed that quality PD, scaffold teachers' pedagogy of ELs and quality of instruction. Furthermore, Bowers, Fitts, Quirk, and Jung (2010) pointed out that ongoing PD supported EL teachers to implement the research-based strategies, which were found to be effective to improve ELs' English language proficiency. However, data confirming the impact of PD on students' achievement and teachers' pedagogy and instructional quality

remain inadequate (Avalos, 2011; Lesaux & Gamez, 2012; Lewis & Shaha, 2003; Shaha, Lewis, O'Donnell & Brown, 2004; Tong et al., 2010; Tong, Luo, et al., 2017). It is partially because the traditional format of PD that requires teachers' absence from their daily routine and classrooms results in time burden, extra expense, and faculty reluctance (Shaha, Glassett, & Copas, 2015). Dede (2006) suggested that Online Professional Development (OPD) would be a more efficient and economical approach. Virtual PD (VPD), another form of OPD, provides online professional development and learning activities in real time virtually, which blends the traits of face-to-face PD and OPD (Bates, Phalen, & Moran, 2016).

The impact of OPD on students' academic performance has been found to be positive in an increasing number of studies (e.g., Dash, Magidin de Kramer, O'Dwyer, Masters, & Russell, 2012; Shaha & Ellsworth, 2013). However, the evidence is sparse about how VPD improves teachers' pedagogical practices, especially for ELs (Tong, Irby, & Lara-Alecio, 2015). Smith (2014) called for more investigation in this area with the solid evidence of the impact of effective VPD on teachers' classroom pedagogical practices and on students' academic achievement, including ELs'.

Also interesting, is how researchers measure classroom practices or teachers' pedagogy improvement as the results of VPD. There are mainly three approaches that have been adopted to document classroom practices: self-reports, interviews, and classroom observations (Durlak & Dupre, 2008; Lee, Penfield & Maerten-Rivera, 2009; Lillehoj, Griffin, & Spoth, 2004). Compared with data collected via self-reports and interviews, observational data are more objective and less susceptible (Lee et al., 2009; Noell, 2010) and have a stronger link to the outcome (Dane & Schneider, 1998; Durlak & Dupre, 2008; Dusenbury, Brannigan, Falco, & Hansen, 2003). It was suggested that classroom observation analysis (Snow, 2002; Turner &

Meyer, 2000) and video analysis (Snow, 2002; Stigler, Gallimore, & Hiebert, 2000) have led to improving educational research. Classroom observation provides a direct, objective, and systematic recording of classroom occurrences (Medley, 1992). Moreover, in the context of EL education, classroom observation helps to collect classroom practices of students, teachers, and student-teacher interaction (e.g., Haager, Gersten, Baker, & Graves, 2003; Lara-Alecio & Parker, 1994; Lara-Alecio et al., 2009; Tong, Luo, et al., 2017; Waxman, Tharp, & Hilberg, 2004; Waxman, Wang, Lindvall, & Anderson, 1983). Furthermore, classroom observation has been considered as the most reliable instrument to document classroom activities (e.g., Lara-Alecio et al., 2009) and to evaluate the effectiveness of teacher professional development (e.g., Tong, Luo, et al., 2017). However, classroom observation presents methodological challenges especially in large-scale intervention studies with diverse students (Lee et al., 2009). For example, analyzing video via observation instrument is most likely with multiple observers and coders (Borko, Jacobs, Eiteljorg, & Pittman, 2008), which requires sufficient training for better understanding the intervention focus (Lee et al., 2009) and acceptable inter-rater reliability (Tong, Irby, et al., 2017). It is time-consuming and resource-consuming for a large-scale study to develop and validate observation instruments, train research personnel to collect and analyze observational data, establish and monitor inter-rater reliability, further secure the validity and reliability of observational data analysis.

Significance of the Study

In his meta-analysis of the CCP strategy for ELs, Cole (2014) suggested that future researchers focusing on the CCP strategy should aim not only at determining the effectiveness of this strategy for ELs but also at identifying the processes and pedagogical factors that make this strategy effective for ELs. Bowman-Perrott, Mahadevan, and Etchells (2016) stressed that

although in-service PD should include this strategy with multi-tiered intervention supports, few evidence-based practices have been identified for ELs. The curriculum-based PD can stimulate teachers' cognitive development (Taylor et al., 2015) as well as scaffold teachers' understanding of the curriculum, the pedagogy of ELs and quality of instruction (Lara-Alecio et al.2009; Lee & Buxton, 2013; Tong, Luo, et al., 2017). Moreover, VPD, which is conducted virtually and synchronically, blends the characteristics of face-to-face PD and OPD and makes the distance between each teacher and the PD provider the same. Bowman-Perrott et al. (2016) pointed out that both teachers and ELs can benefit from the CCP strategy because of its flexibility and adaptation to the existing curriculum. Applying the CCP strategy in EL classrooms provides more opportunities for ELs to engage with academic content through reading, speaking, listening, and writing (Short, Fidelman, & Louguit, 2012). Based on Piaget's theory, the interactions among ELs in the English as a second language (ESL) classrooms stimulate ELs' cognitive development circle, therefore improve their English language proficiency.

The findings of my dissertation contribute to the literature in three ways. First, the findings add a research synthesis of the CCP strategy to serve ELs at the elementary level. Second, I used archival data from an ongoing RCT study to describe teachers' pedagogical practice, teachers' and students' language, and students' communication mode when teachers were employing this strategy during the light and dense cognitive content instruction, which provides a clearer picture of classroom occurrences when the CCP strategy was implemented. Third, by explicitly examining teachers' time allocation in implementing the CCP strategy among first grade ELs in a randomized control trial (RCT), in this dissertation, I identified the direct impact of intensive and ongoing curriculum-based VPD and the moderate effect of the quality CCP learning on ELs' English reading fluency, reading comprehension, and oral expression

proficiency, which served as a reference for ELs educators and researchers to find a viable solution to improving ELs' English language proficiency.

Purpose of the Study and Research Questions

The purpose of my study was to investigate the impact of VPD and teachers' employment of the CCP strategy on grade 1 ELs' improvement in reading and oral proficiency. Eight research questions guided this dissertation, which served three individual studies: (a) a research synthesis for the impact of the CCP strategy on ELs' English reading and speaking proficiency at elementary level, (b) a frequency assessment of teachers' pedagogical practices, students' communication mode, teachers'/students' language when the CCP strategy was employed in light and dense cognitive instructional time, and (c) an investigation of the impact of curriculum-based VPD and teachers' application of the CCP strategy on ELs' development in English reading comprehension, oral reading fluency, and oral expression. Within each of these studies, the following research questions were addressed:

1. How were ELs' reading and oral proficiency developed in the empirical studies with the CCP strategy in kindergarten through sixth-grade settings varying in sample characteristics, intervention features, design characteristics, outcome characteristics?
2. What is the impact of the CCP strategy on ELs' reading and oral proficiency?
3. What were the differences of time allocation in the CCP strategy between the teachers in the treatment (with the support of the curriculum-based VPD) and the control conditions (with the support of district-aligned PD) when the teachers were teaching different levels of language content?
4. What were the pedagogical differences (i.e., teachers' instructional activities and languages, and students' language modes and responding languages) between the treatment

teachers and the control teachers during the dense and light cognitive content instruction time when they employed the CCP strategy as documented by classroom observation instrument TBOP?

5. When the students were responding in English, did the instructional time allocated to quality CCP learning significantly differ between conditions (as a result of VPD) as observed in Project ELLA-V via the classroom observation instrument TBOP?

6. Did ELs' English oral expression, reading comprehension, and reading fluency achievement vary across different classrooms and schools?

7. Did ELs' English oral expression, reading comprehension, and reading fluency differ between the treatment (as a result of instructional intervention including VPD) and the control conditions, controlling for ELs' initial achievement on these measures?

8. What was the effective dosage of time allocation in quality CCP learning that contributed to the improvement of Grade 1 ELs' reading comprehension, reading fluency and oral expression?

Chapter Overview Structure of the Dissertation

Unlike the traditional five-chapter dissertation, a three journal-article format was adopted. In Chapter I, I provide the background of the dissertation, the significance of the dissertation, the general review, the reviews of each chapter, and the research questions. Chapters II, III, and IV were written in a journal-article format and are self-contained studies. Chapter V serves as a synthesis of Chapters II, III, and IV.

In my dissertation, Chapter I is the introduction and overview of the dissertation. Chapter II is a research synthesis on the impact of the CCP strategy on ELs' English reading and speaking proficiency at the elementary level, seeking to answer the first and second research questions.

Chapter II is a research synthesis of the studies that employed the CCP strategy on ELs to improve their English reading and speaking proficiency including reading comprehension, reading fluency, phonics reading, word reading/decoding, phonological awareness, oral language, and oral expression. Furthermore, it serves as a reference of how effective the CCP strategy improved elementary ELs' English reading and speaking proficiency for educators and researchers in EL education field. In Chapter III, I tried to answer the research questions 3 and 4 which were to examine classroom occurrences during the light and dense cognitive content instruction time in both treatment and control conditions when the CCP strategy was applied. The classroom observation instrument adopted in this dissertation was TBOP, which was a multidimensional occurrence-frequency measurement. Teachers' instructional activities and instructional languages as well as students' language modes and responding languages were examined and compared between the treatment and the control conditions during the light and dense cognitive instruction time. The differences between two conditions served to illustrate the impact of the ongoing and intensive curriculum-based VPD that Project ELLA-V provided to the treatment teachers. In Chapter IV, hierarchical linear models were adopted to examine the impact of the quality CCP learning on ELs' improvement in English reading and speaking proficiency. Quality CCP learning was evaluated via the low-inference classroom observation instrument TBOP. Quality CCP learning was defined as the time allocation in three conditions: (a) dense instructional time, (b) the CCP strategy, and (c) ELs’ responding in English. To be specific, it is the time that the teacher was using the CCP strategy to teach the dense cognitive content when the students were responding in English. In Chapter IV, the research questions 5-8 were addressed. Due to the nature of the sample which was: (a) students were clustered in different classrooms; (b) classrooms were nested in different schools; and (c) random selection were

placed at school level, a multi-level analysis was conducted to see the impact of the curriculum-based VPD on teachers' application of the CCP strategy and ELs' English reading and speaking proficiency. Further, by transferring the percentage of teachers' time allocation in the CCP strategy in English dense cognitive content instruction to a categorical variable, I further examined the effective dosage of the quality CCP learning on ELs' reading comprehension, oral expression, and oral reading fluency. Chapters II-V were addressed separately in the following sections.

Chapter II

Chapter II was a research synthesis titled "The Effects of the Cooperative/Collaborative/Peer-tutoring Strategies on English Learners' Reading and Speaking Proficiency in an English-medium Context: A Research Synthesis." It will be submitted to Reviews of Educational Research. In Chapter II, I provide a research synthesis to examine the effectiveness of the CCP strategy on improving ELs' English reading and oral proficiency through Kindergarten to Grade 6 by applying evidence standards. In this chapter, I primarily present a narrative review of related theories and summarize and critic the previous synthesis related to the CCP strategy on ELs' English reading and speaking proficiency. Moreover, in this study, I systematically analyze sample characteristics, intervention features, design characteristics, and outcome characteristics. Furthermore, in this study, I also provide a comprehensive picture of the phenomenon by statistically integrating effect size to quantify the effectiveness of this specific instructional practice/strategies.

Chapter III

Chapter III was named "Grade 1 Bilingual Teachers' Pedagogical Differences in the Application of the Cooperative/Collaborative/Peer-tutoring Strategies during CALP Instruction

in an RCT Study." In Chapter III, I present my investigation of observation data recorded by a multidimensional classroom observation instrument TBOP in Project ELLA-V, which documented teachers' pedagogical practices in English language literacy instruction for Grade 1 ELs.

Chapter IV

Chapter IV was entitled, "The Impact of the Cooperative/Collaborative/Peer-tutoring Strategies on Grade 1 English Learners' Reading and Oral Expression in an RCT Study: A Multi-level Analysis." In this study, I addressed the causal inference of VPD during Grade 1 implementation of a randomized control trial Project ELLA-V. Project ELLA-V was a two-level intervention, beginning from Grade 3 and working backward to Kindergarten, to determine: (a) the impact of the bi-weekly, ongoing, intensive and curriculum-based VPD on the ESL/bilingual teachers; (b) the impact of the intervention components on ELs' English reading and speaking proficiency.

In this study, a hierarchical linear modeling (HLM) approach was adopted to analyze the impact of the curriculum-based VPD and the CCP strategy on grade 1 ELs' English reading comprehension, reading fluency and speaking proficiency. In the HLM model, students' post-test score was treated as an outcome, students' pre-test score was included as a level-1 predictor, instructional time allocated in the CCP learning was included as a level 2 predictor, and condition (treatment teacher receiving VPD vs. control teacher receiving typical district PD) as a level-3 predictor. I further examined the effective dosage of the quality CCP learning on students' English reading and speaking proficiency.

Chapter V

Chapter V is a summary of Chapters II, III, and IV, in which I discuss the significance and key findings of these studies. I also provide implications and recommendations for further study.

CHAPTER II

THE EFFECTS OF THE COOPERATIVE/COLLABORATIVE/PEER-TUTORING STRATEGIES ON ENGLISH LEARNERS' READING AND SPEAKING PROFICIENCY IN AN ENGLISH-MEDIUM CONTEXT: A RESEARCH SYNTHESIS

Introduction

English learners (ELs) refer to the individuals who are "in the process of actively acquiring English and whose primary language is one other than English" (Bardack, 2010, p.7). According to National Center for Education Statistics (NCES, 2015), the enrollment of fourth grade ELs in the United States increased from 5% in 1998 to 10% in 2015. However, for over a decade, ELs have lagged in all content areas including reading, compared with their English-speaking peers (NCES, 2015).

The National Research Council (1998) stated that "reading is essential to success in our society" (p. 17). Learning to read is an essential and challenging process, especially for children at early grade levels (Rubin, 2016). This challenge is even more daunting for English learners (ELs) whose first language is not English (Ross & Begeny, 2011). Due to the difficulty in acquiring reading skills, ELs have been placed at a disadvantage in learning content subjects (August & Shanahan, 2006; Tong, Luo, et al., 2017), and they have a higher grade retention rate (U.S. Department of Education, 2016).

Oral language proficiency almost always develops ahead of reading and writing (Snow, Burns, & Griffin, 1998; Tinajero & Hurley , 2000), and can predict students' reading proficiency (Carlisle, Beeman, Davis, & Spharim, 1999; Miller, Heilmann, Nockerts, Iglesias, Fabiano, & Francis 2006; Proctor, Carlo, August, & Snow, 2005; Zhang, Anderson, & Nguyen-Jahiel, 2013). Moreover, Peregoy and Boyle (2005) point out that comprehensible input and social interaction

are the two most essential elements of oral language proficiency development. All the students, including ELs and non-ELs, need be surrounded by spoken English language to build their English oral language (Echevarria, Vogt, & Short, 2000).

Bowman-Perrott et al. (2016) pointed out ELs have specific learning needs, which make finding an effective instructional strategy of ELs more imperative. However, literacy instruction for ELs is too often teacher-led, whole class instruction (Zhang et al., 2013). ELs receive a low level of engagement instruction for more than 50-70% of instruction time (Arreaga-Mayer & Perdomo-Rivera, 1996; Simmons, Fuchs, Fuchs, Mathes, & Hodge, 1995) and are less engaged in English oral language practice and academic activities (Zhang et al., 2013). One-on-one instruction has been recommended as one of the most effective instructional practice for students with diverse linguistic and cultural backgrounds (Glass, Cahen, Smith, & Filby, 1982). However, it is impossible to provide this form of instruction to every student who is in need of this support during limited instructional time. Among evidence-based strategies that have been proven effective, the cooperative strategy, collaborative strategy and peer-tutoring strategy (the CCP strategy) allows individualization of instruction at different levels of content among pairs or small group of students. This strategy has been recommended as a cost-effective practice to support interaction among the students with various needs (August, McCardle, Shanahan, 2014; Fuchs, Fuchs, Mathes, & Simmons, 1997; Greenwood, Delquadri, & Hall, 1989; Roswal, Mims, & Evans, 1995). Moreover, highly-structured CCP learning enables each student to have opportunities to engage in classroom activities simultaneously (Bowman-Perrott, Mahadevan, & Etchells, 2016; Maheady, Mallette, & Harper, 2006; Short, Fidelman, & Louguit, 2012).

In the CCP learning setting, ELs are more likely to receive one-on-one feedback or correction as well as gain social support from their peers (Fuchs et al., 1997; Greenwood et al.,

1989).Therefore, students have more opportunities to engage in academic activities without adding more instructional time (Mathes et al., 2003). When carefully designed, this strategy has repeatedly demonstrated improvement in both ELs' and non-ELs' academic achievement (Almaguer, 2005; Bowman-Perrott et al., 2016; Calderón, Hertz-Lazarowitz, & Slavin, 2000; Calhoon, Al Otaiba, Cihak, King, & Avalos, 2007; Greenwood et al., 2001; Mathes, Fuchs, Fuchs, & Henley, 1994; McMaster, Kung, Han, & Cao, 2008; Saenz, Fuchs, & Fuchs, 2005; Shanahan, 1998; Short et al., 2012; Zhang et al., 2013).

August and Shanahan (2006) concluded in their synthesis of effective instructional practice that too few experimental studies existed to investigate the effectiveness of specific instructional practice, including the CCP strategy. Though in the past decade there have been an increasing number of empirical studies examining the effectiveness of instructional practice, few studies published in the peer-reviewed journals have investigated the impact of teachers' employment of the CCP strategy in EL classrooms in the United States, especially at the elementary school level (Cole, 2014). The purpose of my study was to further extend the CCP-related literature by providing the first research synthesis of the impact of the CCP strategy on ELs' reading and oral achievement at the elementary school level in the United States.

Theoretical Framework

The CCP strategy is a research-based ESL strategy that has been used to support students to learn academic content (King, Staffieri, & Adelgais, 1998). It is a strategy for quality instruction (Topping, Buchs, Duran, & Van Keer, 2017), which creates an optimal environment where students' differences of knowledge are not seen as a problem but as an opportunity for them to learn (Stainback & Stainback, 1992). Applying the CCP strategy in the EL classrooms to improve ELs' English language proficiency and academic knowledge is aligned with socio-

cognitive theory (Piaget 1932), sociocultural theory (Vygotsky, 1978), four-dimensional transitional bilingual pedagogical theory (Lara-Alecio & Parker, 1994), and second language acquisition theory (Cummins, 1980; Krashen, 1985).

Sociocultural Theories and Socio-cognitive Theory of Language Development, CCP Learning, and ELs

Ramirez, Yuen, Ramey, Pasta, and Bilings (1991) found that in teacher-dominated classrooms, ELs were not provided with enough opportunities to practice. They further specified that during half of the observed instructional time, ELs had no verbal responses. Juxtaposed with teacher-centered instruction, the CCP strategy developed from Vygotskian sociocultural theory of language development and Piaget's socio-cognitive theory of cognitive development, provides student-centered instruction or dialogical instruction to ELs to improve their literacy outcomes (Barton, 2009; Cole, 2014; Lee & Smagorinsky, 2000; Topping et al., 2017).

Vygotskian sociocultural theories defined the zones of unassisted performance and assisted performance with the former referring to the stage of content mastered in past development and the latter referring to the process of mastering content in the current development (Cole, 2014). Wood, Bruner, and Ross (1976) further construed the latter as scaffolding, a term used to describe the process that children were supported by adults or experts to build up their knowledge with utterance, gestures, or facial expressions. Second language researchers (i.e., Donato, 1994; Lantolf, 2000; Swain & Lapkin, 1995) have specified that language scaffolding could also be provided by EL peers. When ELs work together in a cooperative/collaborative manner, one can scaffold the other to co-build a performance which might exceed what is possible when students work individually.

Piaget (1932) emphasized the importance of cooperation among peers. He pointed out that the cognitive conflict between what the child has in his/her mind, and the new information that he/she comes across with his/her peers may prompt him/her to remove his/her misconception and switch to a more accurate conception, which is the process of cognitive development. Based on the works of Piaget and Vygotsky, Klingner and Vaughn (2004) summarized that for ELs, peer discourse could stimulate cognitive growth by creating conflicts during peer interaction in mutual problem-solving. As reported by Vaughn, Schumm, Kingner, and Saumell (1995), compared with assistance from adults, students prefer to receive support from their peers.

Second Language Learning Theories, the CCP Learning, and ELs

The CCP strategy is also linked to Cummins' (1980) theory of language proficiency, in which he proposed that language proficiency encompass two levels: Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP). According to Cummins, ELs need two or three years to master BICS, while they might need five to seven years to acquire CALP. To learn CALP, ELs need instruction that takes place in the context of academic language and academic content learning. The CCP strategy was recommended to provide cognitively demanding instruction to ELs at both BICS level (Baca & de Venezuela, 1998) and CALP level (Saenz et al., 2005).

Lara-Alecio and Parker (1994) proposed that CALP can be defined as two layers, the dense cognitive content and the light cognitive content, which further clarifies the levels of language content delivered in ELs classrooms. They argued, "the dichotomy between CALP and BICS has obscured the large amount of classroom communication which exists on a continuum between BICS and CALP" (p.122). According to Lara-Alecio et al. (2009) and Irby, Tong, Lara-

Alecio, Meyer, and Rodríguez (2007), how teachers allocate instructional time in ELs classroom is an essential indicator of the quality of instruction students receive. Therefore, it is more meaningful to examine teachers' instructional practice /strategies (e.g., the CCP strategy that occurs during higher quality instructional time—the dense cognitive content or CALP time).

According to Krashen's second language acquisition theory (1985), teachers can make language and content input more comprehensible for ELs via instructional activities and strategies (Krashen, 1985). As compared with whole-class undifferentiated instruction, one-on-one instruction or small group instruction has been recommended as one of the most effective instructional practice for students with diverse linguistic and cultural background (Elbaum, Vaughn, Tejero Hughes, & Watson Moody, 2000; Glass et al., 1982). Mathes et al. (2003) also argues that CCP learning is better than one-on-one or small-group instruction because it can support every student in the classroom instead of leaving some students without support during one-on-one or small group instruction.

Taken together, researchers from sociocultural, socio-cognitive, applied linguistics, and psycholinguistic fields suggested that the CCP strategy can scaffold and improve ELs language learning and cognitive development more than traditional teacher-centered or individualized instruction. Therefore, ELs who have been struggling with English language learning and academic achievement for decades could use support of this strategy to help them meet the demands of school.

Description of the CCP Strategy

The CCP strategy was broadly defined as "the instructional use of small groups in which students work together to maximize their own and each other's learning" (Johnson & Johnson, 1999, p. 73). Cole (2014) summarized the ESL cooperative learning strategy as three varieties:

(a) cooperative learning, (b) collaborative learning, and (c) peer tutoring. Cooperative learning was defined as emphasizing students' role in a carefully structured group instruction, which was referred to by Slavin (1996) as “one of the greatest success stories in the history of education research” (p.43). Collaborative learning is very similar to cooperative learning (Cohen, 1994), however, it has a smaller emphasis on students’ role in less-structured instruction (Cole, 2014; Topping et al., 2017). Peer tutoring varies widely, but in general, this strategy is highly academically structured and pairs an older or more capable student with a younger or less academically successful student (Cole, 2014; Topping et al., 2017). Although each of the three varieties is distinct by the supporting literature, in previous syntheses most researchers treated them as similar terms (Bowman-Perrott et al., 2016; Cohen, 1994; Slavin, 1996; Swain, Brooks, & Tocalli-Beller, 2002). In this dissertation study, the three varieties of this strategy were treated as one type of strategy, named in this dissertation as the cooperative/collaborative/peer-tutoring (CCP) strategy, in an effort to appropriately reflect as much literature as possible.

Previous WWC Reports, Systematic Reviews, and Meta-analyses on the CCP Strategy with ELs

In this section, I included three What Works Clearinghouse reports (WWC, 2007a, 2007b, 2010), two systematic reviews (i.e., Bowman-Perrott et al., 2016; Pyle, Pyle, Lignugaris/Kraft, Duran, and Akers, 2017), and two meta-analyses (Cole 2013, 2014). All the studies reviewed in these WWC reports were conducted at the elementary school level, including (a) studies related to bilingual cooperative, integrated reading, and composition (WWC, 2007a), (b) peer-assisted strategies for improving ELs' reading proficiency (WWC, 2010), and (c) peer tutoring strategies to improve ELs' academic achievement by grouping English native speakers, bilingual students, and ELs to work on academic tasks (WWC, 2007b). Out of eight studies

evaluated by three WWC reports, four studies met WWC standards without reservation, and one met WWC standards with reservation. In the WWC report titled "Bilingual Cooperative Integrated Reading and Composition" (WWC, 2007a), the evaluators concluded that based on one study that met WWC standards with reservation, cooperative strategy in bilingual classrooms help ELs with their reading achievement and English language development. In the WWC report titled "Peer-Assisted Learning Strategies" (WWC, 2010), evaluators included four studies. However, only one study met WWC standards. Based on this study, WWC evaluators concluded that peer-assisted strategies might have potential positive effects on ELs' reading achievement. In the WWC report titled “Peer Tutoring and Response Groups” (WWC, 2007b), the evaluators reviewed three studies that met standards, but none of them addressed ELs’ reading achievement.

Two meta-analyses conducted by Cole (2013, 2014) examined ELs' achievement outcomes at all school levels. Cole (2013, 2014) differentiated peer tutoring, collaborative strategy, and cooperative strategy. In Cole’s (2013) meta-analysis, he examined the effect of peer-assisted strategies including peer tutoring, collaborative strategy, and cooperative strategy on ELs' oral and writing proficiency. Peer tutoring was analyzed separately from collaborative and cooperative strategies. He included 32 studies from 12 different countries at the elementary and secondary school level and found that overall peer-assisted learning had a positive mediating effect on ELs' oral and writing proficiency. Out of these 32 studies, only 6 were published in peer-reviewed journals and focused on the elementary level in the United States. However, it is unclear which studies were included in the meta-analysis since he did not specify that in the references. Moreover, in the studies that were included, peer-assisted learning was not always the

intervention-focused strategy in the classrooms. Therefore, the effect of peer-assisted learning might be confounded by other instructional strategies applied in the same studies.

Cole (2014) examined 28 studies in a meta-analysis that tested the effect of peer-mediated strategies including cooperative, collaborative, and peer-tutoring on ELs' literacy outcomes. The overall finding was that peer-mediated strategy was more effective for ELs as compared with individualized or teacher-centered instruction. He also identified that peer-mediated strategies worked better at the elementary level as compared with middle and high school level. Moreover, as compared with collaborative strategy and peer-tutoring, the cooperative strategy had the highest effect size. He suggested future studies need to investigate the process and pedagogical factors that make peer-mediated strategies effective for ELs. Notably, out of 28 studies that were included in the analysis, only nine were conducted at the elementary level in the United States with only two of them published in peer-reviewed journals, which were coded as low quality in this meta-analysis.

In Bowman-Perrott et al.'s (2016) systematic review, they examined the impact of peer-tutoring on ELs' academic, social, and linguistic outcomes. They examined 17 studies from pre-K to grade 12 published in the last 40 years. It was found that cross-age and same-age peer-tutoring had a positive impact on ELs' English language proficiency. Among these 17 studies, 12 were conducted at the elementary level. The authors suggested school level should be examined in the future systematic reviews since ELs have a different level of needs regarding the content vocabulary and concepts, especially between elementary and secondary level. Further, the duration of tutoring was also recommended to be investigated in the future systematic reviews because the relation between duration and impact of peer-mediated strategies on ELs' outcomes remains unclear. Overall, Bowman-Perrott et al. (2016) focused on the impact of the CCP

strategy on ELs' academic achievement rather than ELs' language proficiency. However, the countries where these studies were conducted were not reported. Further, the specific grade level was not mentioned.

In a systematic review by Pyle et al. (2017), researchers examined the impact of the peer-mediated intervention (PMIs) on ELs from Kindergarten to grade 12. Fourteen peer-reviewed journal publications from 1983 to 2013 were included. The authors analyzed selected studies regarding intervention characteristics, the effect of PMIs on ELs' academic outcomes, and methodological quality. They also compared the effectiveness of PMIs by types of peer-assisted learning (i.e., pairing and cooperative/collaborative). Pyle et al. (2017) emphasized fidelity of implementation as an importation quality indicator. Therefore, only eight studies had high methodology quality. It was found that PMIs could support ELs' development in phonemic awareness, vocabulary, and comprehension. The authors called for future studies to investigate the impact of PMIs on improving ELs' language proficiency, including speaking, listening, reading, and writing. A limitation in the Pyle et al. (2017) study was that their research synthesis did not follow any standard protocol or procedure, creating concerning gaps.

For example, whether and how the professional development was provided in the intervention studies for participating teachers to improve the quality of implementation was not examined (i.e., Cole, 2013, 2014; Pyle et al., 2017). Fidelity of implementation was also neglected in the meta-analysis studies (i.e., Cole, 2013, 2014). In all the previous reviews, the researchers focused more on the impact of different types of peer-assisted learning on ELs' achievement instead of differentiating the impact of peer-assisted learning on different language proficiency (i.e., speaking, reading, listening, and writing), especially when they included

multiple school levels (i.e., elementary, middle, and high) and different intervention focuses (e.g., writing, reading, and social science).

Purpose of the Review

The purpose of this research synthesis was to extend the CCP strategy literature by providing the first synthesis of the impact of the CCP strategy with the focus on elementary ELs' reading and oral proficiency in the United States. Sample characteristics, intervention features, design characteristics, outcome characteristics, and the impact of this strategy on ELs' reading and oral language were fully examined. Effect sizes and their confidence intervals (CI) were calculated and reported when sufficient data allowed. This work addressed a gap in the CCP strategy literacy by providing needed information about its effectiveness for elementary-age ELs' English reading and oral proficiency.

Method

The purpose of this study was to examine the impact of the CCP strategy on elementary school ELs' oral and reading proficiency. Therefore, a logic model was constructed in Figure 1 to synthesize the studies that investigated the impact of the CCP strategy on ELs' reading and speaking. The approach first started broadly by bringing together studies where the CCP strategy was employed to improve ELs' English reading and oral proficiency. Then, related studies were examined for the characteristics of participants (i.e., grade level, sample size, primary language), research designs (i.e., experimental, quasi-experimental, assessment, intervention implementation, intervention fidelity, teacher practices evaluation), and outcomes (i.e., reading domain, assessment, effect size).

Search Procedure

The constructed logic model guided this research synthesis in collecting and analyzing the effect of the CCP strategy in intervention studies for ELs' reading and oral proficiency. The general process involved an extensive literature collected from five databases Educational Resources Information Center (ERIC), PsycINFO database, Education Full Text, Academic Search Ultimate, and Scopus. Exclusive and inclusive criteria were applied to select the most related studies. Coding sheet included four parts: sample characteristics, intervention features, design characteristics, and outcome characteristics. Selected studies were coded with the coding sheet, and research synthesis was based on coding findings. The flow of arrows in Figure 1 demonstrates the process to search, screen, and code the related studies. Table 1 showed the results after all inclusion/exclusion criteria were applied.

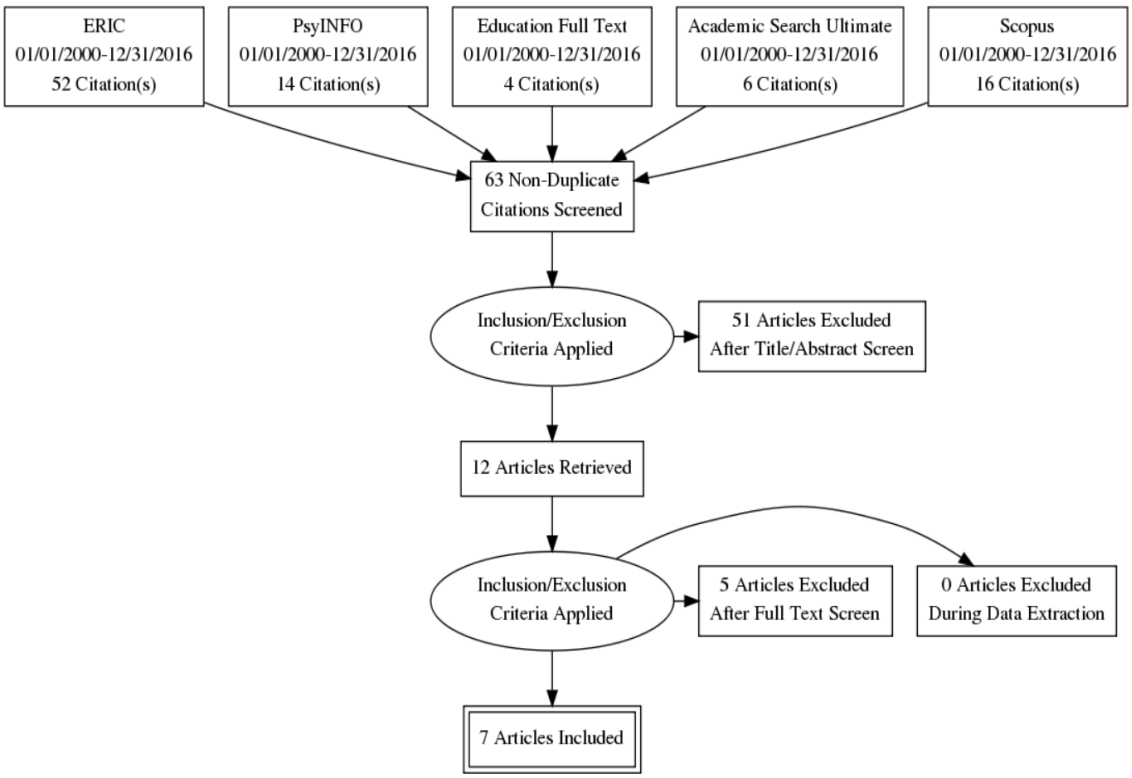

Figure 1. The flowchart of research synthesis.

Table 1
Selection Results after Inclusion/Exclusion Criteria Applied

Category	Number
Included in the final analysis	8
Review or introduction	3
Not elementary Level	2
Not a journal paper	12
Not ELs	6
Not peer-assisted	27
Not about reading or speaking	2
Not in the U.S.	3
Not in the year range	1

By following the steps in Figure 1 and inclusive and exclusive criterion, I included the studies in the final dataset to answer the following two research questions:

1. How did empirical CCP studies developing English reading and oral proficiency with ELs in kindergarten through sixth-grade settings vary in their methodological characteristics (i.e., sample characteristics, intervention features, design characteristics, and outcome characteristics)?

2. What was the impact of the CCP strategy on ELs' reading and oral proficiency?

In this review, I extended the findings of the previous systematic reviews, meta-analyses, and WWC reports summarized in the previous section. I included all the studies focusing on the CCP strategy applied on ELs with a focus on improving their English reading and speaking proficiency. I followed the suggestions recommended by Cook, Mulrow, and Haynes (1997) and Garg, Hackam, and Tonelli (2008) to appraise the features of research/intervention design, summarize the results and the critical findings, identify reasons for different results across

studies, and relate syntheses with current knowledge. Since I was particularly interested in the impact of the CCP strategies on ELs' reading and oral proficiency at the elementary level in the United States, the effect size of each study is reported. In order to ensure that this study was conducted rigorously and systematically, Cooper's (1982) standard procedures for research synthesis were followed: (a) problem formulation stage, (b) data collection stage, (c) data evaluation stage, and (d) analysis and interpretation stage.

Problem Formulation and Data Collection

First, in problem formulation stage, I reviewed and synthesized past systematic review, WWC reports, and meta-analyses on the CCP strategy with ELs. In this stage, variables of interest were defined. Second, in data collection stage, four databases, Educational Resources Information Center (ERIC), PsycINFO, Education Full Text, and Academic Search Ultimate, were employed. A comprehensive computerized search on the CCP strategy on ELs' reading and oral proficiency was conducted. Searching ranged from January 1, 2000, to December 31, 2016. Search terms and variable combinations included the notions of cooperative strategy, collaborative strategy, peer tutoring, elementary school, English language learners, ESL, bilingual, reading, and oral proficiency. Furthermore, the Scopus database was used for an additional subject search (Sampson, McGowan, Cogo, & Horsley, 2006). With this step, I was able to include the relevant studies which cited the studies identified by the previous four databases. As shown in Figure 1, these searches yielded a total of 92 publications from which 29 were identified as duplicates. There were 63 publications included in the data evaluation stage.

Data Evaluation and Inclusion and Exclusion Criteria

A more in-depth examination of the remaining publications was conducted by applying inclusion/exclusion criteria. Selection criteria for inclusion were (a) directly related to the topic,

i.e., those involved ELs' English reading or speaking proficiency and used the CCP strategy, (b) conducted in the United States, (c) limited to those published in English, and (d) English as the major medium language in interventions or program. Furthermore, the literature reviews and the meta-analysis studies related to the topic were summarized in the previous synthesis section but were not included in the coding procedure. The studies were excluded if they were: (a) not published in the peer-reviewed journals or (b) at conducted with secondary, postsecondary or adult learners.

Details of Study Coding Categories

Various study characteristics were coded to answer the first research question related to sample characteristics, intervention features, design characteristics, and outcome characteristics before evaluating the level of effectiveness of the CCP strategy and its relative impact on ELs' reading and oral proficiency. The codes were classified into four categories. First, *sample characteristics* included age/grade, primary language background, language groups, social-economic status. Second, *study/intervention features* included study/intervention focus, instructional time (duration, frequency, intensity, and total time), group size, use of randomized assignment, fidelity checking, and professional development. Third, *design characteristics* included randomization, pre- and post-testing measures, attrition rate, baseline equivalence, demographic and instructional characteristics of comparison group, and data analysis methods. Fourth, *outcome features* included target language proficiency (e.g., phonological awareness, word reading, reading fluency, reading comprehension, and oral language), assessment type (e.g., standardized test, curriculum-based test, researcher-created test, and high-stakes test), reliability and validity of outcome measures, the impact of the CCP strategy on ELs' reading and

oral proficiency, and implication for the future studies. All four features were assigned a group if applicable.

Furthermore, the included studies were further analyzed to see how strong the effect of the CCP strategy on students' outcomes was. Cohen's *d* was reported as the indicator of effect size. The effect sizes were not calculated on some studies for one or both of the following reasons: a) the authors did not provide sufficient or accurate data or b) there was no significant difference between the treatment and the control groups. The effect sizes were calculated according to formulas published by Lipsey and Wilson (2001) with an online calculator developed by Wilson (n.d.), which has also been widely used for other systematic reviews and meta-analysis studies (e.g., Maher, Lewis, Ferrar, Marshall, De Bourdeaudhuij, & Vandelanotte, 2014; Rubak, Sandbæk, Lauritzen, & Christensen, 2005). The interpretation of Cohen's *d* was based on the recommendation of Cohen (1988): small, $d = .2$; medium, $d = .5$; and large, $d = .8$.

Results

Seven journal publication met inclusion criteria and were included in the analysis. The studies in this review were published from 2001 to 2016 (M=2007.7, *S.D.* = 4.86). Notably, most studies were published over the past decade, showing more significant interest in this field over the span of ten years from 2001 to 2010 (*n*= 5).

The following sections address the research questions. To answer the first research question, I presented a summary of sample characteristics, intervention features, design characteristics, and outcome characteristics. The second research question was addressed by evaluating the effectiveness of the studies with effect sizes drawn from available data.

Sample Characteristics

A total of 961 ELs participated across seven studies, with a range of 75 to 351 participants per study (Mean = 137.3, *SD* = 96.4). Terms used to describe the participants across the studies included English language learners (Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; Liu & Wang, 2015; Saenz et al., 2005; Zhang et al., 2013) and ELs (McMaster et al., 2008). Although various terms were adopted to define this particular population, the participants in these studies similarly referred to students who first language was not English and were identified as ELs based on home language survey (McMaster et al., 2008; Zhang et al., 2013), district level information (Liu & Wang, 2015; Saenz et al., 2005), and/or if they did or did not meet a certain level of English language proficiency assessment (Almaguer, 2005; Calhoon et al., 2007; Saenz et al., 2005). In one study, the resesarchers did not disclose how ELs were defined (Greenwood et al., 2001). The participants in the included studies received instruction in a variety of program models, ESL (Greenwood et al., 2001), bilingual education (Almaguer, 2005; Calhoon et al., 2007; McMaster et al., 2008; Saenz et al., 2005), both bilingual and mainstream program (Zhang et al., 2013), and in two studies the reserachers did not specify the program models (Liu & Wang, 2015; McMaster et al., 2008).

Grade levels. The distribution of age groups for the included studies was prescreened and limited to the elementary level. In six studies, the instruction was delivered in a single grade level: Kindergarten (McMaster et al., 2008), grade 1 (Calhoon et al., 2007), grade 3 (Almaguer, 2005), grade 4 (Liu & Wang, 2015; Saenz et al., 2005), and grade 5 (Zhang et al., 2013). In one study, multiple grade levels were involved (Greenwood et al., 2001). Four CCP studies focusing on the upper grades (3-5) indicated a great interest in the CCP research among this grade group. There were 112, 123, 24, 100, 506, and 96 ELs in grades K to 5, respectively.

Language groups. The participants in three studies were all Spanish-speaking ELs (n=3; Almaguer, 2005; Saenz et al., 2005; Zhang et al., 2013). A mix of language samples were identified in four other studies (n=4; Calhoon et al., 2007; Greenwood et al., 2001; Liu & Wang, 2015; McMaster et al., 2008), with one study being Spanish-speaking dominant (n=1; Calhoon et al., 2007). The authors of studies with mixed language sample did not disclose what languages ELs spoke other than Spanish. However, as suggested by the demographic information, ELs in these studies also include Asian, African American, European American, Brazilian, and Pacific Islander.

SES background. In three studies, the authors reported the SES background of participants. In one study, the researchers specified that the participating school served low to middle SES families (n=1; Zhang et al., 2013). It was stated in two studies that the participating schools were located in Mexico-U.S. border area and were classified as high-poverty (n=2; Almaguer, 2005; Calhoon et al., 2007), among which one specified that more than 80% of students receive free or reduced lunch (Calhoon et al., 2007). SES information of the participants was not disclosed in four studies (Greenwood et al., 2001; Liu & Wang, 2015; McMaster et al., 2008; Saenz et al., 2005).

Study/intervention Features

Several study/intervention features were coded as a means of understanding how the CCP strategy may impact ELs' English reading and oral proficiency. These features included the focus of study/intervention, group size, instructional time, frequency and density, L1 support, the CCP strategy application during instruction, fidelity of implementation and professional development/support provided to teachers' delivering of instruction.

The focus of study/intervention. The researchers of six included studies focused on more than one aspect of ELs' reading/oral proficiency (n=6). In three studies, the authors focused on oral reading fluency (n=3; Almaguer, 2005; McMaster et al., 2008; Saenz et al., 2005). The researchers of four studies focused on reading comprehension (Almaguer, 2005; Liu & Wang, 2015; Saenz et al., 2005; Zhang et al., 2013). In two studies, the researchers investigated ELs' phonemic awareness, letter-word identification (n=2; Calhoon et al., 2007; McMaster et al., 2008). The studies conducted in lower grades (Grades k-1) had a focus on letter name/sound and phonemic awareness (e.g., Calhoon et al., 2007; McMaster et al., 2008), while in the studies conducted in higher grades (Grades 3-5) the researchers focused on higher level English proficiency like reading comprehension (Almaguer, 2005; Liu & Wang, 2015; Saenz et al., 2005; Zhang et al., 2013). Oral reading fluency was measured through lower grades and upper grades (Almaguer, 2005; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013).

Instructional time. The amount of time devoted to the CCP strategy instruction varied in each study based on its duration, frequency, and intensity. The duration referred to the number of weeks in which the CCP embedded instruction was delivered. The researchers of the included studies reported the duration of the implementation which ranged from four weeks to an academic year. The intervention in four studies was implemented from 15 weeks to 20 weeks (n=4; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005). The intervention in one study was implemented for only four weeks (Zhang et al., 2013) and in one study for nine weeks (Almaguer, 2005). Within each week, the instruction was delivered between 2 days a week to 5 days per week (M=3.5, SD=1.05). The intervention was implemented two times per week in one study (n=1; Zhang et al., 2013), three times per week in two studies

(n=2; Calhoon et al., 2007; Saenz et al., 2005), four times per week in two studies (n=2; Greenwood et al., 2001; McMaster et al., 2008), and five times per week in one study (n=1; Almaguer, 2005). The intensity of instruction was defined as minutes of the CCP strategy embedded instruction per day. The researchers of six included studies reported intensity of the CCP strategy embedded instruction. The intensity of six studies ranged from 20 to 35 minutes (*M*= 27, *SD*= 6.41). Liu and Wang (2015) drew EL data from two large data sets Progress in International Reading Literacy Study (PIRLS) and National Assessment of Educational Progress (NAEP). Liu and Wang (2015) coded frequency of utilizing small group instruction, pair work, and independent reading. For example, in the PIRLS dataset, the frequency for each type of instruction was coded as "1=never or almost never; 2= once or twice a month; 3=once or twice a week; and 4= every day or almost every day". However, the summary of how often each instruction was employed in EL classroom was not reported in Liu and Wang's (2015) study.

Group size. In six studies, the instruction was organized in a whole group or an entire class (n=6; Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013). In one study, the authors investigated the frequency of three types of grouping: small group, pair work, and independent reading in large-scale data sets (Liu & Wang, 2015).

Measures of the level of exposure/fidelity of implementation. The measures of exposure to the CCP strategy were coded by what means that researchers used to record ELs' exposure to this strategy besides duration, frequency, and intensity reported previously in this chapter. In five studies, the researchers checked fidelity of implementation to document how much the CCP strategy was implemented with ELs during the intervention (n=5; Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005).

In one study, the researcher used student attendance records of intervention as FOI, which was above 91% during the period of intervention (Almaguer, 2005). In four studies, classroom observation checklists were adopted to record fidelity of implementation (Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005). Fidelity of implementation ranged from 80% to 96%. In one of these four studies which was a sub-study of a large-scale randomized control trial, the researchers only reported data on the teachers whose fidelity of implementation was 0.9 and above (McMaster et al., 2008). In one study, the researchers did not report fidelity of implementation (Zhang et al., 2013). In Liu and Wang's (2015) study, they did not involve intervention, so the fidelity of implementation was not applicable.

Professional development. Professional development components were reported in six intervention studies (n=6; Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013). Additionally, in all six studies, researchers confirmed that the initial professional development session was conducted before the study/intervention began (n=6; Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013). In two studies, the teachers were offered professional development training and coaching at the before and throughout the intervention (n=2; Calhoon et al., 2007; Greenwood et al., 2001). To be more specific, in one study, participating teachers received individualized 1-hour PD session for 5-7 times then received biweekly PD consultation (Greenwood et al., 2001). Researchers in the other study offered on-going feedback and support three times a week via classroom observation (Calhoon et al., 2007). All professional development/training sessions in the included studies were traditional/face-to-face (Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013).

Design Characteristics

To avoid bias, I analyzed all the included studies' design characteristics in the following sections: random assignment, attrition rate, baseline equivalence, comparison group, and data analysis.

Random assignment. Three studies were identified as using quasi-experimental design, in which the researchers did not use random selection to assign multiple groups to a different instructional approach (n=3; Almaguer, 2005; McMaster et al., 2008; Zhang et al., 2013). In these three quasi-experimental studies, a pretest and a posttest were administered. Two studies were identified as a randomized control study, in which the researchers randomly assigned the groups to treatment or control conditions (n=2; Calhoon et al., 2007; Saenz et al., 2005). In both studies, randomization was conducted at the classroom/teacher level. Pretests and posttests were also administered to participating students. One study was a single-subject design, which only had one group of students receive the intervention (n=1; Greenwood et al., 2001). The participants served as their own control comparison. In this study, both pretest and posttest achievement were recorded. One study did not involve any intervention (n=1; Liu & Wang, 2015). EL data were drawn from large-scale datasets, the Progress in International Reading Literacy Study (PIRLS) and the National Assessment of Educational Progress (NAEP). Researchers investigated the impacts of different activities (small group instruction, pair-work instruction, and independent reading instruction) on ELs' reading proficiency.

Attrition rate. In two studies, the researchers provided attrition information (Calhoon et al., 2007; Saenz et al., 2005). Based on WWC's (2015a) Standard Brief for Attrition, the attrition rate was 23.7% in Calhoon et al. (2007) and 10.9% in Saenz et al. (2005). However, no researchers reported the number of students who withdrew from intervention by the condition in

the included studies. Therefore, the equivalence of attrition rate between conditions could not be examined.

Baseline equivalence. Due to the potential lack of baseline equivalence among conditions, quasi-experimental studies should assess whether the equivalent baseline was achieved between the treatment and the control conditions (WWC, 2015a). Only in one quasi-experimental study did the researchers examine the baseline equivalence and report that there was no statistically significant difference between the treatment and the control conditions regarding students' pretest achievement (n=1; Zhang et al., 2013). In the other two quasi-experimental studies, although the researchers did not examine the baseline equivalence, they used a statistics technique, Analysis of Covariance (ANCOVA), to adjust the baseline difference (n=2; Almaguer, 2005; McMaster et al., 2008). In two randomized experimental studies, the researchers examined the baseline equivalence and reported there was no statistically significant difference between the treatment and the control conditions (n=2; Calhoon et al., 2007; Saenz et al., 2005). There were two studies not eligible for reporting baseline equivalence (n=2; Greenwood et al., 2001; Liu & Wang, 2015). Greenwood et al. (2001) was a single subject design and Liu and Wang's study (2015) was correlation study with no intervention involved. Therefore, baseline equivalence was not necessary for these two studies.

Comparison group type. Two studies were not eligible for reporting comparison group type (n=2; Greenwood et al., 2001; Liu & Wang, 2015). Greenwood et al. (2001) was a single subject design, which used the subjects' pretesting achievement as the comparison. Liu and Wang (2015) utilized large-scale data sets which had no intervention involved. In four studies, the researchers reported and compared demographic data and instruction practice in both treatment and control conditions (n=4; Almaguer, 2005; Calhoon et al., 2007; McMaster et al.,

2008; Saenz et al., 2005). In one study, researchers reported and compared the demographic data of treatment and the control conditions but did not report the instructional practices in the control condition (n=1; Zhang et al., 2013). The comparison group received the same instructional time as the treatment group in all studies except for Almaguer's (2005) study in which an extra 30 minutes of instruction was offered to the treatment students.

Data analysis. To investigate the impact of the intervention or the CCP strategy on ELs' reading and oral proficiency, the statistical tests in seven studies included Analysis of Variance (ANOVA; n=2, Calhoon et al., 2007; Saenz et al., 2005), Analysis of Covariance (ANCOVA; n=3; Almaguer, 2005; McMaster et al., 2008; Zhang et al., 2013), correlation analysis (n=1; Liu & Wang, 2015), and a chi-square analysis (n=1; Greenwood et al., 2001). Single-level analyses at the student level were adopted in five studies (n=5; Almaguer, 2005; Calhoon et al., 2007; Liu & Wang, 2015; McMaster et al., 2008; Zhang et al., 2013). Aggregated class mean or teacher mean was used in two studies (n=2; Greenwood et al., 2001; Saenz et al., 2005). No cluster-level analysis was conducted in any of these seven studies.

Outcome Characteristics

In this review, outcome measures that were used to assess ELs' English reading and oral language proficiency were examined by assessment type and target language proficiency. Description of the outcome measures included assessment types used (i.e., standardized, high-stakes, curriculum-based, or researcher-created) and language proficiency measured (i.e., phonological awareness, reading comprehension, oral expression, or reading fluency). Reliability and validity of outcome measures were also examined.

Assessment types and ability measured. The researchers adopted curriculum-based measures to assess ELs' phonological awareness (n=1; Greenwood et al., 2001) and oral reading

fluency (n=1; McMaster et al., 2008). Standardized measures were used to measure ELs' letter-word identification and word fluency (n=2; Calhoon et al., 2007; McMaster et al., 2008), reading comprehension (n=2; Liu & Wang, 2015; Zhang et al., 2013), and reading fluency (n=1; Calhoon et al., 2007). In four studies, the researchers used the researcher-created assessment to measure students' reading comprehension (n=3; Almaguer, 2005; Saenz et al., 2005; Zhang et al., 2013), phonemic awareness (n=1; McMaster et al., 2008), oral expression (n=1; Zhang et al., 2013), oral reading fluency (n=2; Almaguer, 2005; McMaster et al., 2008), and reading comprehension (n=1; Almaguer, 2005).

The same measures were adopted for measuring both pretest and posttest scores (n=5; Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005). In two studies, the researchers used different measures for pretest and posttest (n=2; McMaster et al., 2008; Zhang et al., 2013). In one study, there was one-time testing was administered (n=1; Liu & Wang, 2015). Moreover, in three studies, the researchers used multiple measures assessing ELs' different abilities including reading comprehension, phonemic awareness, oral reading fluency, and oral expression (n=4; Calhoon et al., 2007; McMaster et al., 2008; Zhang et al., 2013). There were four studies only involving a single measure to measure an aspect of ELs' language proficiency such as reading comprehension (n=3; Almaguer, 2005; Liu & Wang, 2015; Saenz et al., 2005) or reading vocabulary (n=1; Greenwood et al., 2001).

Reliability of outcome measures. In two studies, the researchers reported the reliability of their standardized measures, ranging from 0.9-0.96 (n=2; Calhoon et al., 2007; McMaster et al., 2008). Among these two studies, McMaster et al. (2008) reported the concurrent validity as 0.9. The researchers (Liu & Wang, 2015; Zhang et al., 2013) did not report the reliability or validity of the standardized tests in their studies (n=2;). It was reported in two studies that the

test-retest reliability of researcher-created measures was 0.93-0.95 and concurrent validity was 0.91(n=2; Almaguer, 2005; Saenz et al., 2005). In one study, the researchers reported their researcher-created measures with test-retest reliability as 0.47-0.87 without validity reported (n=1; Zhang et al., 2013). The researchers of two studies did not report the reliability and validity of their curriculum-based measure (n=2; Greenwood et al., 2001; McMaster et al., 2008) and researcher-created measures (n=1; McMaster et al., 2008).

The Impact of the CCP Strategy on ELs' Reading and Speaking Proficiency

Oral language. In one study, the researchers investigated the impact of collaborative reasoning (CR) on ELs' English oral language development (n=1; Zhang et al., 2013). They found that CR significantly improved participants' (both ELs and Native-English speakers) coherent narratives in the storytelling task (Cohen's d = 0.37, 95% CI [-0.09, 0.83]). The authors reported that ELs produced significantly longer stories than Native-English speakers did, but spoke slower and with more mazes and paused more often, but there was no available data for calculating the effect sizes for these two subcategories. There was no interaction effect between the CR intervention and the program type (bilingual vs. mainstream). The authors did not report the comparison between the CR ELs and control ELs regarding their English oral language development. Therefore, it was unclear how strong the CR learning impacted ELs' English oral language development.

Oral reading fluency. The researchers of three studies examined the impact of the CCP strategy on improving ELs' reading fluency (n=3; Almaguer, 2005; Calhoon et al., 2007; McMaster et al., 2008). There were mixed results in the treatment-control comparison regarding students' achievement in oral reading fluency. Almaguer et al. (2005) found that the CCP reading activities significantly improved Grade 3 ELs' reading fluency with large effect

(Cohen's *d* = 2.53, 95% CI [1.70, 3.37]) based on Cohen's criterion (1988). However, Calhoon et al. (2007) and McMaster (2008), working with early grades (Kindergarten and grade 1) found that the CCP learning did not significantly improve ELs' oral reading fluency as compared with the ELs in the control condition.

Reading comprehension. In four studies, the researchers measured the impact of the CCP learning on ELs' reading comprehension (n=4; Almaguer, 2005; Liu & Wang, 2015; Saenz et al., 2005; Zhang et al., 2013). The authors of two studies adopted reading multiple-choice comprehension questions to measure the impact of the CCP learning on ELs' reading comprehension outcome (n=2; Almaguer, 2005; Saenz et al., 2005). Only in two studies did the researchers provide adequate and accurate data for calculating effect sizes (n=2; Almaguer, 2005; Saenz et al., 2005). Based on Cohen's criterion (1988), I found the CCP learning had a positive impact on ELs' reading comprehension (measured by multiple-choice questions) with effect sizes Cohen's *d* = 1.03, 95% CI [0.37, 1.69] in Almaguer et al. (2005) and Cohen's *d* = 2.67, 95% CI [1.11, 4.22] in Saenz et al. (2005).

Furthermore, in two studies, the researchers used cloze reading assessment to measure ELs' reading comprehension (n=2; Almaguer, 2005; Zhang et al., 2013). The mixed results were found in the impact of the CCP strategy on ELs' reading comprehension as measured by cloze test. Almaguer et al. (2005) found that the CCP activities improved ELs' reading comprehension with small to medium effect size (Cohen's *d* = 0.33, 95% CI [-0.29, 0.96]). Zhang et al. (2013) found there was no statistical difference between the treatment ELs who conducted collaboration reasoning activities and the ELs in the control condition regarding their improvement in reading comprehension as measured by cloze assessment.

In one study, the researchers did not disclose the measures of reading comprehension (n=1; Liu & Wang, 2015) because it involved large-scale public data sets which used plausible value methods to estimate the reading comprehension achievement. However, the negative impact of small-group and pair-work reading instruction on ELs' reading comprehension was identified. They found that for the fourth grade ELs, the more small-group (Cohen's *d* = 0.92, 95% CI [0.29, 0.72]) and pair-work reading instruction (Cohen's *d* = 0.34, 95% CI [0.13, 0.55]) used in the classroom, the lower ELs' reading achievement was. They concluded that independent learning worked best for grade 4 ELs' improvement in reading.

Phoneme awareness, letter-word identification, and word fluency. Researchers of two studies examined ELs' phoneme awareness, letter-word identification, and word fluency (Calhoon et al., 2007; McMaster et al., 2008). Mix results were identified. Calhoon et al. (2007) found that the CCP learning significantly better supported grade 1 ELs' nonsense word fluency (Cohen's *d* = 0.75, 95% CI [0.28, 1.22]) and letter naming fluency (Cohen's *d* = 0.67, 95% CI [0.20, 1.14]) with large effect sizes, while no statistical significance was found in the improvement in ELs' phoneme segmentation as compared with the ELs in the control condition. McMaster et al. (2008) found for Kindergarten ELs, the CCP strategy better supported the development of rapid letter naming (Cohen's *d* = 1.08, 95% CI [0.17, 2.00]), phonemic awareness, including segmentation (Cohen's *d* = 1.00, 95% CI [0.10, 1.92]), and blending (Cohen's *d* = 1, 95% CI [0.12, 1.93]).

Professional development and ELs' reading vocabulary. In one study, the researchers investigated the effects of consultation on improving ELs' reading vocabulary in class-wide peer-tutoring classrooms from Grades 1-5 (Greenwood et al., 2001). They found that when the

teachers were provided with the ongoing individualized consultation, more students in their classes moved up in level of success (Cohen's d = 6.32, 95% CI [0.21, 1.34])

Synthesis and Discussion of the Findings

Exposure to the CCP activities has developed ELs' reading comprehension and oral reading fluency in varied ways. To understand the impact of the CCP strategy on ELs and the extent of opportunities provided to ELs to improve their English reading and speaking proficiency, researchers need to examine research design, treatment characteristics, and outcome variables of previous empirical studies. Specifically, the research questions guiding this research synthesis were:

1. How were ELs' reading and oral proficiency developed in the empirical studies with the CCP strategy in kindergarten through sixth-grade settings varying in sample characteristics, intervention features, design characteristics, outcome characteristics?
2. What was the impact of the CCP strategy on ELs' reading and oral proficiency?

To answer these two research questions, I conducted a multi-layered analysis to identify which relevant studies which met the inclusive criterion. Further, an investigation was conducted to examine the included studies' characteristics of sample, design, outcome, and the level of effectiveness. In this research synthesis, I placed a particular emphasis on elementary school ELs' English reading and oral development. The findings across randomized control trial, quasi-experimental design, and single-subject design research studies provided the first summary of English reading and oral proficiency improvement for elementary ELs engaged in CCP learning in the past 17 years.

Intervention Features and Design Characteristics of Included Studies

Intervention design, sample, and duration. Five included studies (Almaguer et al., 2005; Calhoon et al., 2007; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013) with RCT or quasi-experimental design were considered to have high methodological quality. For the studies of quasi-experimental design (i.e., Almaguer et al., 2005; McMaster et al., 2008; Zhang et al., 2013), the investigators either examined the baseline equivalence or including students pretesting achievement as covariance or both. The researchers of these five studies reported the comparison of improvement in achievement between the treatment and the control conditions; among these, the researchers of four studies reported demographic information and instructional components in both treatment and control conditions. The researchers of five intervention studies (i.e., Almaguer et al., 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005) reported the total intervention time exceeded 20 hours, which was more than the recommended dosage of CCP learning suggested by Rohrbeck, Ginsberg-Block, Fantuzzo, and Miller (2003). Further, ELs with different language backgrounds benefited from the CCP strategy on one or more aspects of English reading and oral language proficiency.

FOI. Measuring fidelity of implementation was an essential design feature. In the previous reviews, only Pyle et al. (2017) and Bowman-Perrott et al. (2016) examined fidelity of implementation in their syntheses. The findings in this synthesis were consistent with theirs that in the most studies FOI was above 90%, except for Greenwood et al. (2001) with FOI at or above 80%. In this review, Zhang et al. (2013) was the only study without reporting FOI of the intervention. It was also observed in the included intervention studies, that observation checklist was the most common instrument for examining FOI (i.e., Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005). Among these four studies, the observation

checklist was the FOI measure to examine both teachers' and students' implementation of this strategy. Only in one study, the researchers used student attendance records as FOI measure (i.e., Almaguer, 2005).

Measurements. In previous CCP syntheses on ELs' achievement (i.e., Bowman-Perrott et al., 2016; Cole, 2013, 2014; Pyle et al.2017), only Pyle et al. (2017) and Cole (2014) examined the types of measures of each study. In this review, it was observed that curriculum-based measures, standardized measures, and published researcher-created measures were adopted in the included studies to test ELs' reading comprehension, oral reading fluency, oral language/expression, and letter-word identification. I found that standardized measure was the most frequently adopted measure, which accounted for half of the included studies. This finding was consistent with the synthesis conducted by Pyle et al. (2017). Cole (2013) found in his meta-analysis review that the standardized measures were less likely to be adopted in the high-quality studies, but the findings in this review did not support his argument. Moreover, large effect sizes were generated from both researcher-created and standardized measures. Therefore, which type of measures are more sensitive to examine the impact of the CCP strategy on ELs' English language proficiency still needs to be investigated.

Professional development. Professional development was provided via face-to-face in all included the intervention studies. No technology-integrated component or online session was involved in PD of included studies. No researchers examined and analyzed the professional development component in their syntheses (i.e., Bowman-Perrott et al., 2016; Cole, 2013, 2014; Pyle et al.2017). Out of six intervention studies, the researchers provided professional development to the treatment teachers in five studies (i.e., Almaguer, 2005; Calhoon et al., 2007; Greenwood et al., 2001; McMaster et al., 2008; Saenz et al., 2005). However, only in two

studies, treatment teachers were supported with ongoing PD (i.e., Calhoon et al., 2007; Greenwood et al., 2001). As pointed out by Lonigan, Farver, Phillips, and Clancy-Mechetti (2011) many well-designed and intensive PD programs did not show a significant impact on improving students' achievement. It has been emphasized that more PD sessions should be required and should be offered in conjunction with coaching (Hamre, Partee, & Mulcahy, 2017; Piasta et al., 2017). The findings in this synthesis confirmed that there was a positive impact of individualized, ongoing coaching on ELs' reading vocabulary learning (e.g., Greenwood et al., 2001).

Study Outcomes and Effect Sizes

Reading comprehension. I found that there was a consistently significant positive impact of the CCP strategy on ELs' English reading comprehension as measured by reading multiple-choice comprehension questions with large effect sizes (Cohen's *d*s > 0.9). This finding was consistent with the Pyle et al. (2017) conclusion that the CCP strategy supported ELs' development in reading comprehension. However, the mixed results were found in the included studies where cloze reading assessment was used to measure ELs' reading comprehension. For example, Calhoon et al. (2007) found that there was a positive impact of the CCP strategy with small effect size and Zhang et al. (2013) found there was no significant impact of the CCP strategy on ELs' reading comprehension. This discrepancy may be because there were more ask/answer activities involved in the CCP learning, which promoted ELs' reading proficiency of understanding and answering questions. Reading comprehension question, therefore, seems to be a more sensitive or stable measure to examine the impact of the CCP strategy on improving ELs' English reading comprehension. We encourage more researchers to explore the effect of the CCP

strategy with ELs on reading comprehension proficiency measured by reading multiple-choice comprehension questions and cloze reading assessment.

It is worth noting that if there was no intervention involved or no professional development provided to ensure the quality of the CCP learning, no impact or even negative impact can be found by merely pairing or putting ELs work together (i.e., Zhang et al., 2013; Liu & Wang, 2015). The results of this synthesis suggested that the intervention duration is a critical factor for the effectiveness of the CCP strategy on ELs' reading comprehension. According to Rohrbeck et al. (2003), students received less than 19 hours of CCP learning through the intervention were expected to see no difference of improvement as compared with control group. The findings in this review supported the argument that there was no impact of collaborative reasoning as identified in a study by Zhang et al. (2013) that only had 2.67 hours' intervention. Furthermore, I found that grade level was an important indicator. The CCP strategy worked significantly better with the ELs in grades K-3. However, I found no impact or negative impact of the CCP strategy on ELs' reading proficiency in grade 4. Analysis based on large-scale public datasets also indicated that grade 4 ELs might benefit more from independent learning (Liu & Wang, 2015). This finding was consistent with the previous findings that the CCP strategy may work more effectively with younger students (Rohrbeck et al., 2003). The results also extended the findings in Cole's (2014) review that within the elementary school, the CCP learning works better with lower graders (i.e., Grades k – 3).

Phoneme awareness and letter-word identification. Among the seven studies I included, there was a significantly positive impact of the CCP strategy on the improvement of ELs' phonemic awareness and letter-word identification across grade levels with medium to large effect sizes. Since only two studies were identified to examine the impact of the CCP

strategy on ELs' phonemic awareness and letter-word identification, it would be beneficial if future researchers further investigate the area.

Oral reading fluency. I determined mixed results for oral reading fluency among seven included studies. Almaguer (2005) found the CCP strategy significantly improve ELs' oral reading fluency with large effect size, while Calhoon et al. (2007) and McMaster et al. (2008) found there was no significant positive impact on improving ELs' oral reading fluency. Two reasons might explain this discrepancy. First, in the studies conducted by McMaster et al. (2008) and Calhoon et al. (2007), both treatment and control students received the same amount of instructional time while in the study conducted by Almaguer (2005), intervention time was extra English literacy instructional time. The students in Almaguer's (2005) study received more time for English literacy instruction. Third, both the researcher-created and standardized measures were adopted in the included studies. It seems like researcher-created measures are more sensitive to the effect of the CCP strategy. Although Cole (2014) claimed that researcher-created measures were more likely used by high-quality studies measuring the effect of the intervention, I suggested that future researchers investigate which type of measures can best examine the effects of the CCP strategy with ELs' oral reading fluency.

Oral language and oral expression. The CCP strategy was positively identified as supporting ELs' development of English oral language/expression in only one of the seven studies (i.e., Zhang et al., 2013). The CCP strategy positively supported students' development of English coherent narratives for both ELs and mainstreamed students. No difference was identified between the ELs and the mainstream student who received the same CCP intervention, which indicated that the CCP learning supported both language groups. Since there was only one

study which focused on the effect of the CCP strategy on ELs' Englis oral language development, more researchers should explore the effect of this strategy on ELs.

Professional development. Among the seven studies, we found that PD was delivered face-to-face prior to the interventions. The lack of information reported on ongoing PD, requires additional attention. In one study (i.e., Greenwood et al., 2001), I found that individualized, ongoing coaching positively impacts teachers' abilities to influence ELs' reading vocabulary learning. Furthermore, on the basis of the evidence of this research synthesis, it seems fair to suggest that if not intervention of PD was provided to ensure the quality of cooperative/collaborative learning, no impact or even negative impact might be found by merely pairing or putting ELs work together.

Limitations, Implications, and Future Research

While in this research synthesis, I provided a base for future researchers by applying the CCP strategy with ELs; there are several limitations. First, the results of this synthesis and analysis were drawn from the available information as reported by the authors across the included studies. Therefore, the levels of details provided in each study varied, which may have led to some bias in the studies with less information included regarding their intervention design or methodological characteristics. Secondly, the sample sizes in the included intervention studies were relatively small. A finding of no difference or mixed results did not necessarily mean that no difference existed. In the future, researchers can involve a large sample size in the well-designed experimental or quasi-experimental to continue to evaluate the effectiveness of the CCP strategy on ELs' English language proficiency.

There are several implications regarding the design of the future studies. First of all, elementary ELs can benefit from the CCP strategy because it can be embedded into the

curriculum teachers have already used, therefore providing more opportunities for ELs to respond and engage in academic content learning via reading, speaking, listening, and writing (Bowman-Perrott et al., 2016). Secondly, the duration of the intervention focusing on the CCP strategy should be longer, which is also supported by the previous review (Rohrbeck et al., 2003). Thirdly, it is also suggested for future researchers to examine the effective total duration of the CCP intervention. It would be even more beneficial to the field if researchers report the intensity (how many minutes a day), frequency (how many days a week), and duration (how many weeks of implementation) or intervention, which was also supported by the previous review (i.e., Pyle et al., 2017). Fourthly, I agree with the previous reviews (i.e., Browman-Perrott et al., 2016; Pyle et al., 2017) that future researchers should focus on the quality of the CCP learning, which can be a more refined index than simply the description of intensity, frequency, and duration. It should be the exact minutes/frequency that the teachers and students participate in the actual CCP learning activities. Lastly, it is beneficial to involve ongoing PD/coaching and FOI monitoring process during the invention to ensure the quality of implementation. It is an area for the future researcher to conduct more studies that examine the effect of how ongoing PD/coaching and FOI can impact the effectiveness of the CCP strategy for elementary ELs. This is particularly needed in ELs' reading and oral language proficiency at early grade levels because it is a crucial moment for ELs' to develop oral communication for later reading comprehension (Kieffer, 2012) and for reading to learn content knowledge in higher grade levels (Tong, Irby, Lara-Alecio, & Koch, 2014).

CHAPTER III

GRADE 1 BILINGUAL TEACHERS' PEDAGOGICAL DIFFERENCES IN IMPLEMENTING COOPERATIVE/COLLABORATIVE/PEER-TUTORING STRATEGIES DURING CALP INSTRUCTION IN A RANDOMIZED CONTROL TRIAL STUDY

Introduction

Classroom observation assists researchers and practitioners in quantifying teachers' pedagogy (Foorman, Flecher, & Francis, 2004; Foorman & Schatschneider, 2003; Gersten & Baker, 2000; Haager et al., 2003; Irby et al., 2007; Lara-Alecio et al., 2009; Saunders & Fisher, 2006; Tong, Luo, et al., 2017) and reveals instructional quality (Ramirez et al., 1991). Researchers have confirmed that quality curriculum-based PD leads to the improvement in teachers' pedagogy of ELs (Lara-Alecio et al., 2009; Lee & Buxton, 2013; Tong, Luo, et al., 2017). Moreover, researchers also have suggested that the Cooperative/Collaborative/Peer-tutoring (CCP) strategies effectively improves ELs' English proficiency (Bowman-Perrott et al., 2016; Cole, 2013, 2014). The purpose of this study was to investigate the differences in teachers' application of the CCP strategy during different cognitive levels of language content between two conditions in a randomized control study: (a) treatment condition in which teachers were provided with biweekly, intensive, and ongoing curriculum-based VPD, and (b) control condition in which teachers participated in typical district-aligned PD. I intended to answer the over-arching research question – What were the differences in time allocation in the different cognitive levels of language content of the treatment and the control teachers when the teachers were employing the CCP strategy? Because instructional time allocated to teaching cognitive and academic language (CALP) is indicative of the quality of instruction in EL education (Tong, Luo, et al., 2017), I further investigated the pedagogical differences (i.e., teachers' instructional

activity and language and students' communication mode and language) between the treatment and control conditions when CCP strategy was employed during light cognitive and dense cognitive content as defined by TBOP. I intended to answer the following over-arching research question – What were the pedagogical differences (i.e., teachers' instructional activity and language and students' communication mode and language) between the treatment and the control condition during light cognitive and dense cognitive content instruction time when they implemented the CCP strategy as documented by the classroom observation instrument TBOP?

In this study, I conducted a multi-layer cross-tabulation, a multi-layer chi-square of independence, and a nonparametric chi-square analysis on archived data, which came from a large-scale research project – English Language and Literacy Acquisition -Validation (ELLA-V, U411B120047). It was a five-year, large-scale validation study with a focus on investigating the effectiveness of the structured English immersion and transitional bilingual education models on improving Spanish-speaking ELs' English language and literacy acquisition from kindergarten to Grade 3. The purpose of Project ELLA-V was to validate the instructional components of a previous four-year (K-3) Project ELLA (R305P030032). The majority of the student population in this dissertation study were identified by Texas state criteria as Limited English Proficient, all of whom were identified at the time of admission with Spanish as their first language (L1). This five-year RCT sought to validate the practice that was most effective in helping native Spanish-speaking ELs in acquiring English language proficiency and literacy skills. This dissertation study was focused on pedagogical practice when the CCP strategy was used in Grade 1 EL classrooms.

Classroom activities were virtually recorded and coded via a classroom observation instrument, TBOP, for three rounds each school year from Kindergarten to Grade 3. This

dissertation focused on the first-grade teachers' pedagogical practices in EL classrooms. Since the archived observation data was categorical, a three-layer (condition, the domain of ESL strategies, and the domain of language content) cross tabulation analysis and a chi-squared test of independence was used to determine the pedagogical differences between the treatment and the control conditions when teachers were using the CCP strategy. Since I was more interested in the pedagogical differences when teachers were using this specific strategy during the dense and light cognitive content instructional time, I conducted three-layer cross tabulation analyses and three-layer chi-square tests of independence to investigate the differences in teachers' instructional activities and languages as well as students' language mode and responding languages when the CCP strategy was employed. In order to examine the effect size of the differences between treatment and control condition regarding teachers' and students' practices in the CCP learning activities during light and dense cognitive content, nonparametric chi-square analysis was adopted.

Literature Review

Theories Supported ELs' The CCP Learning from Socio-cognitive and Sociocultural Perspectives

Vygotsky (1978) suggested that when students work in a collaborative situation, they tend to respond at a higher academic level, as compared to being asked to work individually. According to Totten, Sills, Digby, and Russ (1991), collaborative learning provided students with an opportunity to participate in a discussion, taking charge of their education and thereby becoming critical thinkers. Bruner (1985) specified that the cooperative learning strategies could improve students' problem-solving abilities because they have to internalize external knowledge and critical thinking skills and at the same time interpret internalized knowledge to their peers.

Teachers can also benefit from using the CCP strategy to help less capable learners learn from their peers (Almaguer, 2005)

ELs whose first language is not English, need to be provided sufficient opportunities to develop their language proficiency and content area knowledge. To succeed in school, ELs need various opportunities to explore and practice the language. Moreover, educators have a responsibility to address the needs of ELs and provide opportunities to help ELs improve English language proficiency. Therefore, if ELs develop their English language proficiency through social interactions, the involvement of ELs in collaborative learning is beneficial for their academic achievement (Krashen, 2004). Slavin (1988) stressed that collaborative learning strategy should be a major component of instruction for ELs because it provides a rich academic context for language learning. Using a collaborative learning strategy, teachers can increase the amount of instructional time for ELs and increase the opportunities for ELs to use English in the classrooms (Enright & McCloskey, 1988). Furthermore, researchers have found that students' self-esteem, intergroup relations, and ability to cooperate with their peers were significantly improved via a process of collaborative learning (Slavin, 1991).

Defining Quality of Instruction from Pedagogical Perspective

Researchers found that quality instruction has been associated with ELs' achievement (Tong et al., 2010, 2014; Slavin, Madden, Calderón, Chamberlain, & Hennessy, 2011; Burchinal, Field, López, Howes, & Pianta 2012). Pianta, Belsky, Vandergrift, Houts, and Morrison (2008) pointed out that the student achievement gap has been related to an inequality of exposure to learning opportunities, and in majority elementary classrooms including ELL classrooms there was a low level of instruction quality for students' cognitive growth. However, due to the lack of a cohesive body of instructional research with ELs (Francis, Rivera, Lesaux, Kieffer, & Rivera,

2006), the definition of quality of instruction varies (Tong, Luo, et al., 2017). Tong, Luo, et al. (2017) pointed out that the quality of instruction should not be equated with the quantity of instruction. Instead, they supported that instructional time allocated in CALP (Cummins, 1986) can be used as an indicator of the quality of instruction (Tong, Luo, et al., 2017) because they considered that minutes spent on the level of instructional content matters. They argued that for ELs, the exposure to learning opportunities where teachers intentionally deliver CALP content to develop ELs' cognitive and academic language proficiency in English was central to advance ELs' academic achievement.

Cummins' (1986) differentiation of CALP and BICS has been challenged by researchers for an overgeneralization of complexity of language content in ELs classrooms (Edelsky, 1996; Trueba, 1989; Wiley, 1996). Lara-Alecio and Parker (1994) proposed that CALP can be further defined as two layers—dense cognitive and light cognitive content—due to the existence of "a continuum between BICS and CALP" (p.122). In their pedagogical model for transitional English bilingual classrooms, Lara-Alecio and Parker reclassified Cummins' (1986) CALP and BICS theory as (a) BICS that subsumed level 1 social routines and level 2 academic routines, and (b) CALP, which included level 3 light cognitive content and level 4 dense cognitive content. How teachers allocate instructional time in ELs classroom is a crucial determinant of quality of instruction (Irby et al., 2007; Lara-Alecio et al., 2009). Therefore, it would be more meaningful to examine teachers' instructional practice that occurs during higher quality instructional time—light cognitive and dense cognitive content under CALP.

Classroom Observation

Classroom observation, an objective and reliable method to document occurrences in the classrooms (Lee et al., 2009; Noell, 2010), has gradually contributed to quantify teachers'

pedagogical practice in ELL classrooms (Foorman et al., 2004; Foorman & Schatschneider, 2003; Gersten & Baker, 2000; Haager et al., 2003; Irby et al., 2007; Lara-Alecio et al., 2009; Saunders & Fisher, 2006; Tong, Luo, et al., 2017). However, there are many challenges to address. For example, using video to obtain a record of the CCP learning activities usually involves multiple observers (Borko et al., 2008) and analyzing the recording also requires multiple collaborative coders (Frederiksen, Sipusic, Sherin, & Wolfe, 1998; Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). Furthermore, to analyze the recordings for research purpose, it is crucial to address how to use the proper observation protocol to capture the essential pieces in the ESL context as well as how to accurately interpret what is recorded in the video (Erickson, 1986; Lara-Alecio et al., 2009; Sherin & Han, 2004). Video observation can generate rich data to capture the teacher-student and student-student interaction, which can help researchers and practitioner to assess the classroom instruction as well as help them examine teachers learning as the results of professional development (Stigler et al., 2000). Recent work that focused on the impact of PD on ELs' improvement in English language proficiency used classroom observation to record the teachers' instructional practice and analyze its mediating effect on ELs (e.g., Tong, Luo, et al., 2017), which enriched knowledge of creating descriptive and explanatory accounts of ELs' language learning improvement.

The CCP Strategy, Quality of Instruction, and Classroom Observation

Researchers have found that collaborative/collaborative strategies (Cole, 2013, 2014; Webb, 1995) and quality of instruction (Tong, Luo, et al., 2017) have positive effects on ELs' learning. Researchers have suggested that teachers should provide explicit instruction (Webb, 1995) and feedback (Tudge, 1992) to ELs during the CCP activities and at the same time monitor and evaluate the interaction among the students (Dillenbourg, 1999). Therefore, comprehensive

analysis of the dynamics of the CCP activities might bring insights on the quality CCP learning. However, researchers pointed out there was a literature gap in how to define and evaluate quality CCP learning in an EL classroom (Bowman-Perrott et al., 2016; Pyle et al., 2017). As I mentioned earlier in this study, classroom observation can provide a direct and reliable description of occurrences in the classroom, which would unveil the complexity of EL classrooms when the CCP strategy is employed.

From the perspectives of socio-cognitive, sociocultural, and pedagogical theories, CCP learning supports ELs' development in language proficiency, content subjects, and self-esteem. It would be more beneficial to encourage teachers to involve EL in quality CCP learning activities that contain higher-order and critical thinking. Therefore, I argue that quality CCP learning can improve ELs' language improvement, which means the time allocation of the CCP strategy in the light and dense cognitive content matters in ELs' learning. Furthermore, by analyzing instructional occurrences during the light and dense cognitive content instruction, we can have a better understanding of the dynamics of quality CCP learning in EL classrooms.

Specifically, two research questions guided this study:

1. What were the differences of time allocation in the CCP strategy between the teachers in the treatment (with the support of the curriculum-based VPD) and the control conditions (with the support of district-aligned PD) when the teachers were teaching different levels of language content?

2. What were the pedagogical differences (i.e., teachers' instructional activities and languages, and students' language modes and responding languages) between the treatment teachers and the control teachers during the dense and light cognitive content instruction time

when they employed the CCP strategy as documented by classroom observation instrument TBOP?

Method

In this study, I analyzed archived data from Project ELLA-V 2015-2016 with the permission of the project's principal investigators. Three rounds of classroom observation of Grade 1 ESL instruction were recorded virtually and coded via Transitional Bilingual Observation Protocol (TBOP). I examined the pedagogical practice of EL classroom instruction when the teachers used the CCP strategy. Since ELLA-V was a randomized control trial (RCT) study, I was able to examine the pedagogical differences between the treatment and the control teachers when they used the CCP strategy.

In Project ELLA-V, schools were closely matched based on their overall ELs' English language proficiency rated by Texas English Language Proficiency Assessment System (TELPAS), the percent of economically disadvantaged students, and the percent of ELs before randomization. Three schools were clustered and were randomly assigned to Treatment 1 (T1), Treatment 2 (T2) and the control conditions. Teachers in all conditions were observed three times: beginning, middle, and end of the school year.

In the present study, the archived data from the database of teachers' pedagogical practices of Project ELLA-V were analyzed. Specifically, I focused on first-grade teachers' practices in T2 and the control conditions. The study design is illustrated in Table 2, which includes the sequence of treatment and measurement occasions for two groups in the present study (i.e., T2 and the control condition). Since archived data illustrated the frequency of pedagogical occurrences, multi-layer frequency cross tabulation analyses and multi-layer chi-squared tests of independence were adopted to determine the differences between T2 and the

control condition regarding teachers' time allocation of the CCP strategy in the dense and light cognitive content, as well as the difference in teachers' instructional language and students' language respectively when the CCP strategy was employed during the dense and light cognitive content.

Table 2
Research Design

	Kindergarten			Grade 1			Grade 2			Grade 3		
	Begin	Mid	End	Begin	Mid	End	Begin	Mid	End	Begin	Mid	End
T2				Ob1	Ob2	Ob3						
C				Ob1	Ob2	Ob3						

Begin=Beginning of the school year, Mid=Middle of the school year, End=End of the school year, Ob1=Observation 1, Ob2=Observation 2, Ob3=Observation 3.

Participants

In this study, a subsample from the T2 and control conditions in the first grade of project ELLA-V was adopted. All selected teachers were first-grade teachers in the schools located in Texas. There are 77 first-grade teachers from 41 schools participated in the project, 38 from T2, and 39 from the control condition in the first grade. In Table 3, the number of the first-grade teachers who were observed in three rounds was summarized.

Table 3
Number of Teachers in the First Grade

	Round		
	1	2	3
T2	38	38	38
Control	39	39	39

All teachers were recorded at the beginning, middle, and end of the school year during ESL block, with a total of 231 videos, 114 in Treatment 2 and 117 in control condition. The average length of observation was 49 minutes for each T2 teacher and 48 minutes for each control teacher. Their recorded observation was coded via TBOP to analyze their pedagogical practices during the ESL block. The detailed coding procedure is introduced in the data collection section. In all, 13,860 20-second video clips were coded: 6,840 clips in T2 and 7,020 clips in the control condition.

Intervention: Student Level

Treatment. The instruction included two blocks: a 35-minutes Storytelling, retelling and Higher Order Thinking for English Language and Literacy Acquisition (STELLA) and a 10-minutes Academic Oral and Written Language in Science (AOWLS). STELLA curriculum aimed to facilitate interactive story reading between teacher and students, develop ELs' reading comprehension via using leveled questioning strategy, and provide ELs with meaningful opportunities to participate and respond in English language and literacy activities and expand their vocabulary and their listening and speaking skill. In AOWLS, curriculum targeted at developing ELs' academic vocabulary development and science concepts by providing sentences using words in context and asking a daily question using the target word. Students discussed with a partner and wrote a response in their journal.

Control condition. Instruction included typical District ESL Instruction which includes 45 minutes of instruction daily. The control condition followed district standards-based guidelines and utilized available resources, which varied from district to district and campus to campus. The teachers in the control condition provided various subjects including language arts, reading in English, or content-area instruction (e.g., math or science). ELs' English learning was

scaffolded via visual aids, anchor charts, and peer-assisted learning activities. The instruction was delivered in full classes or small groups.

Intervention: Teacher Level

Treatment. The treatment teachers received virtual professional development (VPD), a biweekly, 90-minute, curriculum-based VPD via Citrix GoToTraining in which participants could chat. During VPD sessions, the coordinators and the treatment teachers reflected on and discussed student learning and progress and previewed upcoming lessons and materials. L2 acquisition strategies, ESL instruction strategies, and video clips of instruction modeling were also provided to treatment teachers. The treatment teachers were also supported by building networks of practice and conducting self-assessment via virtual professional portfolio. The treatment teachers that were observed with a low rating of fidelity of implementation were provided one-on-one, live virtual coaching.

Control condition. Teachers received state aligned PD provided by school districts. Resources and forms varied from district to district.

Instrumentation

In this dissertation study, TBOP was adopted to rate the frequency of teachers' pedagogical practice. TBOP is an observation analysis tool development from Four-dimensional Transitional Bilingual Pedagogical Theory (Lara-Alecio & Parker, 1994) that was designed to observe: a) occurrences of teachers' practices in four major instructional domains: language of instruction, language of content, communication mode, and activity structure and two minor domains: ESL strategies and group, b) interactions among domains, and c) fidelity of implementation of the intervention at teacher level over time (Figure 2).

TBOP is a multidimensional classroom observation instrument that includes four major dimensions and two minor domains. The domain of language content derived from Cummins' (1986) BICS and CALP concepts since these two tiers distinction is not distinct enough to describe the actual occurrence of instruction content observed in the classroom (Lara-Alecio & Parker, 1994). In TBOP, language content was reformulated into four levels: (a) social routines (i.e., social exchanges), (b) academic routines (i.e., returning books, structuring homework, or handing in assignments,), (c) light cognitive content (i.e., reviewing content, repetitive drills, or skills practice),and (d) dense cognitive content (i.e., critical thinking or new content-area information). The second domain of TBOP is language of instruction. Krashen (1985) pointed out that a content area can provide a rich source for ELs' English language learning. Therefore, Lara-Alecio and Parker (1994) proposed four categories in this domain: (a) content presented in L1, which is Spanish in this study, (b) content presented in L2, which is English in this study, (c) L2 clarified by L1, and (d) L1 introducing L2.

The third domain in TBOP is activity structure, defined as a combination of (a) types of teacher practices (e.g., lecturing, observing, evaluating, and asking) and (b) the expectation for student responses (e.g., listening, answering, cooperating, and asking). Activities, such as the transition from one activity to the next activity, unexpected nonacademic interruption, and discipline feedback were categorized as non-academic activities. During the majority of an observation period, if the teacher is asking a question and the students are answering, the practices were coded as ask/answer (ask/ans).

The fourth domain in TBOP is communication mode, derived from the reciprocal interaction model (Cummins, 1986) and context-specific mode (Diaz, Moll, & Mehan, 1986) to specify students' communication modes, such as receptive modes (aural, reading), expressive

modes (verbal, writing), and combinations of these modes. In addition to four main domains of TBOP, two minor domains are also included: ESL strategies (e.g., the CCP strategy, academic language scaffolding strategies, visual scaffolding strategies, leveled questioning strategies, and et al.) and the physical group that the teacher interacts with (e.g., single, pair, small group, large group, and whole class).

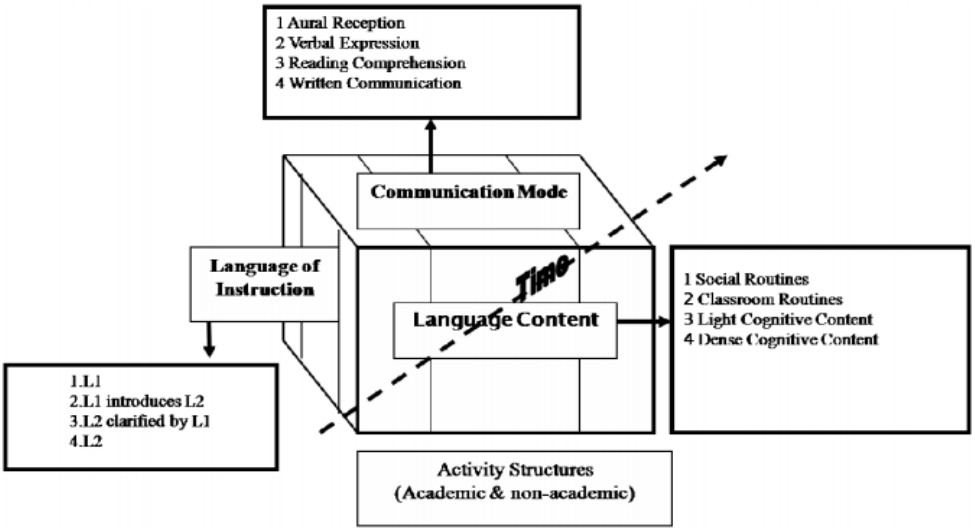

Figure 2. The four-dimensional transitional bilingual pedagogical theory model (reprinted from Lara-Alecio & Parker, 1994).

Data Collection

At the beginning of Project ELLA-V, IRB approval was given for all activities relevant to the purposes of the research project. Consent forms in both Spanish and English were provided to parents to sign. The teachers and the principals also consented to participate in the project.

The first-grade teachers in the treatment and the control conditions were observed and recorded virtually in a secure observation lab at the beginning, middle, and end of the 2015–2016 school year via pre-installed observation cameras in each classroom. The recorded lessons were rated by trained research personnel via an online web tool embedded with TBOP for raters to code teachers' pedagogical practice. For the first five-minutes, the raters coded instruction and every other five-minute-section (third, fifth and seventh); raters coded for a total of 20 coded minutes of instruction. For these four five-minute intervals, raters coded 20-second interval clips and found 15 codes for every five minutes. Therefore, one single teacher received 60 codes for one round of observation.

Fidelity of Implementation

The participating teachers were observed three times a year during the beginning, middle, and end of the intervention for fidelity. In total, 116 videos were collected and evaluated by the Teacher Observation Record (TOR) across conditions for the FOI purpose. TOR was developed by Tong, Lara-Alecio, Irby, Mathes, and Kwok (2008) and modified by Tong, Irby et al. (2017) to fit Grade 1 curriculum. It measured the following nine areas: (a) teacher preparation, (b) material usage and preparation, (c) lesson pacing, (d) student involvement, (e) teacher talking, reading, writing time vs. student talking, reading, writing time, (f) use of leveled questioning, (g) use of ESL strategies, (h) effective and cognitive feedback, and (i) physical environment. Inter-rater reliability was established and monitored through the FOI measuring process, reaching above substantial level (Landis & Koch, 1977). On average, FOI was 0.94 for Treatment 2 and 0.84 for the control condition.

Data Analysis

Data were first analyzed using a three-layer cross-tabulation involving condition, the domain of language content and the domain of ESL strategy to summarize and compare teachers' time allocation of the CCP strategy between conditions. Since I was particularly interested in the pedagogical differences between the treatment and the control condition when the teachers were using the CCP strategy during the light and dense cognitive content instruction, five, three-layer cross-tabulation analyses were adopted. First, a three-layer cross-tabulation involved three layers of the condition, the domain of language content (focusing on the light and dense cognitive content), and the domain of activity structure. The second three-layer cross-tabulation involved three layers of the condition, the domain of language content (focusing on the light and dense cognitive content), and the domain of mode. The third and fourth three-layer cross-tabulations involved three layers of the condition, the domain of language content (focusing on the light and dense cognitive content), and the domain of instructional languages (two subdomains of teacher languages and student languages). Because data was frequency count, multi-layer cross tabulation analyses, multi-layer chi-square tests of independence, and nonparametric chi-square tests were used for data description and analysis. Cohen's *d*s were used for reporting the effect sizes. SPSS 24 and online effect size calculator (Wilson, n.d.) were adopted in this study.

Results

The 13,860 20-second video observations equated to an amount of 4,620 minutes or 77 hours of coded instruction time for the teachers in both treatment and control conditions. The treatment teachers were observed 2,280 minutes or 38 hours, while the teachers in the control condition were observed 2,340 minutes or 39 hours. Although the teachers in both conditions were observed for a similar amount of time to code the pedagogical practices, the teachers in the

treatment condition were expected to differ from the teachers in the control condition. A three-layer cross tabulation analysis was conducted to observe the differences between the teachers in the treatment and the control conditions regarding their time allocation in the CCP strategy (the domain of ESL strategy) in different levels of language content (the domain of language content). A three-layer chi-square test of independence was conducted among the domain of ESL strategy, the domain of language content, and condition. There was a statistically significant association among the domain of language content, condition, and the CCP strategy under the domain of ESL strategy, χ^2 (3) = 58.082, $p < 0.001$, Cramer's V = .224.

From Table 4, the results illustrated a difference between teachers in the treatment and the control conditions regarding teachers' time allocation in applying the CCP strategy in the domain of language content. In fact, the results of nonparametric chi-square analysis indicated that the treatment teachers used this strategy significantly more frequently than the teachers in the control condition in the sub-domains, the light cognitive content ($p<0.001$, Cohen's $d = 1.0$) and the dense cognitive content ($p<0.001$, Cohen's $d = 1.03$). To identify what languages ELs and teachers used with different pedagogies, what instructional activities teachers conducted and what language modes ELs were in when teachers were using the CCP strategy during the light and dense cognitive content instruction time respectively, eight, three-layer chi-square analyses were conducted.

Table 4

Description Statistics of Teachers' Time Allocation in the CCP Strategy by Language Content

		Social	Academic	Light	Dense	Total
Treatment	Count	1	23	120	670	814
	Percentage	0.1%	2.5%	13.0%	72.8%	70.1%
Control	Count	5	47	46	250	348
	Percentage	0.5%	5.1%	5.0%	27.2%	29.9%
Total	Count	6	70	166	920	1162
	Percentage	0.5%	6.0%	14.3%	79.2%	100.0%

I further examined teachers' instructional activities when they were applying The CCP strategy during dense and light cognitive content instruction periods and compared the difference between teachers in the treatment and the control conditions via a three-layer cross tabulation analysis involving the domain of language content, condition and the domain of activity structure. A three-layer chi-square of independence was conducted among the domain of activity structure, the domain of language content, and condition. There was no statistically significant association among the domain of activity structure, condition, and the light cognitive content under the domain of language content, χ^2 (11) = 13.804. There was a statistically significant association among the domain of activity structure, condition, and the dense cognitive content under the domain of language, χ^2 (12) = 25.662, $p < 0.01$, Cramer's V = .17. The selected categories of instruction activities listed above were the ones showing a statistically significant difference between the treatment and the control conditions. The results were reported in Tables 5 and 6. During the light cognitive content instructional time, the treatment teachers differed from teachers in the control condition in three instructional activities: direct/listen, evaluate/performance, and observe/performance. Moreover, the results of nonparametric chi-square indicated that the treatment teachers allocated significantly more instructional time to

these three activities (*p*s < .05, Cohen's *d*s > .9). However, during dense cognitive content instructional time, the treatment teachers allocated significantly more time to direct/performance ($p < .001$, Cohen's $d = 1.7$), ask/answer ($p<.001$, Cohen's $d = 2.07$), evaluate/performance ($p = .013$, Cohen's $d = .78$), evaluate/cooperation ($p<.001$, Cohen's $d = .79$), and observe/cooperation ($p<.001$, Cohen's $d = 1.11$).

Table 5
Description Statistics of Teachers' Time Allocation in Activity Structure by the CCP Strategy and Light Cognitive Content

		dir/per	ev/per	obs/per
Treatment	Count	21	22	34
	Percentage	12.7%	13.3%	20.5%
Control	Count	6	9	9
	Percentage	3.6%	5.4%	5.4%

Table 6
Description Statistics of Teachers' Time Allocation in Activity Structure by the CCP Strategy and Dense Cognitive Content

		dir/per	ask/ans	ev/per	ev/cop	obs/cop
Treatment	Count	89	55	32	323	98
	Percentage	9.7%	6.0%	3.5%	35.1%	10.7%
Control	Count	19	9	15	149	34
	Percentage	2.1%	1.0%	1.6%	16.2%	3.7%

I examined students' language mode when the teachers were using the CCP strategy during the light and dense cognitive content instructional period via a three-layer cross tabulation analysis involving the domain of language content, condition and the domain of communication mode. A three-layer chi-square of independence was conducted among the domain of communication mode, the domain of language content, and condition. There was a statistically

significant association among the domain of communication mode, condition, and the light cognitive content under the domain of language content, χ^2 (8) = 21.112, $p < 0.01$, Cramer's V = .357. There was a statistically significant association among the domain of communication mode, condition, and the dense cognitive content under the domain of language content, χ^2 (12) = 134.714, $p < 0.001$, Cramer's V = .383. The communication modes listed in Table 7 and 8 were the ones that showed a statistically significant difference between the treatment and the control conditions. In Table 7, there were only two student language modes that differed significantly between the treatment and the control conditions: verbal/aural and aural/verbal. The results of nonparametric chi-square indicated that the treatment students spent significantly more instructional time in speaking/listen ($p < .001$, Cohen's $d = 1.48$) and listening/speaking ($p < .001$, Cohen's $d = 1.53$) during the light cognitive content instruction time when the teachers were using the CCP strategy. In Table 8, the students in both treatment and control conditions spent more time in the communication modes of speaking/listening (49% in treatment; 14.6% in control) and listening/speaking (17.1% in treatment; 3.6% in control), during the dense cognitive content instructional time when the teachers were using the CCP strategy. In the dense cognitive content instructional time, the treatment students differed from the students in the control condition in the following modes: aural, aural/writing, verbal/aural, and aura/verbal. The results of nonparametric chi-square indicated that as compared with the students in the control condition, the students in the treatment condition statistically significantly prefer to communicate with each other in the modes of listening (p =.013, Cohen's $d = 0.97$), speaking/listening ($p < .001$, Cohen's $d = 1.29$) and listening/speaking ($p < .001$, Cohen's $d = 1.72$) when the teachers were using the CCP strategy during the dense cognitive content instructional period. Students in the control condition spent more time in the mode of listening/writing ($p < .01$, Cohen's d =

1.65) as compared with the treatment students. However, the total percentage of time that students in the control condition allocated to the listening/writing mode only counted for 2% out all overall the dense cognitive content instructional time.

Table 7
Description Statistics of ELs' Time Allocation in Communication Mode by the CCP Strategy and Light Cognitive Content

		ver-au	au-ver
Treatment	Count	55	41
	Percentage	33.1%	24.7%
Control	Count	14	10
	Percentage	8.4%	6.0%

Table 8
Description Statistics of ELs' Time Allocation in Communication Mode by the CCP Strategy and Dense Cognitive Content

		aural	au-wr	ver-au	au-ver
Treatment	Count	23	4	451	157
	Percentage	2.5%	0.4%	49.0%	17.1%
Control	Count	9	18	134	33
	Percentage	1.0%	2.0%	14.6%	3.6%

A three-layer cross-tabulation analysis was also conducted to examine the domain of instruction language. To be more specific, I examined what languages the teachers and the students used when the teachers were using the CCP strategy during the light and dense cognitive content instruction. A three-layer chi-square of independence was conducted among the domain of teacher instructional language, the domain of language content, and condition. There was a statistically significant association among the domain of teacher instructional language, condition, and the light cognitive content under the domain of language content, χ^2 (3)

= 9.047, $p < 0.05$, Cramer's V = .233. There was a statistically significant association among the domain of teacher instructional language, condition, and the dense cognitive content under the domain of language content, χ^2 (3) = 8.02, $p < 0.05$, Cramer's V = .093. The results of teachers' languages are reported in Tables 9 and 10, while the results of students' languages are reported in Tables 11 and 12. The results of nonparametric chi-square indicated when the teachers were using the CCP strategy during the light cognitive content instructional time, the treatment teachers were observed to use English more often as compared with the teachers in the control condition ($p < .001$, Cohen's $d = 0.95$). It was also observed that the treatment teachers kept silent more often than the teachers in the control condition ($p < .001$, Cohen's $d = 1.5$). The results of nonparametric chi-square indicated that when the teachers were using this strategy during the dense cognitive content instruction period, the treatment teachers were also observed to use English more often ($p < .01$, Cohen's $d = 1.12$) than the control teachers did. The treatment teachers were also observed to allocate more time to silence when using this strategy during the dense cognitive instruction as compared with the teachers in the control condition ($p < .001$, Cohen's $d = .71$).

Table 9
Description Statistics of Teachers' Time Allocation in Instructional Language by the CCP Strategy and Light Cognitive Content

		L1	L2	L2-L1	NA	Total
Treatment 2	Count	0	88	0	32	120
	Percentage	0.0%	53.0%	0.0%	19.3%	72.3%
Control	Count	1	35	2	8	46
	Percentage	0.6%	21.1%	1.2%	4.8%	27.7%
Total	Count	1	123	2	40	166
	Percentage	0.6%	74.1%	1.2%	24.1%	100.0%

Table 10
Description Statistics of Teachers' Time Allocation in Instructional Language by the CCP Strategy and Dense Cognitive Content

		L1	L2	L1-L2	NA	Total
Treatment 2	Count	2	542	0	126	670
	Percentage	0.2%	58.9%	0.0%	13.7%	72.8%
Control	Count	0	186	1	63	250
	Percentage	0.0%	20.2%	0.1%	6.8%	27.2%
Total	Count	2	728	1	189	920
	Percentage	0.2%	79.1%	0.1%	20.5%	100.0%

A three-layer chi-square of independence was conducted among the domain of student responding language, the domain of language content, and condition. There was a statistically significant association among the domain of student responding language, condition, and the light cognitive content under the domain of language content, χ^2 (3) = 15.183, $p < 0.01$, Cramer's V = .302. There was a statistically significant association among the domain of student responding language, condition, and the dense cognitive content under the domain of language content, χ^2 (3) = 27.212, $p < 0.001$, Cramer's V = .172. The results of students' responding languages are reflected in Tables 11 and 12. The same pattern was observed in students' language when teachers were using the CCP strategy during CALP instructional time. The results of nonparametric chi-square indicated that the treatment students were observed to respond more in English in both light and dense cognitive content instructional time when the teachers were using the CCP strategy (*p*s < .01, Cohen's *d*s > 1.1). It was also observed that there was no statistically significant difference between the students in the treatment and the control conditions regarding their silent time when the teachers were using the CCP strategy during the light and dense cognitive content instructional time.

Table 11
Description Statistics of ELs' Time Allocation in Instructional Language by the CCP Strategy and Light Cognitive Content

		L1	L2	L1-L2	NA	Total
Treatment 2	Count	0	102	1	17	120
	Percentage	0.0%	61.4%	0.6%	10.2%	72.3%
Control	Count	2	28	0	16	46
	Percentage	1.2%	16.9%	0.0%	9.6%	27.7%
Total	Count	2	130	1	33	166
	Percentage	1.2%	78.3%	0.6%	19.9%	100.0%

Table 12
Description Statistics of ELs' Time Allocation in Instructional Language by the CCP Strategy and Dense Cognitive Content

		L1	L2	NA	Total
Treatment 2	Count	4	628	38	670
	Percentage	0.4%	68.3%	4.1%	72.8%
Control	Count	9	208	33	250
	Percentage	1.0%	22.6%	3.6%	27.2%
Total	Count	13	836	71	920
	Percentage	1.4%	90.9%	7.7%	100.0%

Discussion

The quality CCP learning scaffolds ELs' development in language proficiency and academic subject, which was supported by socio-cognitive, sociocultural, and pedagogical theories. The first and most important step is to describe and understand the quality of the CCP learning in EL classroom and to examine the potential value of the ongoing VPD of Project ELLA-V on teachers' ability of providing more CCP learning opportunities to ELs. A multidimensional classroom observation instrument, TBOP, and multi-level cross tabulation analyses were adopted in this study to identify variation in treatment and control EL classrooms

regarding time allocation in language content, language of instruction, communication mode, and activity structures when the CCP strategy was applied.

The treatment teachers were observed to spend more instructional time in the CCP strategy. Moreover, the teachers in the treatment condition were observed to allocate more time to the light and dense cognitive content instruction while the teachers in the control condition were observed to allocate more time to social communication and academic transition. The treatment teachers who were supported by the curriculum-based VPD involved the students more often in the CCP learning activities in the ESL classrooms. In addition, the treatment teachers exposed ELs to more opportunities for the light and dense cognitively challenging content during ESL instruction.

During Project ELLA-V, the classroom observation protocol TBOP was utilized for several purposes, to record teachers' pedagogical practice as well as the interactions among students. Since the treatment teachers received various resources from VPD, including ESL strategy training, scripted lesson plans, curriculum material, and technology support, it was not surprising to see a distinction in pedagogy between the teachers in the treatment and the control conditions. On the other hand, the teachers in the control condition following the state-aligned curriculum were provided state-aligned PD. They were limited to the material that they were provided to shape their pedagogy in the ESL classrooms.

I further examined the pedagogy differences in the major domains (i.e., teachers' activity structure, students' language mode, teachers' instructional languages, and students' responding languages) when the teachers were using the CCP strategy in the light cognitive and dense cognitive content respectively under the domain of language content. Statistically significant differences were found in each pedagogical domain when the teachers were using this strategy

with both light cognitive and dense cognitive content. More pedagogical differences in the domains of activity structure and mode were observed when the teachers were using the CCP strategy in the dense cognitive content than in the light cognitive content. For example, when the teachers employed the CCP strategy, the treatment teachers were observed to allocate more time to three instructional activities: direct/performance, evaluation/performance and observe/performance in light cognitive content. Besides these three activities, during dense cognitive instruction, the treatment teachers allocated more instructional time to the instructional activities, e.g., evaluation/cooperation and ask/answer. This finding was consistent with what I found in the domain of mode. The students in the treatment condition were observed to spend more time in listening/speaking and speaking/listening in both light and dense cognitive content. Moreover, when the CCP strategy was applied in the ESL classrooms, the treatment teachers were observed to spend more than half of instructional time in delivering the light and dense cognitive content in English, while the teachers in the control condition were observed to spend around one-fifth of English instruction in both light and dense cognitive content. The same pattern was observed in students' responding languages. The treatment students were observed to spend more than 60% of the time responding in English to light cognitive content and almost 70% of the time to dense cognitive content when the CCP strategy was applied. On the contrary, the students in the control condition were observed to spend around 20% of the time responding in English to both light and dense cognitive content. According to the results, in EL classrooms, ELs tended to mirror their teachers' instructional language. This finding was also consistent with a previous study (i.e., Lara-Alecio et al., 2009).

Limitation

In this study, I aggregated three rounds of TBOP observations conducted during the beginning, middle, and end of a Grade 1 intervention. Analyzing three rounds of observations separately provides a growth trajectory of how the treatment teachers improved their pedagogy in different domains. Moreover, as observed by the observation instrument TBOP, there were more ESL strategies other than the CCP strategy being used in the EL classrooms. Due to the purpose of this study, I did not explore pedagogical differences when teachers employed other ESL strategies (e.g., academic language scaffolding or visual language scaffolding) which might also be beneficial to ELs.

Conclusion and Implication

Project ELLA-V treatment teachers received a biweekly, intensive, and curriculum-based VPD, which led to differences in teachers' time allocation of the CCP strategy in different levels of cognitive content as observed via the multidimensional classroom observation protocol, TBOP. The evidence was further constructed by examining the domains of instructional language, mode, and instruction activity. The treatment teachers were observed to employ the CCP strategy more often than the teachers in the control condition, especially in a dense cognitive content. The treatment students were provided more opportunities for discussion and CCP learning to practice their English speaking and listening proficiency. At the same time, the treatment teachers evaluated the students' CCP activities and provided feedback. This process is the instructional practice that Tudge (1992) and Dillenbourg (1999) recommended. After observing the findings, I also suggest that teachers in the control condition would have the potential to employ this strategy in quality instruction more often if they were provided with a research-based curriculum and the curriculum-based VPD service. The students in the control

condition would also have been provided more opportunities to improve their English language proficiency. An intervention like Project ELLA-V, which provided teachers a research-based curriculum and a curriculum-based VPD, could scaffold teachers' quality of instruction and application of ESL strategies, which could further bring ELs more opportunities to participate in higher-order thinking and CCP learning activities.

CHAPTER IV

THE IMPACT OF THE COOPERATIVE/COLLABORATIVE/PEER-TUTORING STRATEGIES ON GRADE 1 ENGLISH LEARNERS' READING COMPREHENSION, ORAL EXPRESSION, AND ORAL READING FLUENCY IN AN RCT RANDOMIZED CONTROL STUDY: A MULTI-LEVEL ANALYSIS

Introduction

The percentage of English learners (ELs) in the United States' school-aged population has increased in last decade (National Center for Educational Statics [NCES], 2015). ELs have underperformed in content subjects, including math and science (NCES, 2010, 2011, 2015) due to their limited academic English language proficiency. The educational consequences of limited English language proficiency will create a devastating impact on ELs' overall success in school. ELs have to advance their English language and literacy skills to catch up with their English-speaking peers to succeed in content subjects and on high-stakes exams (Cisco & Padrón, 2012; Lee, 2005).

The Every Student Succeeds Act (ESSA) requires that ELs' English language proficiency and academic achievement should be indicators of accountability for every public school (ESSA, 2015). It has significant implication for states and schools about the effectiveness of teachers' practices that promote ELs' English language and literacy development as well as content subject learning. In many cases, teachers are prepared with the content and pedagogy to teach grade-level standards, but most of them lack specific knowledge and skills to support ELs’ ability to access the curriculum (Samson & Collins, 2012).

Researchers have suggested that combining effective instructional practice with features of language-sensitive strategies can maximize opportunities for ELs to improve their English

language proficiency and academic performance (Arreaga-Mayer, 1998; Gersten & Jimenez, 1994; Greenwood et al., 2001). Klingner and Edwards (2006) suggested that it is vital to investigate specific teacher instruction practice's impact on students' responses in the classroom.

Researchers have specified that ELs who were struggling readers should be provided with adequate time for independent practice and engagement in structural academic interaction (Anderson, Hiebert, Scott, & Wilkinson, 1985; Foorman & Torgesen, 2001), other than teacher-led English literacy instruction (Calhoon et al., 2007). The challenges and the reality that ELs educators and teachers are facing lead to the questions regarding (a) effective instructional practice/strategies, and (b) how to improve teachers' instructional practices. The cooperative/collaborative/peer-tutoring (CCP) strategy incorporates a number of elements suggested to be important for ELs such as repetition, routines, modeling and practice, and frequent opportunities to respond in a peer-mediated format (Gersten & Baker, 2000; Gersten & Geva, 2003; Vaughn et al., 2006). According to Wong Fillmore, Ammon, McLaughlin, and Ammon (1985), the more opportunities that ELs practiced English with their peers, the more they gained in English language comprehension and production. Furthermore, Strong (1983) suggested that teachers should not put ELs together by default, but instead, they should create situations where ELs want and need to interact and communicate with each other to reach a common goal.

Although the effective instruction has a clear advantage, teachers might benefit from additional support, such as the research-based curriculum and curriculum-based professional development to fully understand how they can implement the strategy (Davis & Krajcik, 2005). For instance, the teachers often do not fully take advantage of the information provided in the curriculum. Also, they might even interpret the information via their experiential lenses (McNeil,

2009). Researchers urge that the persistent achievement gap between ELs and their English speaking peers show a need for structured teacher professional development, which can better and more effectively respond to ELs' language, literacy, and content needs (Calderón, Slavin & Sánchez, 2011; Loucks-Horsley, Stiles, Mundry, Love, & Hewson, 2009; Ortiz & Artiles, 2010).Therefore, Taylor et al. (2015) suggested that the curriculum-based professional development should be integrated with the curriculum as the most effective way to improve teachers' understanding of the philosophy and critical components of the curriculum.

Researchers agree that teachers' instructional quality matters most, especially for ELs (Calderón et al., 2011; Calhoon et al., 2007; Gersten, Baker, Haager, & Graves, 2005; Klingner & Edwards, 2006; Tong, Luo, et al., 2017). Instructional quality was further defined as the instructional content focused on improving ELs' cognitive academic language proficiency (Tong, Luo, et al., 2017). ELs, therefore, can benefit from the CCP learning activities conducted during quality instruction with the support of a research-based curriculum and curriculum-based professional development.

Literature Review on Research-based Curriculum, Curriculum-based VPD, and Classroom Observation

The effective, research-based curriculum and instruction which are modified to meet various needs of ELs, support ELs to make comparable progress with non-ELs in reading, especially at early school levels (Gersten, Dimino, Jayanthi, Kim, & Santoro, 2010; Gersten et al., 2007; Shanahan & Beck, 2006). Instructional support will become nearly indispensable as ELs learn English language and literacy and subject content at the same time. The understanding of effective professional development (PD) can improve teachers' pedagogical practices and provide quality instruction for ELs is indispensable (Lara-Alecio et al., 2009; Lee & Buxton,

2013; Tong, Luo, et al., 2017). Therefore, it is crucial to provide a curriculum-based PD for the teachers to understand the philosophy and the essential components of the curriculum.

However, research considering how PD influences teachers' instruction and students' achievement remains inadequate (Avalos, 2011; Lesaux & Gamez 2012; Shaha et al., 2004; Tong et al., 2010; Tong, Luo, et al., 2017). This inadequacy is partially due to the conventional format of PD, which requests the absence of teachers from their daily routine and classroom, resulting in reluctance, burden, and expense (Shaha et al., 2015). Darling-Hammond, Wei, Andree, Richardson, and Orphanos (2009) suggest that adding systematic approaches to PD would lead to more benefits in the improvement of teaching skills and better outcomes for students. As proposed by Darling-Hammond et al. (2009), it is necessary that PD should follow a cycle of continuous improvement by which teachers examine student data to identify what students need, support the students with powerful lessons, apply new strategies to scaffold students' learning, refine the new learning into more powerful lessons, reflect on the impact of PD on students' learning, and repeat the circle with new objectives.

Dede (2006) suggested that online PD is a more effective and economical approach since it comprehensively considered what can be delivered and how it can be delivered and measured based on learners' objectives and training purposes. Bates et al. (2016) further differentiated three types of online PD: synchronous, asynchronous and hybrid. Synchronous online PD refers to "online learning activity happen[ing] in real time" (p.71). Asynchronous online PD refers to "online learning activities happen[ing] at different times for different participants" (p. 71). Hybrid online PD refers to "online learning activities take place as part of larger in-person learning opportunities" (p.71). In my study, virtual PD was a type of online PD that was conducted virtually in real time, which was considered synchronous online PD.

Online PD has shown a positive impact on students' academic performance in an increasing number of studies (Dash et al., 2012; Rienties, Héliot, & Jindal-Snape, 2013; Shaha & Ellsworth, 2013; Wasik & Hindman, 2011). Collins and Liang (2014) suggested in EL education field that online PD allowed the teachers to obtain an individualized experience with affordable opportunities, which supported the teachers in EL classrooms. However, there was no RCT study examining the impact of VPD on curriculum-based training for EL teachers (Tong, Irby, & Lara-Alecio, 2015). Researchers called for more research that provides solid evidence of the effectiveness of online PD/VPD and its' impact on teachers' classroom pedagogical practices as well as on ELs' improvement in English language proficiency (Smith, 2014).

Researchers have suggested that research-based data, especially data collected via the classroom-based measures were more accurate to identify effective classroom practices (Kane, Taylor, Tyler, & Wooten, 2011). The actual occurrences of instructional practices within ESL/bilingual classroom would be revealed by the multidimensional classroom observation, providing corresponding empirical evidence for the pedagogical practices of EL teachers (Lara-Alecio et al., 2009). Classroom observation can be adopted as a reliable approach to identify how teachers' pedagogical practices and instruction quality are improved as the results of VPD and thereby bring forth improving ELs' reading achievement (Tong, Luo, et al., 2017). It was pointed out by Joyce and Calhoun (2010) that with no reliable classroom observation data, it cannot be determined whether the teachers brought what they received in PD training into their classrooms. In the previous studies, scholars examined the impact of effective PD for student gains (Shaha & Ellsworth, 2013; Shaha et al., 2015; Tong, Luo et al., 2017). However, the connection between VPD and teachers' pedagogical practices obtained from classroom observation as well as the coordinated impact of those two factors on student achievement are unknown.

Theoretical Perspectives

Constructivism, the theoretical foundation of my study, framed the CCP strategy employed in the EL classroom. My work was based on three theories: socio-cognitive theory (Piaget 1932), sociocultural theory (Vygotsky, 1978), and Four-Dimensional Bilingual Pedagogical Theory (Lara-Alecio & Parker, 1994). According to Piaget (1932), people produce knowledge based on their previous knowledge and experience. His theory covered learning theories as well as teaching methods. Within Piaget's framework, cognitive development involves five key concepts: schema, assimilation, equilibration, disequilibrium, and accommodation. Schema refers to the basic building blocks of intellectual development (Piaget, 1932). Assimilation refers to a stage that a person is using existing schema/knowledge/experience to deal with a new object/situation. "Every schema is coordinated with other schemata and itself constitutes a totality with differentiation parts" (Piaget, 1932. p.7). Accommodation refers to a creation of a new schema/knowledge/experience. During the accommodation process, a person is connecting new schema/knowledge/experience to their previous schema/knowledge/experience and restructures his or her schema.

Piaget used the concept of equilibration-disequilibrium to explain the causes of cognitive development when a person responds to a new situation. For example, a child's existing knowledge can be challenged by his or her peers during the CCP interaction, which is the state of disequilibrium. It is the same case for teachers during training or teacher professional development. When a teacher is provided with a new teaching strategy or new content, his/her existing knowledge/experience can be challenged during his/her interaction with the coach or the PD provider, which is also the state of disequilibrium. The child or the teacher needs to either develop a new scheme or adapt their existing knowledge until the disequilibrium is solved and

the brain can go back again to an equilibration state. The process of disequilibrium being encountered and solved is the process of growing and developing. The conflicts that arise from the interaction bring children to a higher level of cognitive process, which is the cognitive growth (Zhang et al., 2013). Besides the field of children education, Piaget's cognitive development process has been advocated and supported in teacher preparation (Ward, Pellett, & Perez, 2017) and e-learning environments (Gillani, 2010) in adult education.

Vygotsky's sociocultural theory emphasized the importance of dialogical experience in children's intellectual development (Vygotsky, 1962). Vygotsky emphasized the critical relationship between teachers' prior knowledge (a more competent party) and students' prior knowledge as well as the communication between the teachers and the students. According to Vygotsky, active exchanges of ideas among peers and between teachers and students are critical for children' cognitive growth and language development since language is a major cognitive tool for children to process information (Mercer, 2013). In Vygotsky's sociocultural theory, social interaction is important for children to learn knowledge. Children can communicate with their peers in constructing projects and solving problems. In this way, they exchange their ideas, generate new ideas and develop new knowledge. Moreover, Vygotsky's theory implies that classroom instruction should provide the opportunities for student-to-student interaction or CCP learning activities so that the students can develop concepts from their peers whose understandings and interpretations are at a similar level as theirs.

Four-Dimensional Bilingual Pedagogical Theory (Lara-Alecio & Parker, 1994) was developed to identify the interactions among four major dimensions (language of instruction, language content, communication mode, and activity structures) within bilingual classrooms. Two minor domains (physical group and ESL strategy) were added in the validation study

(Bruce, Lara-Alecio, Parker, Hasbrouck, Weaver, & Irby, 1997) and the application study (Lara-Alecio et al., 2009). In the domain of language content, Four-Dimensional Bilingual Pedagogical Theory further extended Cummins' BICS and CALP theory and differentiated cognitive content levels into social routines, academic routines, light cognitive content and dense cognitive content. Teachers' time allocation in the light and dense cognitive content has been an important indicator of the quality of instruction and plays a crucial role in improving ELs' English language proficiency (Tong, Luo, et al., 2017). This differentiation of four cognitive levels of language content and its interaction with the domains of ESL strategy and language of instruction allowed me to assess the occurrences of the CCP strategy being implemented in the dense cognitive content taught in English.

My Theoretical Framework

For my investigation, I conceptualized learning by considering each individual in the EL classrooms, a teacher or an EL, as experiencing cognitive growth in the circle of assimilation, equilibration, new situation, disequilibrium, and accommodation. A teacher was provided with the curriculum-based PD and interacted with his/her PD provider. He/she assimilates the new content/strategy with his/her existing schema to achieve the state of equilibration. He/she might also encounter new content/strategy that cannot fit into his/her existing schemas. To solve the disequilibrium, he/she progresses through the state of cognitive development by changing the internal schema to create balance with the new content/knowledge. During this process, the teacher modifies the old structure of knowledge and creates a new one to better fit his/her assimilated information (Cook & Cook, 2005). For students, including ELs, when they work with their peers, they receive and respond to the challenges from each other. They might assimilate the new content with their existing knowledge. Alternatively, their existing schema

can be changed to accommodate and include the new knowledge. Aloqaili (2012) stressed that both accommodation and assimilation account for intellectual adaptation and development. Teachers' cognitive development circle is stimulated by the interaction between the teachers and their PD providers, while for the students, the circle is stimulated by the interaction between the students and the teachers as well as the students and their peers. Interaction is an important way for the teacher to learn from their PD providers and for students to learn from their teachers and peers. During the interaction, ELs can accommodate their existing knowledge to fit their English language learning by exchanging their ideas in English with the support of their teachers' applications of a research-based curriculum where the CCP strategy is embedded.

In EL classrooms, teacher-student and student-student interactions are not limited to a single dimension. Instead, the interaction is multidimensional, including what language the teacher/the students are using, what language mode they are using, or what level of cognitive content the teacher is teaching. The Four-Dimensional Bilingual Pedagogical Theory provides a theoretical framework to observe all the dimensions of the EL classrooms. As demonstrated in Figures 3 and 4, the cognitive development circle occurs at both teacher level and student level. Teachers implemented the research-based curriculum with the support of the curriculum-based professional development, which stimulates teachers' cognitive development circle. Since the professional development was conducted virtually, the distances between the PD provider and each teacher are the same. In my specific study, I proposed the ELs could benefit from the CCP strategy that was embedded in the curriculum and the VPD. Moreover, I defined quality CCP learning as the occurrences that happened during the dense cognitive content when the students were responding in English. Therefore, Four-Dimensional Bilingual Pedagogical Theory was applied with a focus on quality CCP learning in the EL classrooms.

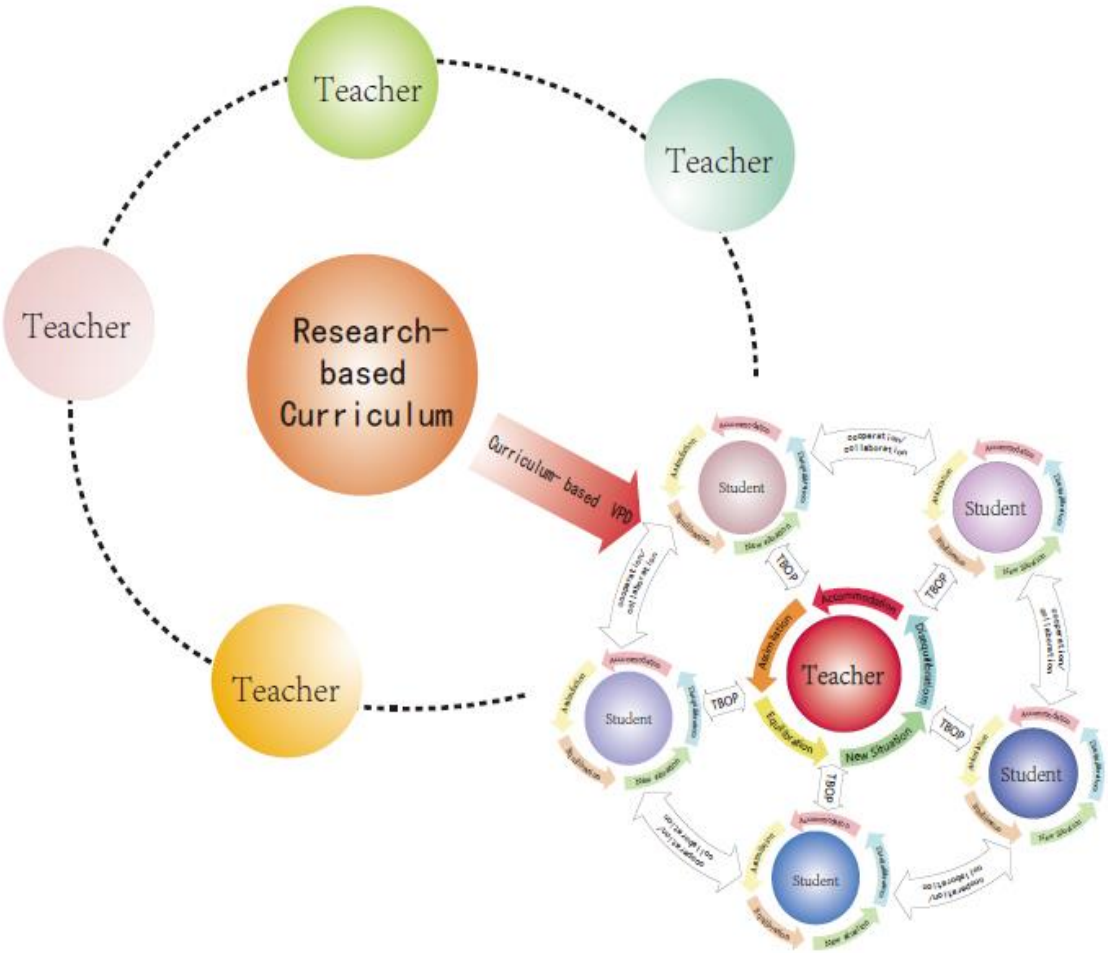

Figure 3. A VPD solar system of cognitive development of ELs and their teachers.

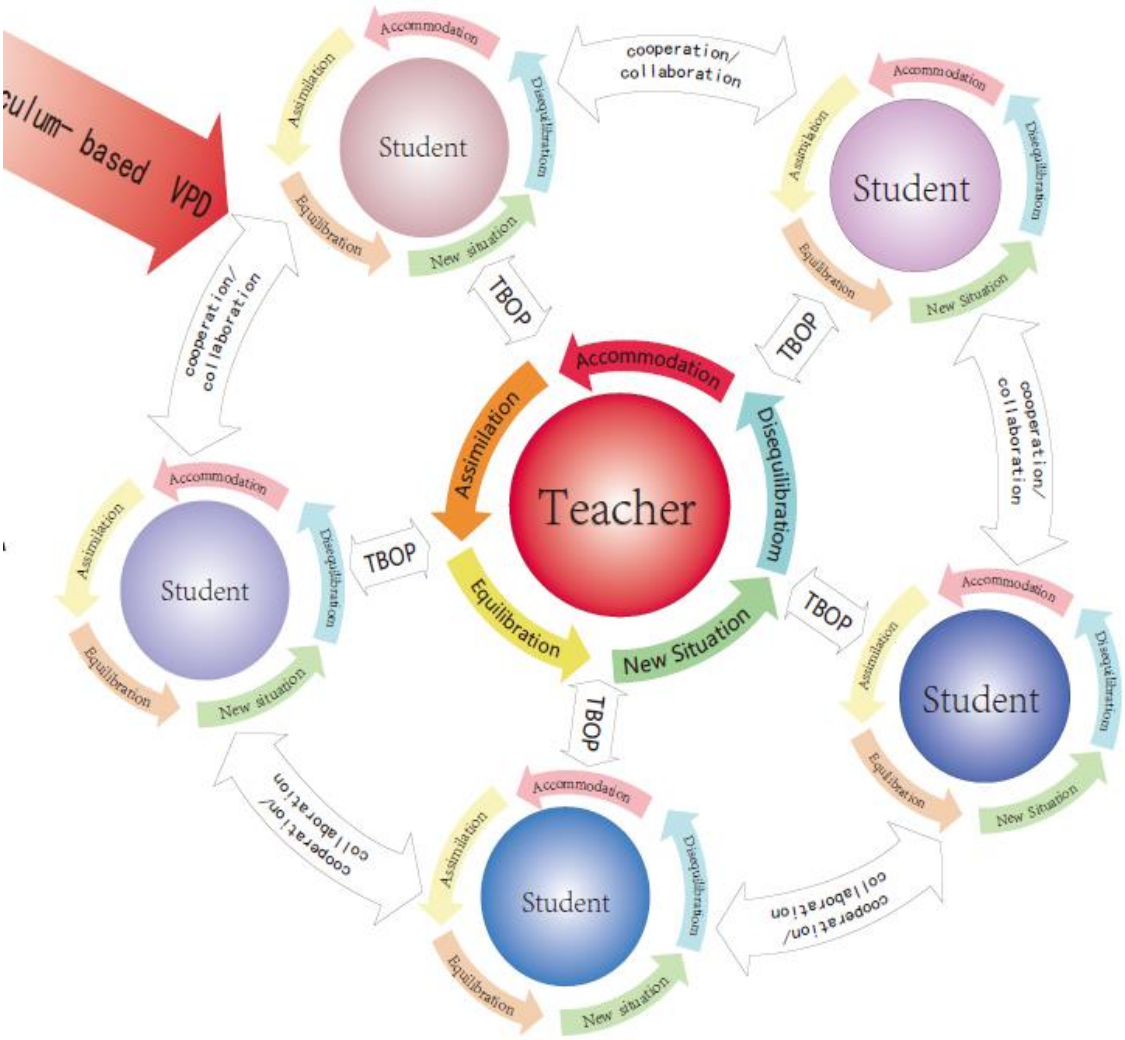

Figure 4. A snapshot of the VPD solar system in an EL classroom.

In my theoretical framework, I proposed that to close the achievement gap between ELs and mainstream students, we need to improve teachers' instructional quality which can be supported by professional development. The research-based curriculum and the curriculum-based professional development emphasize students' opportunities to reconcile their previous ideas and experience, to create new knowledge, and to reflect on and discuss what they are doing

and how their ideas are changing. In this study, I was particularly interested in the interaction among the domains of language of instruction, language content and ESL strategy and how this interaction improves ELs' English reading and oral expression. The multidimensional classroom observation instrument, Transitional Bilingual Observation Protocol (TBOP, Lara-Alecio & Parker, 1994) that was developed from the Four-Dimensional Bilingual Pedagogical Theory, was applied to observe the instructional languages and the levels of language content. To be specific, in this study, I investigated how the research-based curriculum and the curriculum-based PD supported the EL teachers' implementation of quality CCP learning activities in EL classrooms.

To my knowledge, no empirical investigation to date has studied the effect of quality CCP learning on the first-grade ELs' English reading and speaking proficiency with the support from the ongoing and intensive curriculum-based VPD. In this study, which derived from an ongoing RCT, I aimed to (a) observe and compare instructional time in the quality CCP learning when the students were responding in English (coded as CCP_Dense_SL2) between two conditions (i.e., the treatment with VPD vs. the control with traditional district PD), (b) associate the impact of ongoing and intensive VPD with ESL intervention on ELs' English reading and speaking development, and (c) conduct an exploratory analysis to examine the effective dosage of time allocation in the quality CCP learning that contributed to the improvement of ELs' English reading fluency, reading comprehension, and oral expression development within three-level hierarchical linear models. Four research questions guided the study.

1. When the students were responding in English, did the instructional time allocated in quality CCP learning significantly differ between conditions (as a result of VPD) as observed in Project ELLA-V via the classroom observation instrument TBOP?

2. Did ELs' English oral expression, reading comprehension, and reading fluency

achievement vary across different classrooms and schools?

3. Did ELs' English oral expression, reading comprehension, and reading fluency differ between the treatment (as a result of instructional intervention including VPD) and the control conditions, controlling for ELs' initial achievement on these measures?

4. What was the effective dosage of time allocation in quality CCP learning that contributed to the improvement of Grade 1 ELs' reading comprehension, reading fluency, and oral expression?

Method

In this study, I analyzed the archived data from Project ELLA-V 2015-2016 with the permission of the project research team. Two sets of data were included in this study: (a) Grade 1 ELs' reading comprehension and oral expression examined by Woodcock-Muñoz Language Survey-Revised (WMLS-R) and reading fluency examined by Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and (b) Grade 1 instructional time allocated in quality CCP learning, measured by the classroom observation instrument TBOP. I was able to examine the impact of the treatment effect as a result of VPD on teachers' time allocation in the CCP_Dense_SL2 instructional period, the direct impact of the treatment effect and the mediating impact of CCP_Dense_SL2 instructional time on first grade ELs' reading compression, oral expression proficiency. I was able to further identify the effective baseline of teachers' time allocation in the CCP_Dense_SL2 instructional period.

In Project ELLA-V, all the participating districts with more than 6,000 enrollments in the school year of 2011-2012 were selected. Among them, all campuses that had over 60 ELs in Grade 2 were selected. Schools that had similar school demographics including a similar overall English proficiency rated by Texas English Language Proficiency Assessment System

(TELPAS), percentage of economically disadvantaged students, and percentage of ELs were clustered in a group of three. Within the group, each school was randomly assigned to three different conditions: Treatment 1, Treatment 2, and control conditions. The school remained in the same condition through the whole project. The participating schools were mostly rated as highly economically disadvantaged schools with an average of 92.1% of the students on the free/reduced lunch program according to the Academic Excellence Indicator System (AEIS) of the academic year 2011-2012. The average percentage of students classified as ELs in the participating schools was at 64.4%. The average composite scores in TELPAS was 2.1 ranging from 1.0 to 2.7. In this study, I put the focus on the comparison between Treatment 2 and the control condition.

Intervention: Student Level

Treatment. Treatment included two blocks of a 35-minute Story-telling, Retelling and Higher Order Thinking for English Language and Literacy Acquisition (STELLA) and a 10-minutes Academic Oral and Written Language in Science (AOWLS). In STELLA, a curriculum aimed to facilitate interactive story reading between teacher and student and develop ELs' reading comprehension via higher order questioning and thinking strategies, provided ELs with meaningful opportunities to respond and engage in English language and literacy activities and expand their vocabulary and listening and speaking skills. In AOWLS, curriculum targeted developing ELs' academic vocabulary development and science concepts by providing sentences using words in context and asking a daily question using the target word. Students read aloud to the teacher, discussed with their partners and worked together to solve a problem or answer a question about the story, or wrote a response in a journal.

Control condition. Typical district ESL instruction included 45 minutes of instruction daily. The instruction in the control condition followed district standards-based guidelines and resources, which varied from district to district and school to school. Teachers in the control condition provided subjects as language arts, reading in English, or content-area instruction (e.g., math or science). The instruction was conducted in full classes or small groups.

Intervention: Teacher Level

The CCP learning component in treatment condition. The treatment teachers were provided with scripted curriculum, in which guided cooperative questioning, scripted CCP activities, think-pair-share, and share-reading were included purposefully combined with linguistic patterns in content-related activities, which allowed ELs exposure to different levels of complexity of thinking and learning. The highly structured CCP activities aimed at encouraging and ensuring ELs' successful participation in higher-order thinking activities.

The CCP learning component in the control condition. The teachers in the control condition implemented the state-aligned curriculum. The CCP strategy might also be embedded in the curriculum. The emphasis on this strategy differed from teacher to teacher.

Intervention: School Level

Treatment. Bi-weekly VPD was provided to the treatment teachers within their intervention condition. Treatment was an ongoing, 90-minute VPD via Citrix GoToTraining in which the intervention/curriculum coordinators interacted with the treatment teachers via screen sharing, voice, webcam, and chat. During VPD sessions, the coordinators and the treatment teachers reflected on and discussed student learning and previewed upcoming lessons and materials. Teachers received training on L2 acquisition strategies, ESL instruction strategies, and

how to conduct self-assessment via virtual professional portfolio. The treatment teachers were also provided video clips of instruction modeling.

Control condition. The control teachers received state aligned professional development provided by school districts. Resources and forms varied from district to district.

Participants

Since the Project ELLA-V was a three-level intervention, the multi-level analysis was adopted. The participants in this study were reported at three levels: school, teacher, and student.

School level. In this study, I used a subsample from Treatment 2 and the control conditions in the first grade of the Project ELLA-V. All selected schools were located in Texas and had participated in the project for three years. There were 41 schools involved in the project: 20 from Treatment 2 and 21 from the control condition in the first grade.

Teachers level. All selected teachers were the first-grade teachers from Treatment 2 and the control conditions. It was a one-year intervention for these teachers. There were 77 first grade teachers from 41 schools participating in the project, 38 from T2 and 39 from the control condition in the first grade.

Student level. The students in this study were from Treatment 2 and the control conditions in the first-grade intervention of Project ELLA-V. All students were first-grade ELs. There were 1,241 first-grade ELs, 627 from T2 and 614 from the control condition.

Instrumentation

In the current study, three instruments were adopted: a) Transitional Bilingual Observation Protocol (TBOP) to rate quality CCP learning; b) Woodcock-Muñoz Language Survey-Revised (WMLS-R, Alvarado, Ruef, & Schrank, 2005) to assess ELs' reading comprehension and oral expression proficiency, and c) Dynamic Indicators of Basic Early

Literacy Skills Oral Reading Fluency (DIBELS ORF, 6th ed; Good & Kaminski, 2007) to assess ELs' reading fluency.

TBOP. TBOP is an observation analysis tool developed from the Four-Dimensional Transitional Bilingual Pedagogical Theory (Lara-Alecio & Parker, 1994). It was designed to observe: (a) the occurrences of teachers' practices in four major instructional domains—language of instruction, language of content, communication mode, and activity structure, and two minor domains—ESL strategies and group, b) the interactions among the domains, and c) the fidelity of implementation of intervention at the teacher level over time (Figure 2).

TBOP is a multidimensional classroom observation instrument, in which one of the minor domains, ESL strategies, include nine strategies: academic language scaffolding, visual scaffolding, questioning strategies, the CCP strategy, advance organizer, manipulatives and realia, content connections, first language clarification, and integrate technology. In this study, I focused on teachers' time allocation on the CCP strategy (the domain of ESL strategy) during the dense cognitive content (the domain of language of content) instruction when ELs were responding in English (the domain of language of instruction-student). Time allocation in this type of CCP learning was defined in this study as quality CCP learning time. TBOP raters received rigorous training at the beginning and middle of the coding process. According to the classical inter-rater benchmark developed by Landis and Koch (1977), the initial inter-rater reliability was established at the substantial level (adjusted Gwet $AC_1 = 0.64$). Inter-rater reliability was also monitored through the coding process, which generated an average continuous inter-rater reliability at the upper bound of the moderate level (adjusted Gwet $AC_1 = 0.60$).

WMLS-R. WMLS-R is a standardized test of CALP, developed by Riverside Publishing. It has seven subtests, which measures English listening, speaking, reading, and writing proficiency. Seven subtests included picture vocabulary, verbal analogies, letter-word, identification, dictation, understanding directions, story recall and passage compression. Different combinations of seven subtests generate 11 clusters, measuring the specific aspects of English academic language proficiency. In the current study, I focused on two clusters of WMLS-R: oral expression and reading. Oral expression includes two subtests: picture vocabulary and story recall, which measures expressive vocabulary, English language comprehension and development, and memory. Reading comprises letter-word identification and passage comprehension, which measures students' ability of letter/word identification and understanding of written passages. The median reliability of reading and oral expression is 0.92 and 0.95 respectively for the 5 to 19 age range.

DIBELS ORF. DIBELS ORF is a standardized, individually administered test of students' accuracy and fluency to identify whether the students need additional English language instructional support and to monitor students' progress towards instructional goals. DIBELS ORF (6th ed; Good & Kaminski, 2007) was adopted by Project ELLA-V to measure Grade 1 ELs' reading fluency in the school year 2015-16. It consisted of three separate reading passages that the students were asked to read aloud for 1 minute. If any words were omitted or substituted or the students showed hesitation of more than three seconds, it would count as an error. The number of words that were accurately read in that minute was calculated. The median number of accurately-read words from three passages was used as the final score. Tindal, Marston, and Deno (1983) reported that the reliability of test-retest was from .92 to .97 at the elementary level and alternate form reliability ranged from .89 to .94.

Data Collection

At the beginning of Project ELLA-V, IRB approval was given for all research related to the project (IRB2012-0762D). Consent forms in both Spanish and English were provided to parents to be signed. Teachers and principals also consented to participation in the project.

Student data. Six subtests of WMLS-R were administered among first-grade ELs at the beginning and end of the school year 2015-2016: picture vocabulary, verbal analogies, letter-word identification, understanding directions, story recall, and passage comprehension. In this dissertation study, I focused on participants' oral expression and reading measured by WMLS-R and reading fluency measured by DIBELS ORF. The WMLS-R raw scores were converted into grade-based standardized scores to analyze. DIBELS ORF standardized scores were the median scores from three passages with which the participants were tested.

Teacher data. Grade 1 teachers in the treatment and the control conditions were observed and recorded virtually in a secure observation lab at the beginning, middle, and end of the school year of 2015–2016 via pre-installed observation cameras in each classroom. All teachers were recorded at the beginning, middle, and end of the school year during ESL block, creating a total of 231 videos, 114 in Treatment 2 and 117 in the control condition. The average length of observation was 49 minutes for a Treatment 2 teacher and 48 minutes for a control teacher. Their recorded observation was coded via TBOP to analyze their pedagogical practice during the ESL block. The detail coding procedure is introduced in the data collection section. There were 13,860 20-second video clips coded: 6,840 in T2 and 7,020 in the control condition.

The recorded observations were rated by trained research personnel via an online web tool embedded with TBOP for raters to code teachers' pedagogical practice occurrences. Raters coded the first, third, fifth and seventh five-minute period of the 45 minute recorded lessons,

producing 20-minutes of coded instruction. For these four, five-minute intervals, raters coded 20-second interval clips and got 15 codes for each five-minute interval, producing 60 codes for one single teacher. The percent of instructional time allocated in the CCP strategy during the dense cognitive content when the students responded in English was collected as one of the predictors of student achievement.

Fidelity of Implementation

The Teacher Observation Record (TOR) was adopted to measure FOI across conditions. TOR was developed and utilized in previous RCT studies (e.g., Tong et al., 2008; Lara-Alecio, Tong, Irby, Guerrero, Huerta, & Fan, 2012). It was later modified into nine items to fit the Grade 1 curriculum (Tong, Irby, et al., 2017). It measured nine areas (a) teacher preparation, (b) material usage and preparation, (c) lesson pacing, (d) student involvement, (e) teacher talking, reading, writing time vs. student talking, reading, and writing time, (f) use of leveled questioning, (g) use of ESL strategies, (h) effective and cognitive feedback, and (i) physical environment. Overall, FOI was 0.94 for Treatment 2 and 0.84 for the control condition. Raters received rigorous training to meet inter-rater reliability (IRR) above the substantial level (Landis & Koch, 1977). IRR was established and monitored through the FOI measuring process to ensure rating quality.

Data Analysis

A multi-level analysis was adopted in this study. Therefore, the variables described in this section included three levels: student level variables as level 1 variables, teacher level variables as level 2 variables, and school level variable as level 3 variable.

Student level variables. Six student-level variables were involved in this study: pre-testing reading comprehension achievement, post-testing reading comprehension achievement,

pre-testing oral expression achievement, post-testing oral expression achievement, pre-resting reading fluency achievement and post-test reading fluency achievement. Students' reading comprehension and oral expression proficiency were measured by WMLS-R and then were converted to the grade-based standardized scores that were included in this study. Students' reading fluency was measured by DIBELS ORF. The raw scores were transferred to standardized scores and included in this study. Student's' reading comprehension, oral expression, and reading fluency achievement that were collected at the beginning of school year were students' pre-testing achievement in reading comprehension, oral expression, and reading fluency, while the scores that were collected at the end of the school year were students' post-testing achievement in these three measures. Since achievement data have missing values in pre- and post-test and Hierarchical Linear Model (HLM) allows missing values at level 1, HLM was adopted as the major analytic method (Raudenbush & Bryk, 2002).

Teacher level variables. Instructional time in the quality CCP learning was the only teacher level variable involved in this study, which was coded as CCP_Dense_SL2. Instructional time is a continuous variable measured by TBOP. It was the overall percentage of instructional time where teachers used the CCP strategy during dense cognitive content period across three rounds of classroom observation when the ELs were responding in English. The CCP_Dense_SL2 instruction time was further classified as a four-level ordinal variable CCP_Dense_SL2_C: lower than the 25th percentile, 25th percentile to 50th percentile, 50th percentile to 75th percentile, and above the75th percentile.

School level variable. The condition was the school level variable. The condition was a binomial variable with "1" representing the treatment and "0" representing the control. The

teachers in the treatment condition received structure and ongoing, intensive VPD, and teachers in the control condition received state-aligned PD provided by the district.

Statistical Procedure

In order to answer the first research question, an independent-sample t-test was used to examine whether there was any difference in the time allocation of the CCP strategy during dense cognitive content instruction while the ELs were responding in English (CCP_Dense_SL2) between the treatment and the control conditions. Moreover, SAS PROC MIXED was adopted to estimate the variance of four three-level Hierarchical Linear Models that were involved in this study. The first model was a three-level baseline equivalence HLM model using students' pretest achievement as the dependent variable and the condition that the teachers were randomly assigned to as the level 3 predictor. This model was replicated in three outcomes: reading comprehension, oral expression, and oral reading fluency. The second model was a null model that only included student post-test achievement as the dependent variable. The null model was replicated in three scenarios: reading comprehension, oral expression, and oral reading fluency. The third three-level Hierarchical Linear Models was established as (a) WMLS-R reading posttest scores as the dependent variable (W_Reading_Post) with WMLS-R reading pretest scores (W_Reading_Pre) as level 1 predictor and the condition (Treatment) that the teachers were randomly assigned to as level 3 predictor, (b) WMLS-R oral expression posttest scores as the dependent variable (W_OE_Post) with WMLS-R oral expression pretest scores (W_OE_Pre) as level 1 predictor and the condition (Treatment) that the teachers were randomly assigned to as level 3 predictor; and (c) DIBELS ORF posttest scores (ORF_Post) as the dependent variable with DIBELS ORF pretest scores (ORF_Pre) as level 1 predictor and the condition (Treatment) that the teachers were randomly assigned to as level 3 predictor.

The fourth three-level Hierarchical Linear Model was established as (a) WMLS-R reading posttest scores as the dependent variable (W_Reading_Post) with WMLS-R reading pretest scores (W_Reading_Pre) as level 1 predictor, ordinal variable CCP_Dense_SL2_C as level 2 predictor and condition (Treatment) that teachers were randomly assigned to as level 3 predictor, (b) WMLS-R oral expression posttest scores as the dependent variable (W_OE_Post) with WMLS-R oral expression pretest scores (W_OE_Pre) as level 1 predictor, ordinal variable CCP_Dense_SL2_C as level 2 predictor and condition (Treatment) that teachers were randomly assigned to as level 3 predictor, and (c) DIBELS ORF posttest scores (ORF_Post) as the dependent variable with DIBELS ORF pretest scores (ORF_Pre) as level 1 predictor, ordinal variable CCP_Dense_SL2_C as level 2 predictor and condition (Treatment) that teachers were randomly assigned to as level 3 predictor. I used SPSS 24 to clean data and obtain the percentage of time that the teachers allocated for quality CCP learning (CCP_Dense_SL2). SPSS 24 was also adopted to conduct the independent-sample t-test. SAS 9.4 was adopted to estimate four three-level Hierarchical Linear Models.

Model Specification

Four HLM modes was established to examine in three circumstances (oral expression, reading comprehension, and oral reading fluency): (a) whether the baseline equivalence between the students in the treatment and the control conditions was established, (b) what variances at classroom level and school level were, (c) how the curriculum-based VPD impacted ELs' improvement in reading comprehension, oral expression, and oral reading fluency, and (d) what the effective dosage of CCP_Dense_SL2 learning was to improve ELs' achievement in reading comprehension, oral expression, and oral reading fluency. Four models addressed the research questions 2-5 respectively.

Baseline Equivalence. In order to answer the second research question, a three-level HLM model was adopted. It utilized students' pretest achievement at the beginning of the intervention as the dependent variable in three separate circumstances: reading comprehension, reading fluency, and oral expression (one for each of the three dependent variables). The purpose of this step was to examine whether there was any statistically significant difference between students in the treatment and the control conditions regarding their performance in reading comprehension, reading fluency, and oral expression. The model was specified as follow:

Level 1: $PRE_{ijk} = \beta_{0jk} + r_{ijk}$

- PRE_{ijk} is the pre-testing achievement (reading comprehension, reading fluency, and oral expression) for student *i* in class *j* in school *k*
- β_{0jk} is the adjusted mean achievement for class *j* in school *k*
- r_{ijk} is the random effect associated with student *i* in class *j* in school *k*

Level 2: $\beta_{0jk} = \gamma_{00k} + u_{0jk}$

- γ_{00k} is the average mean achievement (reading, oral expression, and reading fluency) in school *k*
- u_{0jk} is the random effect associated with class *j* in school *k*

Level 3: $\gamma_{00k} = \eta_{000} + \eta_{001}(\text{CONDITION}_k) + v_{00k}$

- η_{000} is the grand mean achievement overall schools (reading, oral expression, and reading fluency)
- η_{001} is school k deviation from the mean
- v_{00k} is a level 3 random effect across schools.

Three-level HLM unconditional model. In order to address research question three, the unconditional model was adopted. It was the model that was unconditional on all covariates. The

three-level unconditional model for the analysis of post-testing utilized students' post-testing achievement as dependent variables in three separate situations: reading comprehension, reading fluency, and oral expression (one for each of the three dependent variables). The model was specified as:

Level 1: $POST_{ijk} = \beta_{0jk} + r_{ijk}$

- $POST_{ijk}$ is the post-testing achievement (reading comprehension, reading fluency, and oral expression) for student *i* in class *j* in school *k*
- β_{0jk} is the adjusted mean achievement for class *j* in school *k*
- r_{ijk} is the random effect associated with student *i* in class *j* in school *k*

Level 2: $\beta_{0jk} = \gamma_{00k} + u_{0jk}$

- γ_{00k} is the average mean achievement (reading, oral expression, and reading fluency) in school *k*
- u_{0jk} is the random effect associated with class *j* in school *k*

Level 3: $\gamma_{00k} = \eta_{000} + +v_{00k}$

- η_{000} is the grand mean achievement overall schools (reading, oral expression, and reading fluency)
- v_{00k} is a level 3 random effect across schools.

The purpose is to estimate the variance of components of Level 1 (σ_e^2), Level 2 (σ_{u2}^2), and Level 3 (σ_{u3}^2). The unconditional ICCs for teacher level is $\rho_2 = \sigma_{u2}^2 / (\sigma_e^2 + \sigma_{u2}^2 + \sigma_{u3}^2)$ and for school level is $\rho_3 = \sigma_{u3}^2 / (\sigma_e^2 + \sigma_{u2}^2 + \sigma_{u3}^2)$.

Three-level pretest covariate and condition HLM model. In order to address the fourth research question, a three-level pretest covariate and condition HLM model was adopted. I used

the pretest grand mean centering as suggested by Hedges and Hedberg (2007, 2013) since it could make the estimates of variance components and interpretation of results more meaningful. This model utilized students' post-testing achievement at the end of intervention as dependent variables in three separate instances: reading comprehension, reading fluency, and oral expression (one for each of the three dependent variables). Also, students' pre-testing grand-centered achievement at the beginning of the intervention was add in level-1 of the model as student level covariate (reading comprehension, reading fluency, and oral expression, one for each of the three covariates in three separate instances). A condition that schools were randomly assigned to was added in as a level-3 predictor.

Level 1: $\text{POST}_{ijk} = \beta_{0jk} + \beta_{1jk}*(PRE\_C_{ijk}) + r_{ijk}$

- $POST_{ijk}$ is the post-testing achievement (reading comprehension, reading fluency, and oral expression) for student *i* in class *j* in school *k*
- PRE_C_{ijk} is the grand-centered pre-testing achievement (reading comprehension, reading fluency, and oral expression) for student *i* in class *j* in school *k*
- β_{0jk} is the adjusted mean achievement for class *j* in school *k*
- r_{ijk} is the random effect associated with student *i* in class *j* in school *k*

Level 2: $\beta_{0jk} = \gamma_{00k} + u_{0jk}$

$\beta_{1jk} = \gamma_{10k}$

- γ_{00k} is the average mean achievement (reading, oral expression, and reading fluency) in school *k*
- u_{0jk} is the random effect associated with class *j* in school *k*

Level 3: $\gamma_{00k} = \eta_{000} + \eta_{001}(\text{CONDITION}_k) + v_{00k}$

$\gamma_{10k} = \eta_{100}$

- η_{000} is the grand mean achievement overall schools (reading, oral expression, and reading fluency)
- η_{001} is school k deviation from the mean
- v_{00k} is a level 3 random effect across schools.

Combining the individual models gives the following mixed regression model:

$$POST_{ijk} = \eta_{000} + \eta_{100}*(PRE\_C_{ijk}) + \eta_{001}(CONDITION_k) ++v_{00k} + u_{0jk} + r_{ijk}$$

Three-level pretest covariate, condition and the CCP strategy HLM model. In this model, the students' post-testing achievement was the dependent variable in three separate instances: reading comprehension, reading fluency, and oral expression (one for each of the three dependent variables). This model was established to answer the last research question. Also, students' grand-centered pre-testing achievement was added in level-1 of the model as student level covariate (reading comprehension, reading fluency, and oral expression, one for each of the three covariates in three separate instances). The instructional time in the quality CCP learning was classified by four levels: below 25th, between 25th and 50th, between 50th and 75th, and beyond the 75th percentile. This four-level ordinal variable was coded as CCP_Dense_SL2_C and was added in level 2. The condition that schools were randomly assigned to was added in as a level-3 predictor.

Level 1: $POST_{ijk} = \beta_{0jk} + \beta_{1jk}*(PRE\_C_{ijk}) + r_{ijk}$

- $POST_{ijk}$ is the post-testing achievement (reading comprehension, reading fluency, and oral expression) for student *i* in class *j* in school *k*
- PRE_C_{ijk} is the grand-centered pre-testing achievement (reading comprehension, reading fluency, and oral expression) for student *i* in class *j* in school *k*
- β_{0jk} is the adjusted mean achievement for class *j* in school *k*

- r_{ijk} is the random effect associated with student *i* in class *j* in school *k*

Level 2: $\beta_{0jk} = \gamma_{00k} + \gamma_{01k}$ CCP_Dense_SL2_C$_{jk} + u_{0jk}$

$\beta_{1jk} = \gamma_{10k}$

- γ_{00k} is the average mean achievement (reading comprehension, reading fluency, and oral expression) in school *k*
- γ_{01k} is the predicted change to overall teachers' CCP_Dense_SL2_C instructional time
- u_{0jk} is the random effect associated with class *j* in school *k*

Level 3: $\gamma_{00k} = \eta_{000} + \eta_{001}$(CONDITION$_k$) $+v_{00k}$

$\gamma_{01k} = \eta_{010}$

$\gamma_{10k} = \eta_{100}$

- η_{000} is the grand mean achievement overall schools (reading comprehension, reading fluency, and oral expression)
- η_{001} is school k deviation from the mean
- v_{00k} is a level 3 random effect across schools.

Combining the individual models gives the following mixed regression model:

POST$_{ijk} = \eta_{000} + \eta_{100}$*(*PRE_C*$_{ijk}$) + η_{010} CCP_Dense_SL2_C$_{jk} + \eta_{001}$(CONDITION$_k$) ++v_{00k} + u_{0jk} + r_{ijk}

Results

In all, 13,860 coded 20-second video observations equated to 4,620 minutes of coded instruction time. A total of 1,241 Grade 1 ELs, 77 Grade 1 bilingual teachers, and 41 schools were included in this study. In Table 13, I described the participants by level and by condition. There were 2,280 minutes of instructional hours observed in the treatment condition and 2,340

minutes observed in the control condition. There were 627 ELs in the treatment condition, and 614 ELs in the control condition included in this study. Thirty-eight bilingual teachers from 20 schools in the treatment condition and 39 bilingual teachers from 21 schools in the control condition participated in the study.

Table 13
Descriptive Statistics of Participants and Observation Time in Treatment and Control Conditions

	Coded Instructional Time (Minutes)	Schools	Teachers	Students
Treatment	2,280	20	38	627
Control	2,340	21	39	614

Table 14
t-test Results Comparing Treatment and Control Condition on Instructional Time Allocated in Quality CCP Learning

Condition	n	Mean	SD	*t*	*df*	*p*
Treatment	38	9.20%	0.05	5.417	75	<0.001
Control	39	3%	0.05			

To answer the first research question, an independent-sample t-test was conducted to determine the differences in the instructional time allocated in quality CCP learning. The results are demonstrated in Table 14. There were 38 treatment teachers and 39 control teachers included in the analysis. The teachers in treatment conditions were observed to use the CCP strategy more often during the dense cognitive instruction time and encourage ELs to respond in their second language, English (M = 9.2%, SD = 0.05) than the teachers in the control condition (M = 3%, SD = 0.05) with a statistically significant difference of 6.22% (95% CI, 3.93% to 8.51%), t (75) = 5.417, $p < 0.001$, Cohen's d = 1.23.

Before answering research questions 2-4, I examined the baseline equivalence of the students' achievement in reading, oral expression, and oral reading fluency between treatment and control conditions. A two-level HLM model was adopted with pretest scores as the dependent variable and the condition (treatment vs. control) as the level 2 predictor. As demonstrated in Table 15, there were 626 treatment students and 572 control students participated in reading fluency test as measured by DIBELS. The mean scores for the treatment students were 16.9 points (*SD* = 19.0) in the pretest and 46 points (*SD* = 29.4) in the posttest. The mean scores for the students in the control condition were 18.2 points (*SD* = 19.1) in the pretest and 46.5 points (*SD* = 29.0) in the posttest. There were 626 treatment students and 572 control students participated in reading comprehension test as measured by WMLS-R. The mean scores for the treatment students were 90.9 points (*SD* = 20.9) in the pretest and 106.2 points (*SD* = 18.0) in the posttest. The mean scores for the students in the control condition were 90.5 points (*SD* = 21.2) in the pretest and 106.4 points (*SD* = 17.0) in the posttest. There were 626 treatment students and 572 control students participated in oral expression test as measured by WMLS-R. The mean scores for the treatment students were 90.9 points (*SD* = 20.9) in the pretest and 106.2 points (*SD* = 18.0) in the posttest. The mean scores for the students in the control condition were 90.5 points (*SD* = 21.2) in the pretest and 106.4 points (*SD* = 17.0) in the posttest. The results in Table 16 indicated that there was no statistically significant difference of students' pre-testing achievement in English oral expression (p = .58, Cohen's d = 0.18), reading comprehension (p = .94, Cohen's d = 0.02), and reading fluency (p = .62, Cohen's d = 0.15).

Table 15
Descriptive Statistics of Grade 1 ELs' Achievement in Reading Fluency, Reading Comprehension and Oral Expression

		$Mean_{pre}$	SD_{pre}	N_{pre}	$Mean_{post}$	SD_{post}	N_{post}
Oral Reading Fluency	Treatment	16.9	19.0	626	46.0	29.4	572
	Control	18.2	19.1	611	46.5	29.0	563
Reading Comprehension	Treatment	90.9	20.9	626	106.2	18.0	572
	Control	90.5	21.2	612	106.4	17.0	563
Oral Expression	Treatment	59.2	26.7	626	74.8	23.3	572
	Control	57.1	26.0	612	70.4	22.7	563

Table 16
Baseline Equivalence of Pretest Achievement in Reading Fluency, Reading Comprehension, and Oral Expression

	Predictor	Estimate	S.E.	*df*	*t*	*p*	*d*
Reading Fluency Pre	Treatment	-1.1	2.1	42	-0.51	0.62	0.15
Reading Comprehension Pre	Treatment	0.2	2.2	73	0.08	0.94	0.02
Oral Expression Pre	Treatment	2.0	3.5	40	0.57	0.58	0.18

Fully Unconditional Model

The fully unconditional model addressed the second research question, whether the students' reading comprehension, reading fluency, and oral expression significantly varied across classrooms and schools. Table 17 demonstrated the fixed and random effects of the fully unconditional model in three separate circumstances: reading comprehension, oral expression, and reading fluency. The fixed effect was the grand mean of reading comprehension and oral expression as measured by WMLS-R and reading fluency as measured by DIBELS. The estimate for the grand mean was 106.5 in reading comprehension, 74.2 in oral expression, and 46.1 in reading fluency. The full maximum likelihood estimates of the variance components for reading comprehension was 247.9 at the student level, 58.9 ($p < 0.001$) at the teacher level, and 0 at the

school level. The full maximum likelihood estimates of the variance components for oral expression was 429 at the student level, 42 ($p < 0.05$) at the teacher level, and 65.2 ($p < 0.01$) at the school level. The full maximum likelihood estimates of the variance components for reading fluency was 665.4 at the student level, 176.9 ($p < 0.001$) at the teacher level, and 16.7 ($p = 0.345$) at the school level. The random effects pertaining to the teacher and school level means for oral expression were both statistically significant. However, the school level random effect was not significant in reading comprehension and reading fluency.

Moreover, it is worth noting that the amount of variation in reading comprehension, oral expression, and reading fluency between teachers and schools was quite small. The intraclass correlation coefficient (ICC) was .192 at the teacher level and 0 at the school level in reading comprehension. In the oral expression model, ICC was .078 at the teacher level and .121 at the school level. In the reading fluency model, ICC was .20 at teacher level and .019 at the school level. Due to the small amount of variance found in reading comprehension and reading fluency at the school level, I suggest that school factors had a small influence on ELs' reading comprehension and reading fluency. For oral expression, the bigger variance was found at the school level, indicating that school factors had a bigger influence on ELs' oral expression development.

Table 17

Null Model to Examine the Variance at School Level and Teacher Level

WMLS-R Reading Comprehension					
Fixed Effect	Coefficients	SE	*df*	*t*	*p*
WMLS-R Reading Mean (η_{001})	106.5	1	75	105.6	<.0001
Random Effects	Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Reading (r_{ijk})	247.9	10.8		23	<.0001
Random Effect for Class Mean for WMLS-R Reading (u_{0jk})	58.9	12.5		4.7	<.0001
Random Effect for School Mean for WMLS-R Reading (v_{00k})	0				
WMLS-R Oral Expression					
Fixed Effect	Coefficients	SE	*df*	*t*	*p*
WMLS-R Oral Expression Mean (η_{001})	74.2	1.6	39	46.1	<.0001
Random Effects	Coefficients	SE		*z*	*p*
Random Effect for Oral Expression (r_{ijk})	429	18.7		23	<.00
Random Effect for Class Mean for Oral Expression (u_{0jk})	42	18.7		2.3	0.0122
Random Effect for School Mean for Oral Expression (v_{00k})	65.2	26.3		2.5	0.0065
Reading Fluency					
Fixed Effect	Coefficients	SE	*df*	*t*	*p*
Reading Fluency Mean (η_{001})	46.1	1.8	41	25	<.0001
Random Effects	Coefficients	SE		*z*	*p*
Random Effect for Reading Fluency (r_{ijk})	665.4	29		23	<.0001
Random Effect for Class Mean for Reading Fluency (u_{0jk})	176.9	56.3		3.1	0.0008
Random Effect for School Mean for Reading Fluency (v_{00k})	16.7	41.9		0.4	0.345

*p < .05. **p < .01. ***p < .001.

Three-level Pretest Covariate and Conditional HLM Model

The three-level pretest covariate and condition HLM model addressed the third research question: did ELs' English oral expression, reading comprehension, and reading fluency differ between the treatment (as a result of instructional intervention including VPD) and the control condition, controlling for ELs' initial achievement on these measures? This model included ELs' centered pretest achievement at the first level and the condition that schools were randomly assigned at the third level. Table 18 provides the means of reading comprehension, oral expression, and reading fluency that were used for grand centering. The fixed and random effects for this conditional model are listed in Table 19. The estimated grand mean of reading

comprehension post-test was 106.2 ($p < .0001$). The full maximum likelihood estimates of the variance components for reading comprehension were141.6 ($p < .0001$) at the student level, 9.3 ($p < .05$) at teacher level, and 5.7 at the school level. The random effect pertaining to the teacher mean for reading comprehension remained statistically significant. Adding condition as a predictor at schools increased the estimated of variance at school level from 0 to 5.7.

The estimated grand mean of oral expression post-test was 71.2 ($p < .0001$). The full maximum likelihood estimates of the variance components for oral expression were151.6 ($p < .0001$) at the student level, 4.8 at teacher level, and 5.6 at the school level. The random effects pertaining to the teacher and school means for oral expressional were not statistically significant, which meant by adding ELs pretest of oral expression and the condition into the model, the school mean of ELs' oral expression did not significantly vary across schools. The estimated grand mean of reading fluency post-test was 106.2 ($p < .0001$). The full maximum likelihood estimates of the variance components for reading fluency were 254 ($p < .0001$) at the student level, 36.8 ($p < .01$) at teacher level, and 10.5 at the school level. The random effect pertaining to the teacher means for reading fluency remained statistically significant. Adding condition as a predictor at schools decreased the estimated of variance at school level from 16.7 to 10.5.

This pre-test covariate and condition three-level condition model involved ELs' pretests of reading comprehension, oral expression, and reading fluency in three separate instances. In three instances ELs' pretesting achievement statistically significantly predicted their post-testing achievement (*ps* < .0001). However, condition, the school level predictor varied in different instances. The condition was not a statistically significant predictor in the model of reading comprehension ($p = 0.65$) and the model of reading fluency ($p = 0.73$), which meant there was no statistically significant difference between students in the treatment and the control conditions

regarding their improvement in reading comprehension and fluency. However, the condition was a statistically significant predictor in the model of oral expression (p = .045, Cohen's *d* = 0.64). When ELs' pre-testing oral expression achievement was controlled for, the ELs in the treatment condition over-performed their peers in the control condition by 2.4 points in the post-test.

Table 18
Pretest Mean Used for Grand Centering

Measures	Mean
Reading Comprehension pre	90.7
Oral Expression pre	58.2
Reading Fluency	17.5

Table 19

The Results of Three-level Pretest Covariate and Conditional HLM Model

WMLS-R Reading Comprehension						
Fixed Effect		Coefficients	SE	*df*	*t*	*p*
WMLS-R Reading Comprehension Mean (η_{001})		106.2	0.9	42	118.1	<.0001
WMLS-R Reading Pre (η_{100})		0.6	0	1103	31	<.0001
Condition (η_{001})	Treatment	-0.6	1.3	40	-0.5	0.6453
	Control	0	.	.	.	.
Random Effects		Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Reading Comprehension (r_{ijk})		141.6	6.2		23	<.0001
Random Effect for Class Mean for WMLS-R Reading Comprehension (u_{0jk})		9.3	4.8		2	0.0258
Random Effect for School Mean for WMLS-R Reading Comprehension (v_{00k})		5.7	4.4		1.3	0.0966
WMLS-R Oral Expression						
Fixed Effect		Coefficients	SE	*df*	*t*	*p*
WMLS-R Oral Expression Mean (η_{001})		71.2	0.8	45	85.3	<.0001
WMLS-R Oral Expression Pre (η_{100})		0.7	0	1062	48.1	<.0001
Condition (η_{001})	Treatment	2.4	1.2	43	2.1	0.045
	Control	0	.	.	.	.
Random Effects		Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Oral Expression (r_{ijk})		151.6	6.6		23	<.0001
Random Effect for Class Mean for WMLS-R Oral Expression (u_{0jk})		4.8	3.5		1.4	0.0882
Random Effect for School Mean for WMLS-R Oral Expression (v_{00k})		5.6	3.4		1.6	0.0509
WMLS-R Reading Fluency						
Fixed Effect		Coefficients	SE	*df*	*t*	*p*
WMLS-R Reading Fluency Mean (η_{001})		45.6	1.4	44	32.1	<.0001
WMLS-R Reading Fluency Pre (η_{100})		1.2	0	1130	43.7	<.0001
Condition (η_{001})	Treatment	0.7	2	42	0.3	0.7346
	Control	0	.	.	.	.
Random Effects		Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Reading Fluency (r_{ijk})		254	11.1		22.9	<.0001
Random Effect for Class Mean for WMLS-R Reading Fluency (u_{0jk})		36.8	13.8		2.7	0.0039
Random Effect for School Mean for WMLS-R Reading Fluency(v_{00k})		10.5	10.6		1	0.1615

*p < .05. **p < .01. ***p < .001.

Three-level Pretest Covariate, Condition and The CCP Strategy HLM Model

This model addressed the fourth research question: what was the effective dosage of time allocation in the CCP_Dense_SL2 instructional period that contributed to an improvement of

Grade 1 ELs' reading comprehension, reading fluency, and oral expression? Table 20 provided the fixed and random effect for this condition model. Note that this model was comprehensive and contained all the predictors: pre-testing achievement at the student level, condition at the school level, and the quality CCP instruction (CCP_Dense_SL2) at the teacher level. Moreover, CCP_Dense_SL2 instruction time was re-coded by four quartiles as an ordinal variable: level 1 (0% - 3.1%), level 2 (3.1% - 6.1%), level 3 (6.1% - 10.6%), and level 4 (10.6% -22.8%) and was included in the model as a categorical predictor, which enabled me to find the effective dosage of CCP_Dense_SL2 instruction that could improve ELs' reading comprehension, oral expression, and reading fluency. The estimated grand mean of reading comprehension post-test was 105.6 ($p < .0001$). The full maximum likelihood estimates of the variance components for reading comprehension were 141.7 ($p < .0001$) at the student level, 9.4 ($p < .05$) at teacher level, and 4.1 at the school level. The random effect pertaining to the teacher means was statistically significantly different, while the random effect of school was not statistically significant, which means by adding the time allocation in the CCP_Dense_SL2 instruction as a level 2 predictor, teacher means of ELs' pretesting achievement of reading comprehension significantly varied across teachers.

The estimated grand mean of oral expression post-test was 70.7 ($p < .0001$). The full maximum likelihood estimates of the variance components for oral expression were 151.6 ($p < .0001$) at the student level, 3.6 at teacher level and 5.6 at the school level. The random effect pertaining to the teacher means and school means was not statistically significantly different, which means by adding the time allocation in the CCP_Dense_SL2 instruction as level 2 predictor, teacher means and school means of ELs' pretesting achievement of oral expression did not vary across teachers and schools. The estimated grand mean of oral expression post-test was

45 ($p < .0001$). The full maximum likelihood estimates of the variance components for oral reading fluency were 254.2 ($p < .0001$) at the student level, 28.8 ($p < .01$) at teacher level, and 13.3 at the school level. The random effect pertaining to the teacher means was statistically significantly different, which means teacher mean of ELs' pretesting achievement of reading fluency remained statistically significant across teachers after adding CCP_Dense_SL2 instruction as level 2 predictor. Moreover, the random effect of school remained insignificant across schools after the time allocation in the CCP_Dense_SL2 instruction was added to the model as a level 2 predictor.

This three-level condition model involved condition as a school level predictor, CCP_Dense_SL2 instruction time as a teacher level predictor, and ELs' pretesting achievement as a student level predictor. The model was replicated in three separate instances: reading comprehension, oral expression, and reading fluency. In three instances, ELs' pretesting achievement statistically significantly predicted their post-testing achievement (*ps* < .001). However, the condition was not a statistically significant predictor in any of three instances, which means there was no statistically significant difference between students in the treatment and the control conditions regarding their improvement in reading comprehension, oral expression, and reading fluency when the CCP_Dense_SL2 instruction time was included in the model.

CCP_Dense_SL2 instruction time which was re-coded by four quartiles as an ordinal variable. It was included in the model as a teacher level predictor, and level 1 was the reference level. In the model of reading comprehension, when the condition and ELs' pretesting achievement were controlled for, as compared with CCP_Dense_SL2 level 1, other levels of CCP_Dense_SL2 could not improve ELs' post-testing achievement in reading comprehension in

a statistically significant way. It is worth noting that when teachers allocated instructional time to the CCP_Dense_SL2 above 10.6% (level 4), there was a marginal significance of improvement. Compared with the students who nested in the classrooms where the teachers' CCP_Dense_SL2 instructional time was less than 3.1% of instructional time, the students who nested in the classroom where the teachers' CCP_Dense_SL2 instructional time was 3.1% - 6.1%, 6.1% - 10.6%, or 10.6% - 22.8% of the total instructional time, could obtain 1.7, 1.6 or 3.6 ($p = 0.053$, Cohen's $d = 0.50$) points more, respectively.

In the model of oral expression, when the condition and ELs' pretesting achievement were controlled for, as compared with CCP_Dense_SL2 level 1, there was no statistically significant difference among level 1, 2 and 3 regarding their impact on ELs' improving in oral expression. Moreover, if CCP_Dense_SL2 instructional time was at level 3 or under, CCP_Dense_SL2 could not statistically significantly improve ELs' oral expression. However, when the teachers increased the time allocation in the CCP_Dense_SL2 instruction to level 4 (10.6% - 22.8%), CCP_Dense_SL2 leaning statistically significantly improved ELs' oral expression ($p < .05$, Cohen's $d = 0.49$). Compared with the students who nested in the classrooms where the teachers' CCP_Dense_SL2 instructional time was less than 3.1% of instructional time, the students who nested in the classroom where the teachers' CCP_Dense_SL2 instructional time was 3.1% - 6.1%, 6.1% - 10.6%, or 10.6% - 22.8% of the total instructional time, could obtain 1.5, 0.8 or 3.3 points more, respectively.

In the model of reading fluency, when the condition and ELs' pretesting achievement were controlled for, as compared with CCP_Dense_SL2 level 1, there was no statistically significant difference among level 1, 2 and 3 regarding their impact on ELs' improving in reading fluency. However, when the teachers increase the time allocation in the

CCP_Dense_SL2 instruction to level 4 (10.6% - 22.8%), the CCP_Dense_SL2 instruction statistically significantly improved ELs' reading fluency ($p < .05$, Cohen's $d = 0.6$). Compared with the students who nested in the classrooms where the teachers' CCP_Dense_SL2 instructional time is less than 3.1% of instructional time, the students who nested in the classroom where the teachers' CCP_Dense_SL2 instructional time was 3.1% - 6.1%, 6.1% - 10.6%, or 10.6% - 22.8% of the total instructional time, could obtain 2.2, 0.4 or 6.6 more points, respectively.

Table 20

Three-level Pretest Covariate, Condition, and the CCP Strategy HLM Model

WMLS-R Reading Comprehension						
Fixed Effect		Coefficients	SE	*df*	*t*	*p*
WMLS-R Reading Comprehension Mean (η_{001})		105.6	1.0	49.	108.5	<.001
WMLS-R Reading Pre (η_{100})		0.6	0.0	1103.	31.0	<.001
Condition (η_{001})	Treatment	-2.2	1.5	58	-1.5	0.1494
	Control	0.0	.	.	.	.
CCP_Dense_SL2	Level 2	1.7	1.9	76	0.9	0.3767
	Level 3	1.6	1.6	61	1.0	0.3267
	Level 4	3.6	1.8	65	2.0	0.0527
	Level 1	0	.	.	.	.
Random Effects		Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Reading Comprehension (r_{ijk})		141.7	6.2		23.0	<.001
Random Effect for Class Mean for WMLS-R Reading Comprehension (u_{0jk})		9.4	5.0		1.9	0.0294
Random Effect for School Mean for WMLS-R Reading Comprehension (v_{00k})		4.1	4.4		0.9	0.1757
WMLS-R Oral Expression						
Fixed Effect		Coefficients	SE	*df*	*t*	*p*
WMLS-R Oral Expression Mean (η_{001})		70.7	0.9	54	76.8	<.001
WMLS-R Oral Expression Pre (η_{100})		0.7	0.0	1059	48.4	<.001
Condition (η_{001})	Treatment	1.0	1.4	61	0.7	0.5
	Control	0.0	.	.	.	.
CCP_Dense_SL2 _level 2	Level 2	1.5	1.7	75	0.9	0.4
CCP_Dense_SL2 _level 3	Level 3	0.8	1.4	59	0.5	0.6
CCP_Dense_SL2 _level 4	Level 4	3.3	1.6	66	2.0	0.0
	Level 1	0.0	.	.	.	.
Random Effects		Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Oral Expression (r_{ijk})		151.6	6.6		23.0	<.001
Random Effect for Class Mean for WMLS-R Oral Expression (u_{0jk})		3.6	3.3		1.1	0.1333
Random Effect for School Mean for WMLS-R Oral Expression (v_{00k})		5.6	3.2		1.7	0.041
WMLS-R Reading Fluency						
Fixed Effect		Coefficients	SE	*df*	*t*	*p*
WMLS-R Reading Fluency Mean (η_{001})		45.0	1.6	52	28.6	<.001
WMLS-R Reading Fluency Pre (η_{100})		1.2	0.0	1126	43.9	<.001
Condition (η_{001})	Treatment	-1.8	2.4	58	-0.7	0.4667
	Control	0.0	.	.	.	.
CCP_Dense_SL2 _level 2	Level 2	2.2	3.0	67	0.7	0.4717
CCP_Dense_SL2 _level 3	Level 3	0.4	2.6	55	0.2	0.8627
CCP_Dense_SL2 _level 4	Level 4	6.6	2.9	57	2.3	0.0282
	Level 1	0.0	.	.	.	.
Random Effects		Coefficients	SE		*z*	*p*
Random Effect for WMLS-R Reading Fluency (r_{ijk})		254.2	11.1		22.9	<.001
Random Effect for Class Mean for WMLS-R Reading Fluency (u_{0jk})		28.8	12.0		2.4	0.008
Random Effect for School Mean for WMLS-R Reading Fluency(v_{00k})		13.3	9.9		1.3	0.0894

Discussion

The goals of this chapter were to (a) examine the difference between the treatment and the control teachers regarding their time allocation for quality CCP learning, (b) examine the variances at classroom level and school level, (c) investigate the impact of the curriculum-based VPD on ELs' improvement in reading comprehension, oral reading fluency, and oral expression, and (d) determine the effective dosage of the quality CCP learning that improved ELs' achievement in reading comprehension, oral reading fluency, and oral expression. An independent t-test and four three-level HLM models were used to address these question. Five important findings emerged.

Firstly, the quality CCP instructional time was documented via the low-inference multidimensional classroom observation protocol TBOP by examining the overlapping time across three domains—ESL strategies, student language, and language content. The treatment teachers were observed to allocate significantly more time to quality CCP instruction than the control teachers did, which meant the teachers in the treatment conditions used the CCP strategy more often in the dense cognitive content and provided more opportunities for ELs to respond in English. This finding was aligned with the curriculum-based VPD provided by the Project ELLA-V scaffolding of treatment teachers' quality CCP instruction. Supported by ongoing VPD, the treatment teachers purposefully implemented the scripted CCP activities, which encouraged ELs to participate in the higher-order thinking, think-pair-share interactivities. This finding also aligned with the VPD solar system proposed in this study that VPD stimulated teachers' cognitive development and quality instruction delivery in the EL classrooms.

Before investigating the impact of VPD and quality CCP learning on ELs' reading comprehension, oral expression, and oral reading fluency, I examined whether the baseline

equivalence between the treatment and the control conditions was established in three circumstances: oral expression, reading comprehension, and oral reading fluency. To be specific, based on the WWC standard for baseline equivalence (absolute effect size ≤0.05, satisfies baseline equivalence; 0.05 < absolute value of effect size ≤ 0.25, statistical adjustment required to satisfy baseline equivalence; absolute value of effect size > 0.25, not satisfy baseline equivalence), reading comprehension satisfied baseline equivalence while oral reading fluency and oral expression needed to conducted statistical adjustment by controlling pretest achievement.

The second finding in this chapter addressed the second research question. Through the results from the fully unconditional model, I found that (a) ELs' reading comprehension and reading fluency significantly varied across teachers/classrooms, and (b) oral expression of ELs significant varied across teachers/classrooms and schools. Through this finding, I confirmed that multi-level analysis was an appropriate method for this study.

The third finding in this chapter answered the third question of whether the ongoing, structured, and curriculum-based VPD positively impact Grade 1 ELs' oral expression, reading comprehension, and oral reading fluency achievement. Although through the third finding, I found that a two-level HLM might better fit for reading comprehension and oral reading fluency, the three-level HLM model was still adopted in the pretest covariate model considering randomization was conducted at school level. The ongoing, structured, curriculum-based VPD significantly, better supported Grade 1 ELs' oral expression. No difference was identified in reading comprehension and oral reading fluency based on the results of the ongoing, structured, curriculum-based VPD. Considering the focus of student level intervention was higher-order thinking and oral language developed by providing ELs with meaningful opportunities to engage

in English language learning activities and teacher level intervention focusing on cooperative questioning, think-pair-share, and share-reading, I found that the model of pretest covariate and condition confirmed the positivity effect of the ongoing, structured, curriculum-based VPD. This finding was also consistent with the previous study conducted by Tong et al. (2008) that the intervention that supported teachers with quality instruction also supported ELs' oral language development, especially at early grades. It was further validated that the ongoing, structured, curriculum-based VPD support ELs' English oral language development had the same effect as the face-to-face PD that was delivered in the previous project ELLA. Finally, the VPD solar system proposed in this study was used to show that quality instruction supported by the ongoing, structured, curriculum-based VPD facilitated ELs' oral language development.

The fourth finding was to address the last research question. From the results drawn from the full model, I identified a statistically significant effective dosage of quality CCP instruction time in two separate models of oral expression and reading comprehension. In both cases, the quality CCP instruction time needs to exceed 10.6% of the instructional time to improve ELs' oral expression and reading fluency significantly effectively. Although no statistically significantly effective dosage of quality CCP was identified for improving ELs' reading comprehension, the marginal significant difference with medium effect size was found such that when the teachers spent more than 10.6% of instructional time in the quality CCP activities, ELs' reading comprehension was improved. This finding filled the gap in the literature as pointed by Bowman-Perrott et al. (2016) and Pyle et al. (2017). In their systematic reviews, they claimed that the effective dosage of the quality CCP learning should be examined. To be specific, what dosage of quality CCP learning is strong enough to support ELs' reading comprehension, oral reading fluency, and oral expression was addressed in this study. Through my research, I

validated the VPD solar system that quality instructional time embedded with an effective dosage of CCP learning stimulated ELs' cognitive development as well as their language proficiency. Additionally, I validated that the multidimensional classroom observation instrument TBOP that was embedded in the VPD solar system could accurately record the quality instruction and the quality CCP learning activities in the EL classrooms.

In this full model that addressed that last research question, condition (treatment vs. control) became a non-significant predictor in three instances including reading comprehension, oral reading fluency, and oral expression after the quality CCP learning was added into the model. Therefore, teachers' quality CCP instruction played a crucial role in ELs' development of reading comprehension, oral reading fluency, and oral expression. Connecting the first finding of this study, I ascertained that the curriculum-based VPD combined with research-based curriculum improved teachers' instructional quality of the CCP strategy and further improved ELs' language proficiency.

Limitation and Recommendations

After the thorough examination on the impact of the curriculum-based VPD and the quality CCP learning on ELs' improvement in English oral expression, reading comprehension, and oral reading fluency with four sophisticated HLM models via large-scale datasets from an RCT study, there were not many limitations in this study. The only limitation was that I did not identify the higher bound of the effective dosage of the quality CCP learning. Due to the fact that the highest quality CCP time was up to 22.8% as observed via TBOP in the participating classrooms, the higher bound of effective dosage was not identified. It cannot be assumed that there would be a direct correlation between the percent of the CCP strategy applied in the classroom and the positive impact on ELs' improvement. Project ELLA-V embedded various

ESL strategies in the curriculum and the VPD training, and the CCP strategy was one of those strategies. In future studies, researchers should conduct intervention focusing on the CCP learning in ELs. With a higher percent of this strategy being observed, the higher bound of the effective dosage might be observed. Moreover, future studies should replicate the fourth HLM and involve different ESL strategies, by which the effective dosage of other ESL strategies can be identified. If the effective dosage of all ESL strategies can be identified, the research-based curriculum and the curriculum-based VPD can be modified and improved, which will bring significant benefits to the EL education field.

For large-scale intervention studies, the ongoing, structured, curriculum-based VPD seemed like an effective method to provide the accessible and intervention-related PD resources to teachers. Furthermore, the multidimensional classroom observation instrument TBOP has been found to be a reliable and comprehensive instrument for documenting instructional practice, examining the quality of instruction, and evaluating the effectiveness of VPD on teachers. The VPD solar system can also be applied in future studies to examine the impact of the VPD on teachers' instructional practices as well as teacher-student interaction.

Conclusion

Project ELLA-V teachers received intensive, structured, and ongoing, curriculum-based virtual professional development. With the support of VPD, these teachers implemented a research-based curriculum trained at improving ELs' English language proficiency. Via a multidimensional, low-inference classroom observation protocol—TBOP, it was observed that the teachers in the treatment condition allocated more time to the CCP strategy during the dense cognitive content instruction and encouraged ELs to respond in English, all of which were considered as an effective instructional practice/strategy for ELs (August & Shanahan, 2006) and

a reflection of the quality of instruction (Tong, Luo, et al., 2017). During the intervention, the treatment teachers were following the outlined curriculum and the teaching objectives with an emphasis on teaching higher-order thinking content and encouraging students conducting think-pair-share, cooperative, and collaborative activities. Further, provided with the biweekly, structured, curriculum-based VPD, the treatment teachers were more aware of their pedagogical behavior, strategies, and students' interaction, which enhanced the consistency and quality of intervention implementation.

One of the most pertinent questions of this study was to what extent Project ELLA-V intervention impacted ELs' English language learning. In order to answer this question, the pretesting baseline equivalence between students in the treatment and the control conditions was examined. To be specific, the baseline equivalence of ELs' pre-testing achievement in reading comprehension, oral expression, and reading fluency was checked via a three-level hierarchical linear model. The baseline equivalence of three pre-testing achievements was established between the treatment students and the control students, which indicated no-difference of students' achievement between condition at the initial stage of intervention. Furthermore, the students' improvement of reading comprehension, oral expression, and reading fluency was examined between the treatment and the control conditions by controlling their pretesting achievement as suggested by WWC's Baseline Equivalence Brief (2015b).

During the intervention, the treatment students received the research-based curriculum intervention implemented by the treatment teachers who received the curriculum-based VPD. This intervention, embedded with the research-based curriculum and the curriculum-based VPD provided by the Project ELLA-V, statistically significantly improved ELs' English oral expression. However, there was no difference found in the students' progress of reading

comprehension and reading fluency between the conditions. In this study, I identified the effective dosage of the quality CCP learning—the higher level of dosage of quality CCP learning, which counted for 10.6% - 22.8% out of total instrumental time, improved students' reading comprehension, oral expression, and reading fluency.

Project ELLA-V and this study successfully proved that the research-based curriculum combined with an intensive, ongoing, curriculum-based VPD improved the teachers' instruction quality, which further improved their students' proficiency of oral expression, which included expressive vocabulary, language comprehension, and development. Moreover, the CCP strategy that the teachers implemented in the dense cognitive content encouraged ELs to respond to each other in English during their CCP learning improved ELs' English reading comprehension, oral expression, and reading fluency when the amount of this instruction exceed 10.6% of the total instructional time. Further, the multidimensional, low-inference classroom observation protocol TBOP can provide a rigorous analysis of what is happening in the EL classrooms, which seems to be an effective method in generating a representative set of reliable EL classroom observation data. As Lara-Alecio et al. (2009) noted, this flexible and comprehensive multidimensional classroom observation instrument can be used as an evaluation tool that provides a clear picture of teacher pedagogical practice and their interactions with ELs in different educational settings. It can also provide researchers and intervention implementers a good picture of how/whether the key intervention components are implemented. Last but not least, the VPD solar system proposed in this study was validated through the comparison of treatment and control teachers' quality instruction with the CCP strategy as well as the comparison between the treatment ELs' and the control ELs' improvement in English oral expression, reading comprehension, and oral reading fluency.

CHAPTER V

CONCLUSION: A SYNTHESIS

This dissertation consisted of five chapters oriented towards the discussion on the impact of the CCP strategy on ELs' English language proficiency and the impact of VPD on teachers' application of this strategy on ELs' English reading and oral expression proficiency. Chapter I was an overview of the dissertation. Chapter II was a research synthesis of the impact of the CCP strategy on ELs' English language learning, particularly in reading and speaking. In Chapter III, the archived large-scale RCT data recorded by a low-inference classroom observation instrument was adopted to validate the impact of the research-based curriculum and the curriculum-based VPD on the quality of teacher using the CCP strategy. In Chapter IV, I proposed a dynamic model of the research-based curriculum, the curriculum-based VPD, the teacher-student interaction, and the student-student CCP activities. The archived data sets were adopted from the large-scale RCT to validate the model, and further to identify the effective dosage of the CCP strategy in quality ESL instruction. In Chapter V, the findings from preceding chapters were synthesized to present how each chapter related to another.

The overview in Chapter I included the research background, the research questions, and potential journal for publication of Chapters II, III, and IV. The relationship between Chapter I and Chapter II was that Chapter I was a dissertation overview while Chapter II was a research synthesis of the impact of the CCP strategy on ELs' English reading and speaking proficiency. The relationship between first two chapters and Chapters III and IV was that Chapter III and IV addressed the issues that were specified in Chapter I and II. The relationship between Chapter V and the previous chapters is that Chapter V is the process of synthesis from Chapter I to Chapter IV.

In Chapter II, a research synthesis was conducted to identify the impact of the CCP strategy on ELs' English reading and speaking proficiency. In Chapter II, seven studies were fully examined regarding their sample characteristics, intervention features, design characteristics, and outcome characteristics. I found that the CCP learning, in general, helps elementary ELs' reading comprehension, reading fluency, and phonemic awareness. Additionally, I found that ongoing PD and coaching help teachers to improve the quality of implementing this strategy. Combined with the results drew from the previous reviews, the quality of implementing the CCP strategy and actual dosage of this strategy being implemented in the classroom, and to what extent the quality CCP strategy improve ELs' English reading and speaking should be the focus of future research. Based on the findings in Chapter II, the quality of teachers' application of the CCP strategy became the focus of Chapter III. Examining how ongoing VPD impacted ELs' development of reading and oral language and what dosage of the quality CCP learning should be effective became the focus of Chapter IV. Chapters III, IV, and V then drew from the literature and the findings from Chapter II to build a theoretic framework.

In Chapter III, I used the data from Project ELLA-V. Via the low-inference classroom observation instrument TBOP, the pedagogies of the treatment and control teachers were observed and recorded. In Chapter III I found that the research-based curriculum and the curriculum-based VPD had a substantial impact on teachers' instruction and application of the CCP strategy. Chapter III and Chapter IV are related as the impact of VPD on the teachers' pedagogy found in Chapter III was further explored and extended to ELs' achievement, discussed in Chapter IV. In Chapter V, I conclude Chapter III by considering what it adds to the research.

In Chapter IV, I proposed the VPD solar system to validate the findings in Chapter II and Chapter III. In Chapter IV, the dynamics of the research-based curriculum, the curriculum-based VPD, teachers' application of the CCP strategy, and ELs' achievement in reading and speaking was discussed in terms of causal reasoning to rationalize for a causal effect of VPD on teachers and students as validated in Project ELLA-V. The findings in Chapter IV are synthesized and added in Chapter V.

In Chapter V, I draw and synthesize the findings from all other chapters. Chapter V works as a summary to illustrate the process and logic of this dissertation and brings all information together in a meaningful way. Moreover, in Chapter V, after summarizing the findings of each chapter, I also provide recommendations and suggestions.

Summary of Study Significance

Following a rigorous search procedure on five databases, I found seven studies that met the inclusive and exclusive criterion and included them in the research synthesis (referring to Chapter II). After analyzing seven studies' sample characteristics, intervention features, design characteristics, and outcome characteristics, I found the CCP strategy can support ELs to develop their phoneme awareness, letter-word identification, oral language, oral reading fluency, and reading comprehension. In this finding, I extended the previous reviews which focused more on academic achievement and further examined the impact of the CCP strategy on ELs' language proficiency—reading and speaking. The importance of professional development are also emphasized in the results. Not only can the initial training can help teachers get familiar with the intervention, but the ongoing PD can also better support teachers with higher quality implementation and students with better instruction. Moreover, I found that the total intervention hours might relate to the effectiveness of implementation, which was also consistent with the

previous reviews. However, quantity does not mean the quality. After this synthesis, I suggested that future researchers should examine the actual time allocation that the teacher implemented the CCP strategy and the quality of ELs' CCP learning activity. Additionally, the conclusions in Chapter II led to two questions—how ongoing professional development supports teachers to implement the CCP strategy with quality instruction and how the quality implementation of this strategy support ELs' English language proficiency.

Through an examination of archived classroom observation data (referring to Chapter III), I suggest using Project ELLA-V as a validation intervention for empowering teachers and improving ELs' English language proficiency. This suggestion stems from the comparison between the treatment and the control teachers' pedagogy. In Chapter III, I demonstrated how the VPD impacted teachers' application of the CCP strategy regarding each dimension of pedagogy (i.e., teachers' language, students' language, language content, language mode, and instruction activity). Moreover, the teachers' instruction practice within both treatment and control classrooms was comprehensively documented by using TBOP, the low-inference classroom observation instrument. The most visible difference between the teachers in the treatment and the control conditions was the total percent of the instructional time that was allocated in the CCP strategy. Furthermore, in Chapter III, a multi-layer cross-tabulation analysis was adopted to check the differences between the treatment and the control teachers regarding their practice in the pedagogical domains (i.e., Language Mode, Activity Structure, and Language Content) layering across the domain of Language Content when the CCP strategy was implemented. I found that that the treatment teachers' instructional languages and activities significantly differed from the control teachers' when light and dense cognitive contents were delivered (the Domain of Language Content). I also found that the treatment students' responding languages and

communication modes signficantly differed from the control students' when the teachers were using the CCP strategy in light and dense cogntive contents.

As mentioned in the previous chapters, Project ELLA-V was a validation study of Project ELLA with the crucial purpose of validating the effectiveness of ELLA PD components delivered virtually. In Chapter III, I reinforced and validated the possible impact Project ELLA PD components had in improving teachers' pedagogy and instructional practice in the EL classrooms. Additionally, through a multi-layer analysis, I was able to see a "zoom-in" shot of teachers' pedagogical practice during 20 seconds. This micro-analysis of ESL instruction in EL classrooms, aligned with the analytic method of the studies conducted by Lara-Alecio et al. (2009) and Bruce et al. (1997), further added to the conversation concerning analytic methods of observation data collected via a multidimensional classroom observation instrument. Additionally, in Chapter III, I questioned whether teachers' pedagogical differences in their application of the CCP strategy impacted ELs' English reading and oral proficiency, as it did for VPD.

In Chapter IV, I proposed a model of "VPD solar system of cognitive development of ELs and their teachers" to examine the dynamics in EL classrooms regarding research-based curriculum and the curriculum-based VPD, focusing on the cognitive development of both ELs and their teachers. In addition, via the classroom observation data and the ELs' English reading and speaking achievement data collected in Project ELLA-V, the solar VPD system was used to validate the effectiveness of VPD on ELs' English reading and speaking proficiency and then was used to identify the effective dosage of the CCP strategy on ELs' improvement in English reading and speaking proficiency. The rationale for building this system in Chapter IV was to serve a means for causal implication for future intervention with the attention of improving

teachers' cognitive development and instruction quality via curriculum and professional development and by proxy to develop students' cognitive development. By validating the impact of VPD in the VPD solar system, the face-to-face delivered PD components of Project ELLA, delivered virtually in Project ELLA-V, were found to have the same effect to improve Grade 1 ELs' English oral proficiency. Moreover, in Chapter IV, the effective dosage of the quality CCP strategy employed in English dense cognitive content was identified for improving ELs' English reading fluency and oral expression. By validating the VPD solar system, the findings of the impact of the research-based curriculum and the curriculum-based VPD on ELs help confirm the effectiveness of Project ELLA-V to serve ELs.

Summary of Key Findings

In Chapter II, after thoroughly searching through the literature, I found that research considering the CCP strategy that was applied to support ELs' to develop their English reading and speaking proficiency was limited. In this dissertation, only seven studies were identified. Since in this synthesis I examined ELs' reading proficiency by subcategories like reading comprehension, oral reading fluency, letter-word identification, and phoneme awareness, and I examined speaking proficiency by subcategories like oral language and oral expression, only a limited number of studies (two to three) met the criteria for each subcategory of reading and speaking proficiency. In this research, I found that (a) the CCP strategy supported elementary school ELs' reading comprehension especially for the students under Grade 3 with medium to large effect size; (b) the CCP strategy supported ELs' development in phoneme awareness, letter-word identification, and word fluency with large effect size; (c) more studies are needed in the area of oral reading fluency as the findings were mixed; (d) ongoing PD and coaching can better support teachers' implementation of this strategy as well as students' achievement, and (e)

the total CCP intervention duration as well as the actual dosage of this strategy should be examined in future studies.

In Chapter III, via the multidimensional classroom observation instrument, TBOP, both treatment and control teachers' pedagogical practice were comprehensively recorded. Via a multi-layer cross-tabulation analysis, the observation data of teachers' pedagogy were micro-analyzed. Based on the analysis, I found the treatment teachers used the CCP strategy more often than the control teachers in general. More importantly, the treatment teachers displayed a focus on using this strategy in CALP time which was considered a crucial indicator of the quality of instruction. As compared with the teachers in the control condition, the treatment teachers allocated more instructional time to higher cognitive level content and provided more opportunities for ELs to involve in the CCP activities (e.g., think-pair-share and share-reading) which was aligned with the scripted, research-based curriculum and was also one of the focuses of the curriculum-based VPD.

I also found that the treatment teachers also displayed a higher quality of teaching as compared with the teachers in the control condition when the CCP strategy was implemented. For example, the treatment teachers used more instructional time in evaluating ELs' CCP activities and providing feedback. Students in the treatment condition were observed to engage more in speaking-listening and listening-speaking during the dense cognitive content instruction than those teachers who were not using the CCP strategy. The students in both conditions mirrored teachers' language. When the teachers used the majority of instructional time teaching in English, their students were observed to respond in English a majority of the time, which was consistent with the findings in the study of Lara-Alecio et al. (2009). Therefore, I could conclude that the curriculum-based VPD improved the quality of the CCP learning in EL classroom.

The effectiveness of the research-based curriculum and the curriculum-based VPD on ELs was also validated in this study (referring Chapter IV). Project ELLA-V was already considered to have a causal-effect impact on the participants due to its nature of being an RCT study. The ongoing curriculum-based VPD significantly improved the treatment teachers' quality of CCP instruction. Moreover, the VPD solar system proposed in the study was used to validate the two key components' (the research-based curriculum coupled with the curriculum-based VPD and effective dosage of the quality CCP learning) impact on ELs' English reading and speaking proficiency. I found that the research-based curriculum together with curriculum-based VPD significantly improved Grade 1 ELs' English oral expression. Furthermore, the VPD solar system was used to validate the effective dosage of the quality CCP learning in EL classrooms to improve ELs' reading comprehension, oral reading fluency, and oral expression. An important finding drawn from the model is that the teachers should allocate at least 10.6% of instructional time to quality CCP learning to improve ELs' English reading fluency and oral expression.

Limitation

In this dissertation, rigorous methods were used to identify the effectiveness of some components of Project ELLA-V, including the CCP strategy, teachers' pedagogy, the research-based curriculum, and the curriculum-based VPD. This research synthesis followed the research synthesis standard procedure proposed by Cooper (1982) to increase the rigor of this dissertation. Although the research synthesis procedure precipitated a comprehensive search, the findings that drew from the synthesis were limited to the quality of the primary studies.

Although I used the archived data from Project ELLA-V, the data were collected during the school year 2015-2016, which was only one year before this dissertation was conducted. Due to the nature of categorical data, the analysis in Chapter III only involved cross-tabulation and

frequency analysis. However, to minimize weakness related to this study, I instituted the following mechanisms (a) via the low-inference, multidimensional classroom observation instrument TBOP, teachers' pedagogy was rigorously analyzed; (b) inter-rater reliability was established and monitored throughout the TBOP coding procedure; (c) multi-layer cross-tabulation analyses provided a "zoom-in" shot of cross-dimensional analysis,, producing more detailed informatio, which was observed and analyzed, and (d) the sophisticated three-level HLM models were adopted to examine baseline equivalence and the impact of the VPD and the effective dosage of the quality CCP learning on ELs' reading, oral expression, and oral reading fluency.

Another limitation of this study was that I only investigated one of the ESL strategies --- the CCP strategy. There were eight other ESL strategies included in the curriculum and VPD components and observed via the instrument TBOP. These strategies were also believed to improve ELs' English language proficiency; however, they were not examined in this study. Furthermore, in this study, I identified the effective dosage of the quality CCP learning in an English instruction environment. However, only the lower bound of the effective dosage was identified for improving ELs' English reading fluency and oral expression. The higher bound of an effective dosage of the quality CCP learning to improve ELs' English language proficiency was not identified due to the fact that multiple ESL strategies were embedded in both ESL instruction and ELLA-V curriculum which shared the instructional time.

Implications and Recommendations

VPD, The CCP Strategy, and ELs' Reading and Speaking proficiency

Through my research synthesis, I found that not many studies were related to the CCP strategy, and even fewer provided the research-based curriculum or offered the curriculum-based

PD. Socio-cognitive theory should play an important role in professional development serving the teachers who work with ELs. EL teachers can be viewed as learners of the research-based curriculum and the curriculum-based VPD as well as deliverers of the academic instructional components of curriculum to the ELs.

One of the implications of this research is that the treatment teachers who received the curriculum-based VPD of Project ELLA-V demonstrated the critical components of ELLA-V curriculum in EL classrooms and better supported ELs' English language proficiency development. The treatment teachers showed the different pedagogy that was embedded in the curriculum and supported by the VPD. Although in this study I only examined one ESL strategy in the curriculum, Project ELLA-V was believed to effectively support other key components of the curriculum due to the high fidelity of implementation. Furthermore, future researchers can adopt the VPD solar system to identify the effectiveness of the research-based curriculum and the curriculum-based PD/VPD components on teachers as well as students, including ELs.

Although the authors of previous studies emphasized that the CCP strategy facilitates ELs in improving their English reading and speaking proficiency (Almaguer, 2005; Calhoon et al., 2007; McMaster et al., 2008; Saenz et al., 2005; Zhang et al., 2013), in this study, I found the quality CCP learning in English that involved higher-order thinking needed to exceed 10.6% of instructional time to be effective in improving ELs' reading comprehension, reading fluency, and oral expression. Since the upper-bound of the effective dosage of the quality CCP learning in EL classrooms has not been identified this study, I suggest that future researchers design an intervention focusing on the CCP strategy. In this case, the researchers might observe a higher percentage of time allocated for this strategy and further identify the effectiveness of higher dosage of the CCP strategy in EL classrooms.

Methodology of Conducting Analysis Involving Categorical Data and Multi-level Analysis

The multi-layer cross-tabulation was used to analyze the multidimensional data. To be specific, via the multi-layer cross-tabulation, researchers can select different combinations of dimensions of multidimensional data, which might fit the particular research purposes of future researchers dealing with this type of data or similar categorical data sets. For example, in this dissertation study, one of my focuses was to examine the differences between the treatment and the control teachers regarding their implementation of the CCP strategy during dense cognitive and light cognitive contents. Multi-layer analysis, therefore, involved three layers: condition, the domain ESL strategy, and the domain of language content. Via this multi-layer cross-tabulation analysis, I was able to zoom in to observe what percent of instructional time teachers were teaching the dense cognitive content when they were using the CCP strategy and further compare the differences between the treatment and the control teachers regarding this three-cross-dimension segment of instructional time.

In the present study, I also presented an evaluation of the treatment effect of Project ELLA-V. Since the students were nested in the classrooms and the teachers were nested in the schools, a three-level HLM model perfectly fits this type of data. Via the three-level HLM model, I identified the treatment effect of Project ELLA-V on improving ELs' English oral expression. Future researchers can adapt the models and the model building process presented in this study to fit their own research purposes when working with multi-level, clustered datasets. Moreover, the methodology that was used in Chapter IV to identify the effective dosage of the CCP learning can also be replicated to identify the effective dosage of other ESL strategies and even the teachers' pedagogical practice in the EL classrooms.

Classroom Observation Instruments

The multidimensional TBOP has been found to be a reliable and comprehensive classroom observation instrument when applied in ESL classrooms. For researchers and practitioners, it can be utilized to (a) observe teacher pedagogy in general, (b) examine teachers' fidelity of implementation, especially critical components of intervention, (c) examine the impact of teachers' pedagogy on students' achievement, and (d) provide PD feedback to teachers based on TBOP data analysis. Due to the nature of multidimensional analysis, the TBOP instrument can be flexibly applied to collecting classroom observation data by using different combinations of different domains to adapt to diverse research interests or focuses. Moreover, it is worth noting that it is crucial for researchers and practitioners to establish and monitor inter-rater reliability when TBOP or another classroom observation instrument is used to analyze observation data. In this way, the classroom observation data can be reliably analyzed, which leads to reliable and accurate results.

REFERENCES

Almaguer, I. (2005). Effects of dyad reading instruction on the reading achievement of Hispanic third-grade English language learners. *Bilingual Research Journal, 29*(3), 509-526.

Aloqaili, A. S. (2012). The relationship between reading comprehension and critical thinking: A theoretical study. *Journal of King Saud University-Languages and Translation*, *24*(1), 35-41.

Alvarado, R., Ruef, M. L., & Schrank, F. A. (2005). *Woodcock-Munoz Language Survey-Revised*. Itasca, IL: Riverside Publishing.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers*. Champaign, IL: University of Illinois.

Arreaga-Mayer, C. (1998). Increasing active student responding and improving academic performance through classwide peer tutoring. *Intervention in School and Clinic*, *34*(2), 89-94.

Arreaga-Mayer, C., & Perdomo-Rivera, C. (1996). Ecobehavioral analysis of instruction for at-risk language-minority students. *The Elementary School Journal*, *96*(3), 245-258.

August, D., McCardle, P., & Shanahan, T. (2014). Developing literacy in English language learners: Findings from a review of the experimental research. *School Psychology Review*, *43*(4), 490-498.

August, D., & Shanahan, T. (2006). Executive summary. *Developing literacy in second-language learners: A report of the national literacy panel on language minority children and youth.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, *27*(1), 10-20.

Baca, L., & de Valenzuela, J. S. (1998). Background and rationale for bilingual special education In L. Baca & H. Cervantes (Eds.), *The bilingual special education interface* (pp. 4-25). Upper Saddle River, NJ: Simon & Schuster.

Bardack, S. (2010). Common ELL terms and definitions. Washington DC: English Language Learner Centre, American Institutes for Research.

Barton, D. (2009). Understanding textual practices in a changing world. In M. Baynham & M. Prinsloo (Eds.), *The future of literacy studies* (pp. 38–53). Basingstoke, UK: Palgrave Macmillan.

Bates, M. S., Phalen, L., & Moran, C. (2016). Online professional development: A primer. *Phi Delta Kappan*, *97*(5), 70-73.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, *24*(2), 417-436.

Bowers, E., Fitts, S., Quirk, M., & Jung, W. (2010). Effective strategies for developing academic English: Professional development and teacher practices. *Bilingual Research Journal*, *33*(1), 95-110.

Bowman-Perrott, L., Mahadevan, L., & Etchells, M. (2016). Assessing the academic, social, and language production outcomes of English language learners engaged in peer tutoring: a systematic review. *Education and Treatment of Children*, *39*(3), 359-388.

Bruce, K. L., Lara-Alecio, R., Parker, R. I., Hasbrouck, J. E., Weaver, L., & Irby, B. (1997). Inside transitional bilingual classrooms: Accurately describing the language learning process. *Bilingual Research Journal*, *21*(2-3), 123-145.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111-114.

Burchinal, M., Field, S., López, M. L., Howes, C., & Pianta, R. (2012). Instruction in Spanish in pre-kindergarten classrooms and child outcomes for English language learners. *Early Childhood Research Quarterly*, *27*(2), 188-197.

Calderón, M., Hertz-Lazarowitz, R., & Slavin, R. (2000). Effects of bilingual cooperative integrated reading and composition on students making the transition from Spanish to English reading. *Elementary School Journal*, *99*(2), 153-165.

Calderón, M., Slavin, R., & Sánchez, M. (2011). Effective instruction for English learners. *The Future of Children*, *21*(1), 103-127.

Calhoon, M. B., Al Otaiba, S., Cihak, D., King, A., & Avalos, A. (2007). Effects of a peer-mediated program on reading skill acquisition for two-way bilingual first-grade classrooms. *Learning Disability Quarterly*, *30*(3), 169-184.

Carlisle, J. F., Beeman, M., Davis, L. H., & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics*, *20*(4), 459-478.

Cisco, B. K., & Padrón, Y. (2012). Investigating vocabulary and reading strategies with middle grades English language learners: A research synthesis. *RMLE Online*, *36*(4), 1-23.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*(1), 1-35.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hilsdale. NJ: Lawrence Earlbaum Associates.

Cole, M. W. (2013). Rompiendo el Silencio: Meta-analysis of the effectiveness of peer-mediated learning at improving language outcomes for ELLs. *Bilingual Research Journal*, *36*(2), 146-166.

Cole, M. W. (2014). Speaking to read: Meta-analysis of peer-mediated learning for English language learners. *Journal of Literacy Research*, *46*(3), 358-382.

Collins, L. J., & Liang, X. (2014). Task relevance in the design of online professional development for teachers of ELLs: AQ methodology study. *Turkish Online Journal of Distance Education*, *15*(3), 268-281.

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Annals of Internal Medicine*, *126*(5), 376-380.

Cook, J. L., & Cook, G. (2005). *Child development*. London: Pearson education, Allyn & Bacon.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research*, *52*(2), 291-302. Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of educational research*, *49*(2), 222-251.

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 175-187.

Cummins, J. (1981). *Bilingualism and minority-language children. language and literacy series*. Toronto, Ontario: The Ontario Institute for Studies in Education.

Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review*, *56*(1), 18-37.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control?. *Clinical Psychology Review*, *18*(1), 23-45.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). Professional learning in the learning profession: A status report on teacher development in the United States and abroad. Dallas, TX: National Staff Development Council. Retrieved from http://learningforward.org/docs/pdf/nsdcstudy2009.pdf.

Dash, S., Magidin de Kramer, R., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of online professional development or teacher quality and student achievement in fifth grade mathematics. *Journal of Research on Technology in Education*, *45*(1), 1-26.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3-14.

Dede, C. (2006). *Online professional development for teachers: Emerging models and methods*. Cambridge, MA: Harvard Education Press.

Dillenbourg, P. (1999). What do you mean by 'collaborative learning?' In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1–19). Oxford: Elsevier.

Diaz, S., Moll, L. C., & Mehan, H. (1986). Sociocultural resources in instruction: A context-specific approach. In *Beyond language: Social and cultural factors in schooling language minority students (pp. 187–230).* Bilingual Education Office, California State Department of Education: Los Angeles: Evaluation, Dissemination, and Assessment Center, California State University.

Donato, R. (1994). Collective scaffolding in second language learning. In J. Lantolf, &G. Appel (Eds.). *Vygotskian approaches to second language research*, (pp.33-56). Norwood, NJ: Ablex Pub, Corp.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*(3-4), 327-350.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*, *18*(2), 237-256.

Echevarria, J., Vogt, M. E., & Short, D. (2000). *Making content comprehensible for English language learners: The SIOP model*. Needham Heights, MA: Allyn & Bacon.

Edelsky, C. (1996) *With justice and literacy for all: rethinking the social in language and education*. London: Taylor & Francis.

Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*, 605–619.

Enright, D. S., & McCloskey, M. L. (1988). *Integrating English: Developing English language and literacy in the multilingual classroom*. Boston: Addison Wesley Publishing Company.

Erickson, F. (1986). Qualitative methods in research teaching. In M.C. Wittrock (Ed.) *Handbook of Research on Teaching* (3rd.edn). London: Collier Macmillan.

Foorman, B.F., Flecher, J.M., & Francis, D.J. (2004). Early reading assessment. In M. Evers & H. Walberg (Eds.) *Testing student learning, evaluating teaching effectiveness* (p. 81-125). Stanford, CA: Hoover Institution Press.

Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practices during reading/language arts instruction and its relationship to student achievement. In S. Vaughn & K. L. Briggs (Eds.). *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 1–30). Baltimore: Paul H. Brookes Publishing Company.

Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16*(4), 203-212.

Forbes, C. T., & Davis, E. A. (2010). Curriculum design for inquiry: Preservice elementary teachers' mobilization and adaptation of science curriculum materials. *Journal of Research in Science Teaching*, *47*(7), 820-839.

Francis, D. J., Rivera, M., Lesaux, N. K., Kieffer, M. J., & Rivera, H. (2006). *Practical guidelines for the education of English language learners*. Paper presented at LEP Partnership Meeting, Washington, DC. Retrieved from http://www. centeroninstruction. org.

Frederiksen, J. R., Sipusic, M., Sherin, M., & Wolfe, E. W. (1998). Video portfolio assessment: Creating a framework for viewing the functions of teaching. *Educational Assessment*, *5*(4), 225-297.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*(1), 174-206.

Garg, A. X., Hackam, D., & Tonelli, M. (2008). Systematic review and meta-analysis: when one study is just not enough. *Clinical Journal of the American Society of Nephrology*, *3*(1), 253-260.

Genesse, F., Lindolm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in US schools: An overview of research. *Journal of Education for Students Placed at Risk*, *10*(4), 363-385.

Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children*, *66*(4), 454-470.

Gersten, R., Baker, S. K., Haager, D., & Graves, A. W. (2005). Exploring the role of teacher quality in predicting reading outcomes for first-grade English learners: An observational study. *Remedial and Special Education*, *26*(4), 197-206.

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in elementary grades: A practice guide* (NCEE 2007-4011). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher study group: Impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *American Educational Research Journal*, *47*(3), 694-739.

Gersten, R., & Geva, E. (2003). Teaching reading to early language learners. *Educational Leadership*, *60*(7), 44-49.

Gersten, R., & Jimenez, R. T. (1994). A delicate balance: Enhancing literature instruction for students of English as a second language. *The Reading Teacher*, *47*(6), 438-449.

Gillani, B. B. (2010). Inquiry-based training model and the design of e-learning environments. *Issues in Informing Science & Information Technology*, *7*, 1-9.

Glass, G., Cahen, L., Smith, M., & Filby, N. (1982). *School class size: Research and policy*. Beverly Hills, CA: Sage Publications.

Good, R. H., & Kaminski, R. A. (Eds.). (2007). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed). Eugene, OR: Institute for the Development of Educational Achievement.

Greenwood, C. R., Arreaga-Mayer, C., Utley, C. A., Gavin, K. M., & Terry, B. J. (2001). Classwide peer tutoring learning management system: Applications with elementary-level English language learners. *Remedial and Special Education*, *22*(1), 34-47.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, *81*(3), 371.

Haager, D., Gersten, R., Baker, S., & Graves, A. (2003). The English-language learner classroom observation instrument: Observations of beginning reading instruction in urban schools. In S. Vaughn & K. Briggs (Eds.), *Reading instruction observation tools* (pp. 111–144). Baltimore: Brookes.

Hamre, B. K., Partee, A., & Mulcahy, C. (2017). Enhancing the impact of professional development in the context of preschool expansion. *AERA Open*, *3*(4), 1-16.

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*(1), 60-87.

Hedges, L. V., & Hedberg, E. C. (2013). Intraclass correlations and covariate outcome correlations for planning two-and three-level cluster-randomized experiments in education. *Evaluation Review*, *37*(6), 445-489.

Irby, B. J., Tong, F., Lara-Alecio, R., Meyer, D., & Rodriguez, L. (Winter, 2007). The critical nature of language of instruction compared to observed practices and high stakes tests in transitional bilingual classroom. *Research in the Schools*, *14*(2), 26-36.

Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory into Practice*, *38*(2), 67-73.

Joyce, B., & Calhoun, E. (2010). *Models of professional development: A celebration of educators*. Thousand Oaks CA: Corwin.

Kane, Thomas J., Eric S. Taylor, John H. Tyler, and Amy L. Wooten. "Evaluating teacher effectiveness: can classroom observations identify practices that raise achievement?" *Education Next* 11, no. 3 (2011): 54-61.

Kieffer, M. J. (2012). Early oral language and later reading development in Spanish-speaking English language learners: Evidence from a nine-year longitudinal study. *Journal of Applied Developmental Psychology*, *33*(3), 146-157.

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology*, *90*(1), 134.

Klingner, J. K., & Edwards, P. A. (2006). Cultural considerations with response to intervention models. *Reading Research Quarterly*, *41*(1), 108-117.

Klingner, J. K., & Vaughn, S. (2004). Strategies for struggling second-language readers. *Adolescent Literacy Research and Practice*, 183-209.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S. D. (2004). *The power of reading: Insights from the research*. Westport, CT: Libraries Unlimited.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.

Lantolf, J. P. (Ed.). (2000). *Sociocultural theory and second language learning*. Oxford, UK: Oxford University Press.

Lara-Alecio, R., & Parker, R. I. (1994). A pedagogical model for transitional English bilingual classrooms. *Bilingual Research Journal*, *18*(3-4), 119-133.

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

Lara-Alecio, R., Tong, F., Irby, B. J., & Mathes, P. (2009). Teachers' pedagogical differences during ESL block among bilingual and English-immersion kindergarten classrooms in a randomized trial study. *Bilingual Research Journal*, *32*(1), 77-100.

Lee, O. (2005). Science education with English language learners: Synthesis and research agenda. *Review of Educational Research*, *75*(4), 491-530.

Lee, O., & Buxton, C. A. (2013). Teacher professional development to improve science and literacy achievement of English language learners. *Theory Into Practice*, *52*(2), 110-117.

Lee, O., Penfield, R., & Maerten-Rivera, J. (2009). Effects of fidelity of implementation on science achievement gains among English language learners. *Journal of Research in Science Teaching*, *46*(7), 836-859.

Lee, C. D., & Smagorinsky, P. (Eds.). (2000). *Vygotskian perspectives on literacy research: Constructing meaning through collaborative inquiry*. Cambridge, UK: Cambridge University Press.

Lewis, V. K., & Shaha, S. H. (2003). Maximizing learning and attitudinal gains through integrated curricula. *Education*, *123*(3), 537-548.

Lesaux, N. K. & P. Gamez. (2012). *Examining Classroom Talk in the San Diego Unified School District*. Council of the Great City Schools. Language diversity and literacy development research group, Harvard University. Retrieved from http://www.cgcs.org/cms/lib/DC00001581/Centricity/ Domain/85/SanDiego_2012.pdf.

Lillehoj, C. J., Griffin, K. W., & Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and relation to youth outcomes. *Health Education & Behavior*, *31*(2), 242-257.

Lipsey M.W., & Wilson D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Liu, S., & Wang, J. (2015). Reading cooperatively or independently? Study on ELL student reading development. *Reading Matrix: An International Online Journal*, *15*(1), 102-120.

Lonigan, C. J., Farver, J. M., Phillips, B. M., & Clancy-Menchetti, J. (2011). Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. *Reading and Writing*, *24*(3), 305-337.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). *Designing professional development for teachers of science and mathematics*. CA: Corwin Press.

Maheady, L., Mallette, B., & Harper, G. F. (2006). Four classwide peer tutoring models: Similarities, differences, and implications for research and practice. *Reading & Writing Quarterly*, *22*(1), 65-89.

Maher, C. A., Lewis, L. K., Ferrar, K., Marshall, S., De Bourdeaudhuij, I., & Vandelanotte, C. (2014). Are health behavior change interventions that use online social networks effective? A systematic review. *Journal of Medical Internet Research*, *16*(2).

Mathes, P. G., Fuchs, D., Fuchs, L. S., & Henley, A. M. (1994). Increasing strategic reading practice with Peabody Classwide Peer Tutoring. *Learning Disabilities Research & Practice*, *9*(1), 44-48.

Mathes, P. G., Torgesen, J. K., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., & Grek, M. (2003). A comparison of teacher-directed versus peer-assisted instruction to struggling first-grade readers. *The Elementary School Journal*, *103*(5), 459-479.

McMaster, K. L., Kung, S., Han, I., & Cao, M. (2008). Peer-assisted learning strategies: A "tier 1" approach to promoting English learners' response to intervention. *Exceptional Children*, *74*(2), 194-214.

McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education*, *93*(2), 233-268.

Medley, D. M. (1992). Structured observation. In M. Aikin (Ed.), *Encyclopedia of educational research* (Vol. 6, pp. 1310-1315). New York: Macmillan.

Mercer, N. (2013). The social brain, language, and goal-directed collective thinking: A social conception of cognition and its implications for understanding how we think, teach, and learn. *Educational Psychologist*, *48*(3), 148-168.

Miller, J. F., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. J. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research & Practice*, *21*(1), 30-43.

National Center for Education Statistics. (2010). *The Nation's Report Card: Grade 12 Reading and Mathematics 2009 National and Pilot State Results (NCES 2011–455).* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2011). *The Nation's Report Card: Science 2009 (NCES 2011–451).* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2015). *The condition of education 2015* (No. 2015–144). Washington, DC: Institute of Education Sciences.

National Center for Education Statistics. (2017). *The condition of education 2017* (No. 2017–144). Washington, DC: Institute of Education Sciences.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Noell, G. H. (2010). Empirical and pragmatic issues in assessing and supporting intervention implementation in school. In G. G. Peackock, R. A. Ervin, E. J. Daly, & K. W. Merrell (Eds.), *Practical handbook in school psychology* (pp. 513–530). New York, NY: Guilford.

Ortiz, A. A., & Artiles, A. J. (2010). Meeting the needs of ELLs with disabilities: A linguistically and culturally responsive model. In G. Li & P. A. Edwards (Eds.), *Best practices in ELL instruction* (pp. 247–272). New York, NY: Guilford.

Peregoy, S., & Boyle, O. (2005). *Reading, writing, and learning in ESL: A resource book for k-12 teachers* (4th ed). Boston: Pearson/Allyn and Bacon.

Piaget, J. (1932). *The moral judgment of children*. London: Routledge & Kegan-Paul.

Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal*, *45*(2), 365-397.

Piasta, S. B., Justice, L. M., O'Connell, A. A., Mauck, S. A., Weber-Mayrer, M., Schachter, R. E., ... & Spear, C. F. (2017). Effectiveness of large-scale, state-sponsored language and literacy professional development on early childhood educator outcomes. *Journal of Research on Educational Effectiveness*, *10*(2), 354-378.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246.

Pyle, D., Pyle, N., Lignugaris/Kraft, B., Duran, L., & Akers, J. (2017). Academic effects of peer-mediated interventions with English language learners: A research synthesis. *Review of Educational Research*, *87*(1), 103-133.

Ramirez, J. D., Yuen, S. D., Ramey, D. R., Pasta, D. J., & Billings, D. K. (1991). *Final report: Longitudinal study of structured immersion strategy, early-exit, and late-exit transitional bilingual education programs for language-minority children.* San Mateo, CA: Aguirre International.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage.

Rienties, B., Héliot, Y., & Jindal-Snape, D. (2013). Understanding social learning relations of international students in a large classroom using social network analysis. *Higher Education*, *66*(4), 489-504.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, *95*(2), 240.

Ross, S. G., & Begeny, J. C. (2011). Improving Latino, English language learners' reading fluency: The effects of small-group and one-on-one intervention. *Psychology in the Schools*, *48*(6), 604-618.

Roswal, G. M., Mims, A. A., Evans, M. D., Smith, B., Young, M., Burch, M., ... & Block, M. (1995). Effects of collaborative peer tutoring on urban seventh graders. *The Journal of Educational Research*, *88*(5), 275-279.

Rubak, S., Sandbæk, A., Lauritzen, T., & Christensen, B. (2005). Motivational interviewing: a systematic review and meta-analysis. *British Journal of General Practice*, *55*(513), 305-312.

Rubin, D. I. (2016). Growth in oral reading fluency of Spanish ELL students with learning disabilities. *Intervention in School and Clinic*, *52*(1), 34-38.

Saenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children*, *7*1(3), 231-247.

Sampson, M., McGowan, J., Cogo, E., & Horsley, T. (2006). Managing database overlap in systematic reviews using Batch Citation Matcher: case studies using Scopus. *Journal of the Medical Library Association*, *94*(4), 461.

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.

Saunders, K. J., & Fisher, D. L. (2006). An action research approach with primary pre-service teachers to improve university and primary school classroom environments. In D. L.

Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments* (pp. 247–272). Singapore: World Views.

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *Splintered vision: An investigation of US mathematics and science education.* Norwel, MA: Kluwer Academic Publishers.

Shaha, S. H., & Ellsworth, H. (2013). Predictors of success for professional development: Linking student achievement to school and educator successes through on-demand, online professional learning. *Journal of Instructional Psychology*, *40*(1), 19.

Shaha, S. H., Glassett, K. F., & Copas, A. (2015). The impact of teacher observations with coordinated professional development on student performance: A 27-state program evaluation. *Journal of College Teaching & Learning*, *12*(1), 55-64.

Shaha, S., Lewis, V., O'Donnell, T., & Brown, D. (2004). Evaluating professional development. *Journal of Research in Professional Learning*, *1*, 1-18.

Shanahan, T. (1998). On the effectiveness and limitations of tutoring in reading. *Review of research in education*, *23*(1), 217-234.

Shanahan, T., & Beck, I. L. (2006). Effective literacy teaching for English-language learners. In D. August & D. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 415–488). Mahwah, NJ: Erlbaum.

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education*, *20*(2), 163-183.

Short, D. J., Fidelman, C. G., & Louguit, M. (2012). Developing academic language in English language learners through sheltered instruction. *TESOL Quarterly*, *46*(2), 334-361.

Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P., & Hodge, J. P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *The Elementary School Journal*, *95*(5), 387-408.

Slavin, R. E. (1988). *Student team learning: An overview and practical guide*. West Haven, CT: NEA Professional Library.

Slavin, R. E. (1991). Synthesis of research of cooperative learning. *Educational leadership*, *48*(5), 71-82.

Slavin, R. E. (1996). Cooperative learning in middle and secondary schools. *The Clearing House*, *69*(4), 200-204.

Slavin, R. E., Madden, N., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Educational Evaluation and Policy Analysis*, *33*(1), 47-58.

Smith, S. U. (2014). Frameworks shaping an online professional development program for K–12 teachers of ELLs: Toward supporting the sharing of ideas for empowering classroom teachers online. *TESOL Journal*, *5*(3), 444-464.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. CA: Rand Corporation.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children committee on the prevention of reading difficulties in young children.* Washington, DC: National Research Council.

Stainback, S., & Stainback, W. (1992). *Curriculum considerations in inclusive classrooms*. Baltimore: Paul H. Brookes.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist*, *35*(2), 87-100.

Stigler, J., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study*. Washington, DC: US Department of Education. National Center for Education Statistics. Retrieved from http://nces.ed.gov/timss.

Strong, M. (1983). Social styles and the second language acquisition of Spanish-speaking kindergartners. *TESOL Quarterly*, *17*(2), 241-258.

Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, *22*, 171.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*(3), 371-391.

Taylor, J. A., Getty, S. R., Kowalski, S. M., Wilson, C. D., Carlson, J., & Van Scotter, P. (2015). An efficacy trial of research-based curriculum materials with curriculum-based professional development. *American Educational Research Journal*, *52*(5), 984-1017.

Tinajero, J. v., &: Hurley, S. (2000). Exemplary schooling for intermediate and middle school students acquiring English (pp, 172-190). In K. Wood and T. Dickinson, (Eds.), *Promoting literacy in grades 4-9: A handbook for teachers and administrators*. Boston: Allyn & Bacon.

Tindal, G., Marston, D., & Deno, S. L. (1983). *The reliability of direct and repeated measurement* (Research Report No. 9). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities

Tong, F., Irby, B. J., & Lara-Alecio, R. (2015). Teachers' perception of virtual professional development in a randomized control trial. *International Journal of New Technology and Research, 1*(7), 58-61.

Tong, F., Irby, B. J., Lara-Alecio, R., & Koch, J. (2014). Integrating literacy and science for English language learners: From learning-to-read to reading-to-learn. *The Journal of Educational Research, 107*(5), 410-426.

Tong, F., Irby, B. J., Lara-Alecio, R., Yoon, M., & Mathes, P. G. (2010). Hispanic English learners' responses to longitudinal English instructional intervention and the effect of gender: A multilevel analysis. *The Elementary School Journal, 110*(4), 542-566.

Tong, F., Irby, B., Tang, S., Lin, S., Guerrero, C., Lara-Alecio, R., & Lopez, T. (2017 April). *Indicators of inter-rater reliability for classroom observation instruments as fidelity of implementation in large-scale RCTs.* Paper presented at the annual meeting of American Educational Research Association, San Antonio, TX.

Tong, F., Lara-Alecio, R., Irby, B., Mathes, P., & Kwok, O. M. (2008). Accelerating early academic oral English development in transitional bilingual and structured English immersion programs. *American Educational Research Journal, 45*(4), 1011-1044.

Tong, F., Luo, W., Irby, B. J., Lara-Alecio, R., & Rivera, H. (2017). Investigating the impact of professional development on teachers' instructional time and English learners' language development: a multilevel cross-classified approach. *International Journal of Bilingual Education and Bilingualism, 20*(3), 292-313.

Topping, K., Buchs, C., Duran, D., & Van Keer, H. (2017). *Effective peer learning: From principles to practical implementation*. London and New York: Routledge.

Totten, S., Sills, T., Digby, A., & Russ, P. (1991). Collaborative learning: A guide to research. *Scandinavian Journal of Educational Research*, *33*(4), 231-243.

Trueba, H. T. (1989). *Raising silent voices: Educating linguistic minorities for the 21st century.* Boston: Heinle & Heinle.

Tudge, J. R. H. (1992). Processes and consequences of peer collaboration: A Vygotskian analysis. *Child Development*, *63*(6), 1364–1379.

Turner, J. C., & Meyer, D. K. (2000). Studying and understanding the instructional contexts of classrooms: Using our past to forge our future. *Educational Psychologist*, *35*(2), 69-85.

U.S. Department of Education, (2015). *Every student succeeds act summary*. Retrieved from http://www.ed.gov/ESSA.

Usiskin, Z. (1985). We need another revolution in secondary school mathematics. In C. R. Hirsch (Ed.), *The secondary school mathematics curriculum: 1985 Yearbook* (pp. 1-21). Reston, VA: National Council of Teachers of Mathematics.

Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., ... & Francis, D. J. (2006). Effectiveness of Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities*, *39*(1), 56-73.

Vaughn, S., Schumm, J. S., Klingner, J., & Saumell, L. (1995). Students' views of instructional practices: Implications for inclusion. *Learning Disability Quarterly*, *18*(3), 236-248.

Vygotsky, L. (1962) *Language and thought*. Cambridge MA: MIT Press.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34-41.

Vygotsky, L. (1987). *The collected works of L. S. Vygotsky*. New York: Plenum.

U.S. Department of Education (2016). Educational experiences of English learners: Grade retention, high school graduation, and GED attainment, 2011–12. Retrieved from https://www2.ed.gov/rschstat/eval/title-iii/crdc-retention-completion.pdf.

Ward, S., Pellett, H. H., & Perez, M. I. (2017). Cognitive disequilibrium and service-learning in physical education teacher education: Perceptions of pre-service teachers in a study abroad experience. *Journal of Teaching in Physical Education*, *36*(1), 70-82.

Wasik, B. A., & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology*, *103*(2), 455.

Waxman, H. C., Tharp, R. G., & Hilberg, R. S. (Eds.). (2004). *Observational research in US classrooms: New approaches for understanding cultural and linguistic diversity*. Cambridge: Cambridge University Press.

Waxman, H. C., Wang, M. C., Lindvall, C. M., & Anderson, K. A. (1983). *The observation rating scale for features of adaptive instruction.* Pittsburg: Learning Research and Development Center, University of Pittsburgh.

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis*, *17*(2), 239-261.

What Works Clearinghouse (2007a). Bilingual cooperative integrated reading and composition. Retrieved from http://lib-ezproxy.tamu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED495727&site=ehost-live

What Works Clearinghouse (2007b). Peer tutoring and response groups. Retrieved from http://lib-

ezproxy.tamu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED499296&site=ehost-live.

What Works Clearinghouse (2010). Peer-assisted learning strategies. Retrieved from http://lib-ezproxy.tamu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED512003&site=ehost-live.

What Works Clearinghouse. (2015a). *WWC Standard Brief for Attrition*. Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_attrition_080715.pdf.

What Works Clearinghouse. (2015b). *WWC Standard Brief for Baseline Equivalence*. Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_baseline_080715.pdf.

Wiley, T. G. (1996). *Literacy and language diversity in the United States. Language in education: Theory and Practice 87*. Washington, DC: Center for Applied Linguistics and Delta Systems.

Wilson, D. B. (n.d.). Practical meta-analysis effect size calculator [Online calculator]. Retrieved from https://www.campbellcollaboration.org/this-is-a-web-based-effect-size-calculator/explore/this-is-a-web-based-effect-size-calculator.

Wong Fillmore, L., Ammon, P., McLaughlin, B., & Ammon, M. S. (1985). Learning English through bilingual instruction (Final Report). Washington, DC: National Institute of Education.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100.

Zhang, J., Anderson, R. C., & Nguyen-Jahiel, K. (2013). Language-rich discussions for English language learners. *International Journal of Educational Research, 5*8, 44-60.

APPENDIX A

COLLEGE OF EDUCATION
AND HUMAN DEVELOPMENT

Educational Psychology
Bilingual Education Program

TEXAS A&M UNIVERSITY

To whom it may concern,

I hereby grant Ms. Shifang Tang permission to reproduce the Model/theory of the four-dimensional bilingual pedagogical theory in her dissertation study. Analysis of First Grade Bilingual Classroom Observation using the Transitional Bilingual Observation Protocol: Language of Instruction, Language Content, Communication Mode, and Activity Structure.

Name: Dr. Rafael Lara-Alecio

Signed Rafael La

Date: February 21, 2018

For permission to use this theory/mode, contact

Rafael Lara-Alecio, Ph.D.
Regents Professor, Texas A&M University System
Professor & PI for Projects ELLA-V, LISTO, & ETELL
Director, Center for Research & Development in Dual
Language & Literacy Acquisition (CRDLLA)
Director, Bilingual Programs
Department of Educational Psychology
College of Education and Human Development
Texas A&M University, College Station, TX 77843-4225
979-845-3467 (Office), 979-845-2599 (Front Desk)
E-mail: a-lara@tamu.edu

100 Harrington Tower
4225 TAMU
College Station, TX 77843-4225

Tel. 979.845.2599 Fax. 979.458.0192
www.ldn.tamu.edu

COLLEGE OF EDUCATION
AND HUMAN DEVELOPMENT

Educational Psychology
Bilingual Education Program

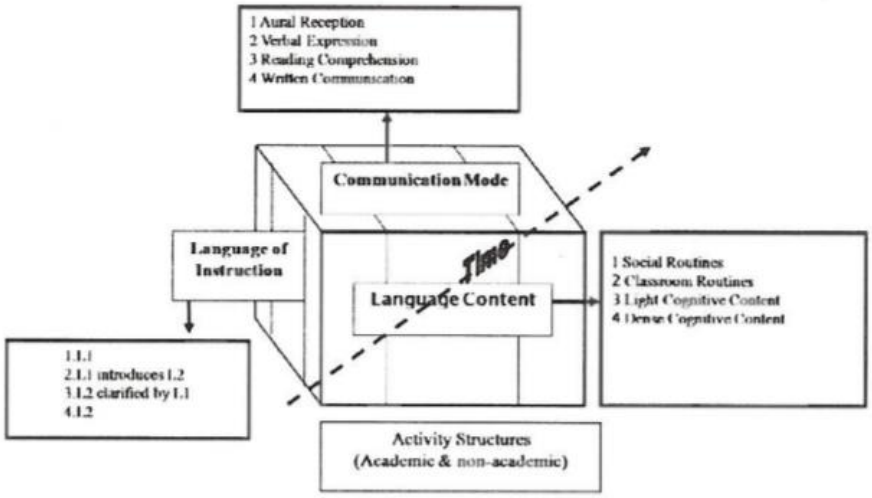

Figure 3.1 Four-dimensional transitional bilingual pedagogical theory model (Lara-Alecio & Parker, 1994)

100 Harrington Tower
4225 TAMU
College Station, TX 77843-4225

Tel. 979.845.2599 Fax. 979.458.0192
www.ldn.tamu.edu

\\

COLLEGE OF EDUCATION
AND HUMAN DEVELOPMENT

Educational Psychology
Bilingual Education Program

TEXAS A&M UNIVERSITY

TRANSITIONAL BILINGUAL OBSERVATION PROTOCOL

Observer: _____ Date: _____
Start Time: _____ End Time: _____
School: _____ Teacher: _____ Grade: _____

20 second intervals	Time	Strategy	Curriculum Area	Physical Group	Activity Structure	Mode	Language Content	Lang of Instruction Teacher/Student	
	1								
	2								
	3								
	4								
	5								
	6								
	7								
	8								
	9								
	10								
	11								
	12								
	13								
	14								
	15								
	CODING	1 QS	1 read/lit	1 TC	1 lec/lis	1 writing	1 social	1 L1	1 L1
		2 ALS	2 math	2 LG	2 lec/per	2 reading	2 academic	2 L2	2 L2
		3 VS	3 spell	3 SG	3 dir/lis	3 aural	3 light cog	3 L1-2	3 L1-2
		4 MR	4 hand	4 Pairs	4 dir/per	4 verbal	4 dns cog	4 L2 -1	4 L2-1
		5 AO	5 science	5 Single	5 dem/lis	5 wr-re		5 NA	5 NA
		6 CG	6 soc sci		6 led/per	6 wr-au			
		7 CC	7 health		7 ask/per	7 wr-ver			
		8 LC	8 PE		8 ask/ans	8 re-wr			
		9 IT	9 Music		9 ans/ask	9 re-au			
		10 NA	10 Art		10 ev/per	10 re-ver			
			11 Lang		11 obs/per	11 au-wr			
			12 Compos		12 ev/dis	12 au-re			
			13 Non-ac		13 ev/cop	13 ver-wr			
			14 ESL		14 obs/dis	14 ver-re			
					15 obs/cop	15 ver-au			
					16 NA/free	16 au-re-ver			
					17 NA/feed	17 NA			
					18 NA/tran	18 au-ver			
					19 NA/int				
					20 NA/out				
					21 interact				

100 Harrington Tower
4225 TAMU
College Station, TX 77843-4225

Tel. 979.845.2599 Fax. 979.458.0192
www.ldn.tamu.edu

APPENDIX B

RESEARCH SYNTHESIS CODING SUMMARY

	Sample Characteristics			
Studies	**Grade Level**	**Size**	**Language Groups**	**SES Background**
Almaguer et al. (2005)	3	80	homogenous Spanish-speaking ELLs	high-poverty
Calhoon et al. (2007)	1	94	heterogeneous language samples	high-poverty
Greenwood et al. (2001)	1-5	117	heterogeneous language samples	NA
Liu & Wang (2015)	4	351	heterogeneous language samples	NA
McMaster et al. (2008)	k	112	heterogeneous language samples	NA
Saenz et al. (2005)	4	132	homogenous Spanish-speaking ELLs	NA
Zhang et al. (2013)	5	75	homogenous Spanish-speaking ELLs	low to middle SES

	Study/intervention Features									
Studies	**Focus of Study/Intervention**	**Instructional Time**				**Group Size**	**Fidelity of Implementation**		**Professional Development**	
		Duration (weeks)	Frequency (n/week)	Intensity (minute)	Total time (hours)		Methods	Fidelity	Process	Form
Almaguer et al. (2005)	oral reading fluency, reading comprehension	9	5	30	22.5	whole group	student attendance records	91%	only initial PD	face-to-face
Calhoon et al. (2007)	phoneme awareness, letter-word identification, and word fluency	20	3	30-35	30-33.3	whole group	classroom observation checklist	94.9% to 98.7%, on average 96.4%	initial PD + through the entire intervention	face-to-face
Greenwood et al. (2001)	professional development supporting ELLs' reading vocabulary	15-21	4	20	20-28	whole group	classroom observation checklist	80% or above	initial PD + through the entire intervention	face-to-face
Liu & Wang (2015)	reading comprehension	NA	NA	NA	NA	NA	NA	NA	NA	NA
McMaster et al. (2008)	oral reading fluency, Phoneme awareness, letter-word identification, and word fluency	18	4	20-30	24-36	whole group	classroom observation checklist	90% or above	only initial PD	face-to-face
Saenz et al. (2005)	oral reading fluency, reading comprehension	15	3	35	26.25	whole group	classroom observation checklist	90% or above	only initial PD	face-to-face
Zhang et al. (2013)	reading comprehension; oral expression	4	2	20	2.67	whole group	not report	Not report	only initial PD	face-to-face

	Design Characteristics							
Studies	**Random Assignment**	**Pre- and Post-Testing**	**Attrition Rate**		**Baseline Equivalence**	**Comparison Group Type**		**Data Analysis**
			Rate	Across Condition		Demographic	Instruction	
Almaguer et al. (2005)	quasi-experimental	pre-& post-test	not reported	not reported	not examined, use ANCOVA to adjust the baseline equivalence	reported	reported	ANCOVA
Calhoon et al. (2007)	randomized control	pre-& post-test	23.70%	not reported	examined and achieved	reported	reported	ANOVA
Greenwood et al. (2001)	single-subject design	pre-& post-test	not reported	not reported	NA	NA	NA	chi-square analysis
Liu & Wang (2015)	NA	NA	NA	NA	NA	NA	NA	correlation analysis
McMaster et al. (2008)	quasi-experimental	pre-& post-test	not reported	not reported	not examined, use ANCOVA adjust the baseline equivalence	reported	reported	ANCOVA
Saenz et al. (2005)	randomized control	pre-& post-test	10.90%	not reported	examined and achieved	reported	reported	ANOVA
Zhang et al. (2013)	quasi-experimental	pre-& post-test	not reported	not reported	examined and achieved	reported	not reported	ANCOVA, MANCOVA

	Outcome Characteristics		
Studies	**Assessment**		**Reliability and Validity of Outcome Measures**
	Type	Ability	
Almaguer et al. (2005)	curriculum-based measures; published measures	phonological awareness; reading comprehension	reported for published measures; not reported for curriculum-based measure
Calhoon et al. (2007)	standardized measures	letter-word identification, and word fluency	reported for standardized test;
Greenwood et al. (2001)	curriculum-based	reading vocabulary	not reported for curriculum-based measure
Liu & Wang (2015)	standardized measures	reading comprehension	not reported for standardized test;
McMaster et al. (2008)	curriculum-based measures; standardized measures	oral reading fluency; letter-word identification, and word fluency	reported for standardized test; not reported for curriculum-based measures and published measures
Saenz et al. (2005)	published measures	reading comprehension	reported for published measures;
Zhang et al. (2013)	standardized measures; published measures	reading comprehension; oral expression	not reported for standardized test; reported for published measures;

	Impact the CCP Strategy on ELs' Reading and Speaking Proficiency				
Studies	**Oral Language (Cohen's *d*)**	**Oral Reading Fluency (Cohen's *d*)**	**Reading Comprehension (Cohen's *d*)**	**Phoneme Awareness, Letter-word Identification, and Word Fluency (Cohen's *d*)**	**Professional Development Supporting ELs' Reading Vocabulary(Cohen's *d*)**
Almaguer et al. (2005)		positive (2.53)	Positive (reading comprehension question, 1.03); Positive (reading comprehension cloze, 0.33)		
Calhoon et al. (2007)		No difference		positive (word fluency, 0.75; letter-naming fluency, 0.67); No difference (Phoneme segmentation)	
Greenwood et al. (2001)					individualized consultation positively impact ELs' reading vocabulary (0.63)
Liu & Wang (2015)			Negative for small group instruction (0.5) and pair work instruction (0.34)		
McMaster et al. (2008)		No difference	Positive (reading comprehension question, 1.03)	positive (letter naming, 1.08; phonemic awareness (segmentation, 1.00; blending, 1.02)	
Saenz et al. (2005)			Positive (reading comprehension question, 2.67)		
Zhang et al. (2013)	positively impact students' coherent narratives (0.37)		No difference (reading comprehension cloze)		