Effects of Academic Intervention of English for Struggling Readers in General and Special Education Setting

August, 2021

Graduate School of Education
Seoul National University
Special Education Major

HyeYun Gladys Shin

Effects of Academic Intervention of English for Struggling Readers in General and Special Education Setting

Dr. Dongil Kim, Adviser

Submitting a Ph.D. Dissertation of
Special Education

August, 2021

Graduate School of Education
Seoul National University
Special Education Major

HyeYun Gladys Shin

Confirming the Ph.D. Dissertation written by
HyeYun Gladys Shin
August, 2021

Chair	Changdae Kim
Vice Chair	Jaehyun Shin
Examiner	Hyejung Koh
Examiner	Heeju Kim
Examiner	Dongil Kim

Acknowledgments

Standing on the edge of my academic journey after receiving B.A. in 2000, Pennsylvania; M.A. in 2004, New York; M.S. in Education in 2007, New York; and Ph.D. in 2021, Seoul, I must have been one blessed and lucky life-long learner. For my doctoral dissertation and continuously joyful learning experience, my gratitude must go to Dr. Dongil Kim for your professionalism, warm-hearted opinions and care throughout the journey, to Dr. Tom McIntyre for your continuous support even nine years after my graduation from your institute and for your encouraging letters, to Dr. Changdae Kim for your academic support and being the Committee Chair for my dissertation thesis, to Dr. Hyejung Koh for your continuous guidance, friendship and academic rigor, to Dr. Jaehyun Shin for your guidance in meta-analysis and companionship at PCRC, Dr. Heeju Kim for being an exemplary senior and scholar in the special education department at Seoul National University. Thank you all for being my mentors, committee members and future colleagues. You guided me through the course of my work whether I was academically inquisitive, culturally fumbling, or personally approaching with lots of questions. Thank you all for your warm and inspiring support in New York and in Seoul. Last but not least, a big "thank-you" to my college pal Inna Ryu, colleagues of NYC Teaching Fellows and Hong Kong Korean International School, and to all of my lab-mates here at Seoul National University for your genuine and loving heart you gave me. Thank you Yeji, HeeEun, Yeonjae and Seohyeon for your support and friendship.

Dedication

This thesis is dedicated to all my parents here and abroad for your endless love and sacrifice near and afar. I have the most amazing parents in the world.

Table of Contents

List of Tables

List of Figures

Abstract

This meta-analysis examined effectiveness of English language interventions on struggling elementary school readers. Findings were drawn from a total of 35 studies with 150 effect sizes of experimental and quasi-experimental study designs. With the random-effects model, the overall mean effect size was g=0.655 (SE=0.054) after removing outlying influential points, and moderator analyses were conducted for possible significant moderator effects from various sub-groups. The variables used for sub-group analyses include (1) participant-related variables (e.g., classroom type, grade level, English as a Foreign Language/English Language Learners/low achievers of English as first language), (2) outcome-related variables (e.g., intervention outcome measures, type of outcome), (3) intervention components (e.g., intervention types), (4) contextual characteristics (e.g., interventionist, frequency, duration). Additionally, meta-regression analysis was also conducted in order to explore the magnitude of effects on covariates and for scatter plots. The results indicated the grade level difference, classroom type, intervention type were not significant moderators while learner type, outcome measure type (e.g., affective measure), interventionist type (e.g., teacher), particular interventions (e.g., multi-sensory/intelligence, shared reading, songs, feeling-based) along with the frequency of the intervention were significant moderators to the effects. Study limitations and educational implications are also discussed in regard to teaching various student types and toward more effective intervention approach.

Keywords: Meta-analysis, English reading intervention, special education setting, moderator analysis, struggling reader, low-achiever, elementary reader, EFL, ELL
Student Number: 2016-30772

Chapter 1. Introduction

1.1. Background and Statement of the Problem

Struggling readers of English, such as low-achievers, language underachievers and students with specific learning disabilities, face difficulties with reading for various reasons, including having problems in both oral language/vocabulary and print-related/phonological knowledge, or having adequate oral language but inadequate printed words (Foorman & Torgesen, 2001). Struggling readers can be defined as "low-achievers, students with unidentified reading difficulties, dyslexia, and/or with reading, learning or speech/language disabilities" (Wanzek et al., 2010, p. 892). To the struggling readers, the nature of reading is quite more complex than non-struggling readers – being far beyond simple understanding of the orally delivered sentences and basic word reading skills. Even in the presence of successful phonemic awareness skills, many struggling readers tend to show a high level of deficiency in comprehension although other reading-related skills are mastered at some point (National Institute of Child Health and Human Development, 2000).

One of the basic ways previous research have framed intensive intervention was to provide more time allotment during intervention. And, interventions through time, dosage, and/or instructional grouping size became the basis for the quantitative method for increasing intensity with multi-tiered interventions (National Center on Intensive Intervention, 2012). In fact, intervention study efforts and research for the past 20 years have shown that such struggling reader population who are at risk for reading failure later is often linked to low-income families, lack of support from school and home, and late detection of intervention needs (Denton, Fletcher, Anthony & Francis, 2006). Some of the common types of experience that the struggling readers face

include teachers' lowered expectations for the student, the student's own increased sense of learned helplessness, lowered self-esteem from repeated academic failure and other undesirable outcomes and experiences with peers and school overall, often expanding to further emotional and social difficulties (Kim et al., 2017). A synthesis research for the past 30 years in seeking extensive knowledge and proof on early intervention for young readers with reading difficulties, some validated instruction -explicit, systematic instruction- was found to be highly effective with upper grades, yet effective instruction that is powerful for across grades has not been proven (Wanzek et al., 2010).

In continuous search for the elements effective to reading instruction on struggling readers, previous research have found and stressed school and teacher-related factors including (1) explicit instruction, (2) early intervention, (3) small group/one-on-one intensive instruction, (4) effective emotional and cognitive relationship between teacher and student, and (5) appropriate instructional level matching the student's individual needs (Foorman & Moats, 2004; Vernon-Feagans et al., 2010)

The Alliance for Excellent Education (2012) stated that too many school students are facing the risk of dropping out of high school not because of any other reasons but because of low literacy skills with poor attendance and class failure. Many of these students have family background that is underserved and underrepresented: students of color, high-mobility students (including foster, migrant, and homeless students), English Language Learners (ELL), students with disabilities, and low-income students (Marchand-Martella, Klingner & Martella, 2013). Furthermore, the U.S. Department of Education (2003) indicated that having difficulties in reading is one of the *most* common reasons students receive special education services and limit their participation in the workplace and in society (Al Otaiba et al., 2009). It is estimated that about 69% of fourth graders cannot read at proficient levels and 36% of the fourth graders are not able to read at or above basic levels

expected by their institution (Daane et al., 2005). It is estimated that 90% of all children identified as having learning disabilities are referred to special education services setting only because of reading problems (Kavale & Forness, 2000). Torgesen and Wagner (1998) stated that (1) "the most severe reading problems of children with learning disabilities lie at the word, instead of text, level of processing," often unable to decode, and (2) the most common cognitive limitation of such children involves a dysfunction in the awareness of the phonological structure of words within oral language (p.226). Those struggling readers in special education whose performance is resistant to interventions also tend to show a second processing problem, a significant deficit in visual naming speed, compared to typically developing readers (Lovett, Steinbach & Frijters, 2000; Wolf, Bowers, & Biddle, 2000).

Normally when the skills relating to reading should be mastered for a developmental shift from 'learning to read' to 'reading to learn' various subject matters (Chall, 1983) nearing 4th grade level, having to learn a second language for non-English speaking struggling readers at that time impose an extra burden with accumulated low self-esteem and confidence (Yeo, 2010). For already-struggling readers, not much attention has been paid to their learning experience and achievement in their second language acquisition. In addition to such negative experience and learning needs that the struggling readers already have, the vast number of students around the globe are also required to learn English as the most important second language. According to the statistics by University of Winnepeg in Canada, 142 out of a total of 193 countries, excluding English speaking countries, mandate English language as part of their national curriculum requirement; EIL, English as an International Language, is globally agreed. (Lee & Chung, 2017). This denotes that the struggling readers around the world are also demanded to learn English in addition to acquiring their first language proficiency.

The students with learning difficulties with English as a

second language are bound to experience dual-disability status when attempting to read and express themselves in English (Kim, 2002). One study that examined the students with learning disabilities specifically has called the students in this situation being analogous to riding a bicycle with a big wheel (native language) and the second wheel (second language) and any haphazard approach would only yield two wheels being dysfunctional, having students with double the challenge to live with (Klinger & Vaughn, 1996). Kang (2010) also found that those students having to learn two languages with the overall underachieving score range compared to other typically developing classmates experienced more learned helplessness and serious isolation in academic settings.

English as a subject is analogous to math; a lack of mastery in previous school years almost always results in accumulated academic difficulties and eventually failure along with continuous low self-esteem (Yeo, 2010). When learning English as a Foreign Language, because the students are learning English in a non-English speaking country, the students almost never experience speaking and utilizing English language outside of their English classes (Lightbown & Spada, 2013; No, Park, & Chang, 2016). English subject matter stays as a unique subject leading to the implication that if underachievement occurs without adequate academic assistance, basic academic recovery based on self-attempt, or self-springing, is nearly impossible in learning English as the Foreign Language (EFL) atmosphere (Lee & Chang, 2017; Yeo, 2010). Because early mastery is absolutely critical for sequent successful episodes, this growing low-achieving rate clearly contributes to the "wait-to-fail" and even to the "watch-them-fail model, as stated by Reynolds and Shaywitz (2009).

Given the importance of diverse student types and school, classroom settings of struggling readers, scrutinizing effective intervention and influential variables is necessary to first analyze the current development and to draw effective factors and intervention methods to teach those struggling reader population as

a whole. Therefore, this meta-analysis study is seeking to analyze effective English interventions and related variables of struggling readers, inclusive of all classroom and population types. And, because English is introduced at the elementary school level as a mandated school subject for the second language learners as well, elementary school students are the major population for the study's purpose. In this meta-analysis, the focus is not only to find effect sizes of interventions, it also intends to find existing relationships and propensities between sub-groups or moderators.

1.2. Research Questions

The present meta-analysis aims to address the following research questions:

1. What is the overall effectiveness of English interventions on struggling readers in elementary school?

2. To what extent do intervention effects vary as per participant-related moderator? (i.e., classroom type, grade level, LA/UA/ELL /EFL status)

3. To what extent do intervention effects vary as per outcome measures and their types? (i.e., English achievement vs. affective, researcher-developed vs. standardized)

4. To what extent do intervention effects vary as per intervention components and their type (i.e., content-based, learner approach-based, teacher/instructional-based intervention)?

5. To what extent do intervention effects vary by other contextual characteristics of interventions (i.e., interventionist, intervention frequency)?

Chapter 2. Literature Review

2.1. Characteristics of Struggling Readers in General and Special Education Setting and Intervention

Researchers broadly use the term "struggling readers" to refer to "significantly at risk for reading difficulties" (Mathes, et al., 2005), "not meeting school standards and expectation" (Vernon-Feagons, et al., 2010), "those needed an extra push" in reading to meet satisfactory level (Duffy, 2000). From this broad definition, the terms "low-achievement" and "underachievement" are also used to refer to students performing below 25th percentile in reading skills compared to typically developing peers. Low-achievement or underachievement is also related to a discrepancy between expected and actual performance (McCoach & Siegle, 2003). Depending on their study purpose, researchers sometimes include/exclude special education students, language impaired students or students with limited English proficiency into struggling reader category as the study participants.

Struggling readers are low-achievers in the sense that "traditionally, the struggling reader has been viewed as a low achiever." (Guthrie & Davis, 2003, p. 60) They are seen as lacking the oral reading and/or reading comprehension skills, word recognition, and reading fluency. Those readers have been found unmotivated and disinterested in school as a whole or school-related work (Vacca & Vacca, 1999). Considerable evidence suggests that the learning disabilities (LD) population, too, is also primarily one of struggling and low achieving populations by academic performance (Ysseldyke, et. al., 1982).

Unfortunately, the nature of reading is quite complex – being far beyond simple understanding of the orally delivered sentences and basic word reading skills. Even in the presence of successful phonemic awareness skills, many struggling readers tend to show a high level of deficiency in comprehension and general understanding

of the text although other reading-related skills are mastered (Nation & Snowling, 2004). As students move up with the grade in school, simple "fun" reading activities turn into "reading to learn" subject matters, and without successfully acquiring comprehension skills in reading, it is extremely difficult to move forward with the grade level expectancy in most subjects when reading text tends to transfer to more nonfiction and expository from the narrative nature (Chall, Jacobs, & Baldwin, 1990; Kim, Shin, & Lee, 2016). These students are also seen with deficiencies in domain-general abilities such as working memory, language, and attentive behavior (Andersson & Lyxell, 2007; Fuchs et al., 2010).

For many struggling readers, past 20 years of research have noticed the importance of direct and explicit teaching, early intervention effort in first 2-3 grades (Foorman & Moats, 2004), small group or one on one teaching, emotional stability between the teacher and student (Foorman & Torgesen, 2001), appropriate instructional level for the student (Foorman & Torgesen, 2001; Vernon-Feagans, et al., 2010).

Traditionally, students had to wait until a significant discrepancy between reading achievement and intelligence was demonstrated to receive reading interventions (Donovan & Cross, 2002). However, in an attempt to prevent this “wait-to-fail” cycle of student’s performance and delayed intervening timing, the Response to Intervention (RTI) educational model, was introduced as part of the reauthorization of IDEA in 2004. It aims to provide a proactive process of early intervention and evidence-based instructions to students with additional intensive and individualized instruction. In doing so, it is expected to prevent student underachievement, including students at risk for severe academic failure and to embrace students of culturally and linguistically diverse background by doing so (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006; Fuchs & Fuchs, 2006).

There also has been more attention on the shifting of the role to the classroom/homeroom teacher in aiding and preventing

reading failure in students. Based on the individual student' needs, Teaching variations utilizing Tier II and Tier III intervention have been found effective in reading intervention in both longitudinal study and short-term study (Mathes et al., 2005). Classroom teachers' instruction of Tier II and Tier 3 are especially essential when there is a lack of other resources in schools with low wealth and the support system (Vernon-Feagans, et. al., 2010)

2.2. English Language Learner (ELL) / English as a Foreign Language (EFL) Learner Characteristics and intervention

The National Literacy Panel defines English Language Learners (ELL) as "students who come from language backgrounds other than English and whose proficiency is not developed enough to where they can profit fully from English only instruction" (Fien et al., 2011, p. 143). The ELL population is a diverse group including anyone whose first language is not English. Learners learning English as a Foreign Language (EFL) has a slightly different definition; an EFL student is learning English language in a non-English-speaking country (Lee & Chang, 2017). For example, students in Korea who are learning English in their schools are considered EFL students since English is not their official language. In contrast, if those same students were in the U.S. learning English, they would be considered ESL/ELL students for the reason that they are learning English in an English-speaking country.

As expected, a disadvantage in oral fluency, limited comprehension and the slender breadth of English vocabulary knowledge characterize ELL students (August, Carlo, Dressler, & Snow, 2005; Verhallen & Schoonen, 1993), and this gap is hard to bridge and each case may pose a challenge to educators. It is estimated that English first-language students on average bring about a vocabulary span of 5,000 to 7,000 words to reading instruction (Biemiller & Slonim, 2001). However, there has been

relatively little research examining the English vocabulary development of ELL children, reading comprehension results in the later grades, and evidence-based instructional programs designed specifically for the ELL students. There is scarcity of rigorous research on what constitutes successful reading instruction for ELLs (Slavin & Cheung, 2003). The belief is that ELL students at risk need programs that implement explicit direct instruction in phonological awareness and the alphabetic principle and supplemental instructional support over a period of months (Lovett et al., 2008).

Research indicates challenges with Response-to-Intervention (RTI) implementation with ELLs (Orosco & Klingner, 2010). There is a growing body of research on RTI implementation; however, evidence-based interventions are not applicable to all students and the impact of interventions on ELL students is still not so clear especially when ELLs require special education or are already struggling readers with their first language. There are additional factors to consider for a successful implementation of this model with this population (Orosco & Klingner 2010). According to Orosco and Klingner's (2010), preparing educators to work with ELLs requires understanding of second language acquisition which can help teachers differentiate between language acquisition and learning disabilities along with appropriate training that provides effective instructional and assessment practices (Orosco & Klingner, 2010). Goldenberg (2008) also pointed out that within the ELL population, which instruction to teach to whom, when, and for how long, has also been a challenging decision. Teaching struggling ELLs is a far more multifaceted step. Because it is often likely to misdiagnose ELLs with other disabilities and learning problems, if a student is not demonstrating adequate progress, educators are strongly encouraged to strictly evaluate their instruction methods and consider all possibilities before they are bound to assume any problems with the student (Klingner & Edwards, 2006; Ortiz, 1997).

The results from different studies of evidence-based

interventions and RTI with ELLs show some advantages of RTI with ELLs but also reveal some limitations of this model with this population. Han (2009) conducted a meta-analysis of evidence-based reading instruction for ELLs from pre-school through sixth grade. This study included 29 studies from peer-reviewed journal from a lengthy period, 1967 through 2009. The overall effect of reading instruction was moderate (ES=0.50). In this study, keyword method, proactive reading, and peer-assisted learning strategies were identified as promising practices. Han (2009) identified more than 10 programs that address phonemic awareness and phonics instructions for pre-school through second grade at both Tier I and Tier II but indicated there are limited vocabulary instructional programs available for ELLs at Tier I. The results of this study indicate the correlation between quality and effect size was not statistically significant.

Instruction methods and interventions that are developed and implemented without consideration of the specific language and learning needs of ELL students could inevitably impact their performance in school (Marchand-Martella, Klingner, & Martella, 2015). In fact, language proficiency and dominance are important variables that may well influence intervention results implemented by teachers (Ortiz, 1997; Klingner & Edwards, 2006). More importantly, ELLs are not all the same. These students have different levels of English language acquisition that can impact their rates. Thus, it is recommended that educators gather information about different factors that can impact ELLs' academic, emotional and linguistic development and response to instruction and intervention. These necessary factors include assessing the student's learning environment, academic achievement and progress, oral language and literacy pro/deficiency, personal and family, physical and psychological, previous schooling experience, and cross-cultural factors that must exist (Hamayan et al., 2015). It is clear that learning to read in a second language requires additional instructional approaches that are carefully designed after a

thorough assessment of the student.

Kangas (2017) further stated the ELLs with special needs remain quite under-researched area. In this study, he found that providing services to ELLs with special needs is not so well organized. Specifically, the findings also proved a school's policy "that both disregarded the bilingual development of ELLs with special needs and misappropriated effective models of bilingual and special education" even when bilingualism was very encouraged in school (Kangas, 2017) and asserted that more dedication and attention to developing in-service and preservice teachers' interdisciplinary knowledge about these students needs and rights must take place. In support of English language learners, Klinger (2010) offered the educational model that includes intensive assistance, culturally and linguistically responsive and differentiated instruction with ongoing progress monitoring and multiple assessments that students need.

Bondie and Zusho (2017) examined the impact of instructional routines, small group discussions and self-regulation on ELLs with learning disabilities in a health, science, technology education (HSTE) program. They found the motivation to learn as a strong indicator and that the student feelings relate to perception of understanding, classroom conversation and confidence. Five key instructional practices were identified in establishing a mastery-oriented and supportive learning environment. That is, clear goals, calm explanation, slow speed, repetition, feedback and chance to improve (Bondie & Zusho, 2017). The National Institute of Child Health and Human Development (2000) concluded that the development of literacy in English is influenced by different factors including age, language proficiency, cognitive skills, previous learning experiences, English oral proficiency, and differences between English and the first language, thereby in need of multiple approaches to teaching. Findings also indicated the instruction that provides *substantial exposure* to the essential reading components including phonemic awareness, phonics, fluency, vocabulary, and

reading comprehension, have shown effectiveness with ELLs.

2.3. EFLs and English Intervention in Korea

Korea, as an example of having students of EFL as its national curriculum mandates, introduces English 3rd grade onward as part of the national curriculum. It first appeared in public elementary schools as a school subject since 1997 (Lee, 2015), and is being continuously taught in all Korean public schools from third grade onward. Korean Ministry of Education (2015) improvised the curriculum to include letter-sound correspondences leading to reading phoneme-based syllables into morpheme-based, to reading English words starting at third and fourth grade level for students' better adaptation to the subject matter (Yoon & Lee, 2017).

English subject being introduced at third grade for all students, English remains as an important subject matter throughout middle and high school years and is also included in the Korean National College Entrance Exam. While English poses its own difficulty level as a language that is not clearly letter-sound transparent, not much research has been done on finding effective factors contributing to reading success especially for struggling readers. English subject, in particular, has shown one of the largest variances in students' skills and abilities.

Korean Ministry of Education in 2015 improvised the English curriculum to include letter-sound correspondences leading to reading phoneme-based syllables into morpheme-based, to reading words starting at third and fourth grade level. English subject in middle and high school become quite content-heavy with an increased level of difficulty in all domains (Lee, Chang, & Cheon, 2017) for students in Korea, not meeting minimum English proficiency standards at the upper primary school level is waiting for the students' failure, only to face more obstacles with the language in middle school (Seo & Bae, 2018). Unfortunately, English struggling readers are rarely identified by the institution for

supplemental academic support. The existing learning gaps at the elementary school level often discourages at-risk students who often experience repeated failure and loss of confidence and interest in the subject (An, 2016; Moon, 2015). In the end, neglecting such needs in students is depriving of the students of earlier opportunities, making them have less optimism for their future (Yoon, 2002).

Statistical findings by the Korean Ministry of Education (2019) indicated that in English low-achiever rate has risen from 3.2% to 5.3% from the previous year for the middle school population, and from 4.1% to 6.2% in high school. Any accumulated underachievement in English from second semester of 3rd grade tends to result in a lack of self-confidence, carrying on to the next grade. And, by 5th grade when full text-based reading expectancy settles in, the rate of struggling readers have shown to increase abruptly and children tend to show bigger gaps in between grades as they move up the grade (An & Lim, 2013; Lim, 2015), and this leads to losing the interests in the language learning and children have shown to give up" (Seo & Bae, 2018).

2.4. Characteristics and Domains (sub-constructs) of English Language Arts and Expectations for Struggling Readers

Compared to other alphabetic languages with more transparent letter-sound correspondence (e.g., Greek and Italian), English is considered relatively an opaque orthography (Frost, 1992). Each phoneme, the smallest unit of spoken language, can be represented by multiple graphemes (e.g. the letters, or units of written language), and each grapheme can represent multiple phonemes. For example, in English, /a/ can be written in two different ways, such as in the words "main" and "mane" and can have multiple sounds as in "cat," "Kate," or "art." This variation also

applies in the translation from letters to sound. In the case of the letter “c,” it can either sound like /s/ when followed by the vowels “e” or “i,” or like /k/ when followed by “a,” “o,” or “u.” The National Reading Panel (2000) evaluated different methods for teaching reading and concluded that alphabetics including phonemic awareness and phonics, reading fluency, vocabulary, and reading comprehension are critical components for teaching reading to young children as well as adolescents even in the general population.

On the basic reading level, phonemic awareness (PA) is defined as the ability to manipulate, blend, and segment sounds or phonemes in oral syllables and words. Unlike phonics instruction, PA does not rely on letter–sound relations when teaching students to read and spell. Correlational studies have demonstrated PA is a strong predictor of how well children later learn to read in early years of instruction (National Reading Panel, 2000). Systematic approach presents phonics in a planned sequence and within an explicit phonics method and within the incidental approach the teacher addresses phonics when given the opportunity and as part of the text (National Reading Panel, 2000). Across grades, good readers improved reading and writing with phonics instruction but these benefits were more substantial with younger students. The systematic phonics instruction also demonstrated benefits for low achieving students and students with learning disabilities (National Reading Panel, 2000).

Reading comprehension is viewed as an active and intentional thinking process “during which meaning is constructed through interactions between text and reader” (National Reading Panel, p. 4–39). Besides being an interactive process, the National Reading Panel notes reading comprehension is a cognitive process that requires complex skills, involving the understanding of vocabulary (National Reading Panel, 2000). The National Reading Panel (2000) concluded that comprehension improves when students relate print materials to prior experiences and knowledge and build mental representations. Studies show that using a combination of

techniques such as comprehension monitoring, cooperative learning, use of graphic and semantic organizers, question answering, question generation, story structure, and summarization improves reading comprehension and yields to better results in standardized tests of reading comprehension. Both reading comprehension and vocabulary involve the meaning of text at different levels (NRP, 2000).

Vocabulary is tied to individual words and comprehension to larger units. There are two types of vocabulary: expressive and receptive. An individual for verbal and written communication relates expressive vocabulary to words produce. Receptive vocabulary is the words individuals recognize by listening and reading. The reading study by the National Reading Panel recognizes the importance of vocabulary for reading but suggested that vocabulary instruction alone does not lead to improvements in reading. (Kim et. al., 2018)

According to NRP's 2000 report, National Institute of Child Health and Human Development (NICHD) in consultation with the Secretary of Education were charged a National Reading Panel (NRP) and was appointed responsible for Preventing Reading Difficulties in Young Children (Snow, Burns, & Griffin, 1998). NRC has been producing reading research and reading instruction. With recent adoptions of topics included alphabetics (phonemic awareness instruction, phonics instruction) fluency, comprehension (vocab instruction, text comprehension, teacher preparation strategies instruction) teaching education and reading comprehension, computer tech and reading instruction (NRP, 2000). Recent large-scale adoption of the Common Core State Standards (CCSS) for English Language Arts (ELA) and Mathematics gained attention for renewing focus on teaching academic English (AE) across the curriculum. Emphasized in the CCSS are the language demands of each content area; students who meet the Standards are described as those who "develop the skills in reading, writing, speaking, and listening that are the *foundation* for any creative and

purposeful expression in language" (CCSS, p. 3). As schools begin to implement the CCSS, the need for defining and understanding how to teach AE is more pertinent than before (DiCerbo, et al., 2014).

2.5. Meta-analysis and Review of Relevant Prior Research

Meta-analysis a research method that adds, integrates and interprets sets of research works and this is done by including only empirical studies that have quantitative findings presenting descriptive and inferential statistics within (Lipsey & Wilson, 2001). It is a statistical approach allowing researchers to formulate inferences on a larger population of various studies. It occurs by comparing and synthesizing results from empirical studies where individual research studies are the actual unit of the analysis (Card, 2012). Meta-analysis is often used in order to provide and filter highly contributing variables among all variables employed in collected studies. It is used for measuring effect sizes in order to yield objective and reliable outcomes (Kavale, 2001). Often, studies collected for meta-analysis are experimental designs utilizing experimental and control group setting (Hall & Burns, 2018). This type of analysis further permits the researchers to aggregate results in order to yield overall mean effect size in studies collected. The advantages of meta-analysis include 1) better parameter estimates, 2) evaluation and assessment of outcomes in multiple domains across, 3) moderator analysis by each variable, 4) minimizing error and bias (Hwang, 2015).

Jung and Choi (2019) meta-analyzed on phonics intervention effectiveness for underachieving elementary school students learning English as a Foreign Language. From 19 studies with 80 effect sizes were analyzed with influential moderators. From this meta-analysis, the mean effect size for English achievement was 1.23 and the mean effect size for affective

measures was 1.09 in hedge's *g*. Intervention duration, frequency of the week, a total number of sessions were shown to have moderated the effectiveness of phonics intervention to underachieving Korean students. In this study, upper elementary graders showed larger effect size for achievement and mid-grade students showed larger effect size for affective measures; however, the difference was not statistically significant. In terms of intervention method, interactive instructional method -neither bottom-up not top-down approach- had the largest effect size and the difference was significant. Session time duration from 40–60 minutes did not moderate significantly. Dependent variables of achievement measure subsets- letter recognition, phonemic awareness, phonics, fluency, vocabulary, reading comprehension and combination type did not play as a significant moderator in the study.

Lee and Chang (2017) meta-analyzed 55 experimental studies of the instruction effectiveness on low achieving elementary school students and found that reading intervention showed the largest effect size of 1.78 followed by technology usage for 1.61. When the intervention lasted for 9-16 weeks, the effect size was the largest (Cohen's d=1.577) for the duration. The result of meta-analysis was that the overall mean effect size of English language instruction for low achievers was 1.319, reflecting a very large effect.

At the middle school level, Lee, Chang, and Jeon (2017) aimed to investigate the overall effects of English language instruction to the underachievers in the secondary school through a quantitative meta-analysis, and drew implications for effective remediation. In this study, 59 experimental studies were collected and analyzed in terms of 7 moderating variables such as publication type, school type, region, treatment period, and dependent variables like linguistic and affective measure domain. The results showed that the overall mean effect size of English instruction with the struggling readers in secondary school was large (d=1.240). It was

also found that studies conducted in small cities had a larger effect than those in the metropolitan areas. Among dependent variables, the effect size in grammar was larger than in other linguistic skills and knowledge; the effect sizes in learning attitude and students' interest were larger than other affective measures.

Wanzek et. al. (2009) had a synthesis of research on reading interventions for students with reading difficulties and disabilities in 4^{th} and 5^{th} grade. 13 studies of treatment/comparison designs and 11 of single group or single subject designs concluded that 24 studies had high effects for comprehension intervention in researcher-developed measures. Word recognition intervention yielded small to moderate effects. An additional meta-analysis study on the effects of RTI Tier II reading intervention done by Wanzek et. al. (2016) found that effects of Tier II reading intervention results with 72 studies had the largest effect sizes for standardized language and comprehension measures (g=1.02) there were no significant differences in effects relating to intervention type, instructional group size, grade level, intervention implementer, or the number of intervention hours.

2.6. Rationale for the Study

The purpose of the study is to meta-analyze and extend previous English intervention meta-analyses by considering all types of struggling readers in either general education or special education setting. In addition to analyzing effectiveness of intervention outcomes and intervention methods on those students, related potential moderators as sub-groups are also analyzed to find any contributors to effective intervention. Because English is introduced at the elementary school level as a mandated school subject for most second language learners as well, elementary school students are the targeted sample population for the study's purpose. In this meta-analysis, the focus is not only to find effect

sizes of interventions, it also intends to find existing relationships and propensities between sub-groups or moderators.

In particular, this study aims to seek struggling readers in general education, special education, ELLs, EFLs, and otherwise low-achievers. In doing so, operational definitions, clear coding procedure and explanations are included as transparency is essential for the readers' assessment of the credibility on such meta-analytic findings (Harwell & Maeda, 2008). In doing so, the overall mean effect was analyzed and interpreted with heterogeneity test between studies, followed by potential moderator analyses in the order of publication type, grade level, learner type, type of outcome, outcome measure domains, intervention frequency as part of the study contextual characteristics. Then, meta-regression and publication bias tests were added for additional interpretation. Through this study, the types of learners, various findings in outcome measures, and types of effective intervention are especially sought for better practice in education.

2.7. Definitions of the Terms in Meta-analysis

The terms needed to be read and interpreted in this study are follows:

• Effect size: defined "as an index of the direction and magnitude of association between two variables and may include differences between groups, correlation between two variables, and contingencies between two dichotomies" (Card, 2012, p. 87). The effect size statistic represents quantitative findings to conduct comparisons and analyses across studies (Lipsey & Wilson, 2001). The effect size allows its standard error to be calculated and give more weight to studies that have small standard errors than those with large standard errors or less precise estimates (Card, 2012).

• Mean effect size: the most important index of central tendency in a

meta-analysis is mean effect size. It allows researchers to describe the typical effect sizes for a particular study. "The mean effect size is calculated by computing the product of each study's effect size by its weight, summing these products across studies, and dividing this value by the sum across the studies" (Card, 2012, p. 181).

• Heterogeneity test: involves calculating the Q-value and shows the amount of heterogeneity of effect sizes across studies. An index of the magnitude of heterogeneity, is used to determine the percentage of variability among effect sizes. I^2 of 25% is considered small, 50% medium, and 75% large effect sizes (Card, 2012).

• Fixed-effect model: this statistical model assumed that all the studies have a single effect in common (Borenstein, et al., 2009; Card, 2012).

• Random-effects model: this statistical model applies that "the true effects in the studies are assumed to have been sampled from the distribution of true effects" (Borenstein et al., 2009, p. 74) The random-effects model allows the researcher to be able to generalize the findings to the general population.

• Publication bias: refers to the possibility that studies, which did not find statistically significant effects, are more likely to be unpublished than studies that reported significant, positive effects. This screening is usually done because published literature may not be wholly representative of the studies conducted on a topic and can yield a stronger overall effect size than if all studies had been included as part of a meta-analysis (Card, 2012). Methods to manage publication bias include moderator analyses, funnel plots, and Trim-and-Fill method (Card, 2012).

Chapter 3. Method

3.1. Search and Screening Procedure

The present study aimed to synthesize data obtained from

available studies, both published peer reviewed journal articles and unpublished dissertations from 2000 to up to date. In actual searching, only those studies from 2003 to 2019 on English language interventions for struggling learner population of elementary school students in both general and special education setting were included. All studies were sought through (a) an electronic database search, (b) a hand search, (c) a reference list search for included studies, and (d) a review of previous systematic reviews search. The following databases were searched for the studies written in Korean: Research Information Sharing Service (RISS), Google Scholar (Korean), National Assembly Library (Korean). The following databases were searched for the studies written in English: Educational Resources Informational Clearinghouse (ERIC), PsycINFO, ProQuest Dissertations, Google Scholar online. The following search terms and keywords were used for searching: "English struggling learn*," "English struggling read*," "English low-achievers," "underachieve* in English," "Elementary English intervention," "English intervention for at-risk," "English intervention for learning difficult*," "English reading intervention for learning dis*, reading dis*," "English for special edu*" "intervention of, for low performance in English," "English language learner/ELL intervention," "English as a second language/ESL intervention," and "learners of English as a foreign language/EFL." The references in published meta-analytic studies and the articles reviewed were also examined. Along with the ancestral searches revisiting reference lists, additional hand search of major journals in special and elementary education were done through: *Journal of Special Education, Remedial & Special Education, Reading and Writing quarterly, journal of Specific Learning Disabilities (Korean), Journal of Foreign Language (Korean) and English Literature and Study (Korean)*. Citations and abstracts identified using these search processes were inspected and any duplicates screened were eliminated. If there were articles published for the dissertations on the same content, the published

articles were chosen over the dissertations. The search was completed in October, 2020.

3.2 Title and Abstract screening and inclusion criteria

The titles and abstracts retrieved were screened for further suitability using the following criteria:

1. The studies that assessed the effects or effectiveness of English language arts intervention, employing English intervention measures, and measurable dependent variables on the constructs of English achievement in sub-construct domain (e.g., phonological awareness, fluency, writing, listening comprehension) and/or affective measure domain (e.g., learner's confidence, motivation, attitude).

2. The studies that had a pretest-posttest control group design that included (a) a randomized controlled trial or (b) a quasi-experimental design (i.e., students not randomly assigned to groups), studies that reported pre-posttest gains within a one (single) group, or studies that reported posttest results only from control and experimental group. Any studies of qualitative articles and studies with single-subject/single-case designs were excluded "because there is no known statistical procedure for valid combination of single-subject and group design studies" (Gersten, Chard et al., 2009, p. 1204).

3. The studies had English intervention on elementary school age students (1^{st} to 6^{th} grade) with the learners' categories of struggling readers, including underachievement, low-achievement, populations of specific learning disabilities, English as a Second Language/English Language Learners (ESL/ELL), and English as a Foreign Language (EFLs). Any English intervention studies on kindergarten-level children, middle school, high school or college students or other adult learner categories were excluded

for this meta-analysis.

4. All types of intervention programs, instructions, and outcome measures were included for English intervention.

5. Studies that were published in English using any of non-English language intervention programs or outcome measures not related to English language were excluded.

6. Studies published for the years of 2000-2020 were included.

3.3. Full-text Screening

The studies after the title and abstract screening underwent more thorough screening for suitability. In addition to the criteria used in the title and abstract screening procedure, the studies were further assessed according to the following criteria in the text:

1. Studies reported data needed to calculate effect size.

2. If not random assignment or true experiment study, then equivalence on key measures at pretest was established in the studies. The control group/pretest phase had either: (a) no intervention or (b) typical instruction as usual before the treatment.

3. Studies conducted with mixed elementary and middle school grade students (e.g., 5^{th} to 9^{th} graders together in the intervention group) were excluded if middle and elementary grades were not possible to disaggregate for the elementary school level analysis only. Similarly, the studies whose participants' grades of kindergarten-1^{st} grade were excluded if grade 1 alone could not be disaggregated for analysis.

4. All learner types of participants were included as long as the authors reported standards or definitions of "struggling" readers for their participants. For this reason, any participant groups of English as a Second Language (ESL)/English Language Learners (ELL), English as a Foreign Language (EFL), Underachievers, Low-achievers were also included as long as they met the

definitions of struggling readers.

5. If the classroom type (i.e., general education or special education inclusion classroom) was not discernable, studies were excluded. Any studies with an extreme level of special education needs-based small classroom, (e.g., students with severe autism, blind and speech impairment) were excluded since those population groups go beyond "struggling" readers.

At the initial stage of screening, a total of 110 studies were extracted from database, hand and ancestral search. After duplicates removed and title and abstract screening, studies were narrowed down to 87. After removing foreign language intervention, measurable dependent variables being other than English language achievement, single-subject design studies, 32 studies were eliminated, leaving 55 studies. Additional 20 studies were further eliminated after full-text screening for studies with inappropriate grade range/elementary and other age groups that cannot be disaggregated, missing data, studies with outlying influential points after initial meta-analysis. A complete process with each study number for the studies found and included in this meta-analysis is in Figure 1 below. The process includes both Korean and English studies for English language intervention as it was the author's intention to find the results of English intervention effects from studies written in either language.

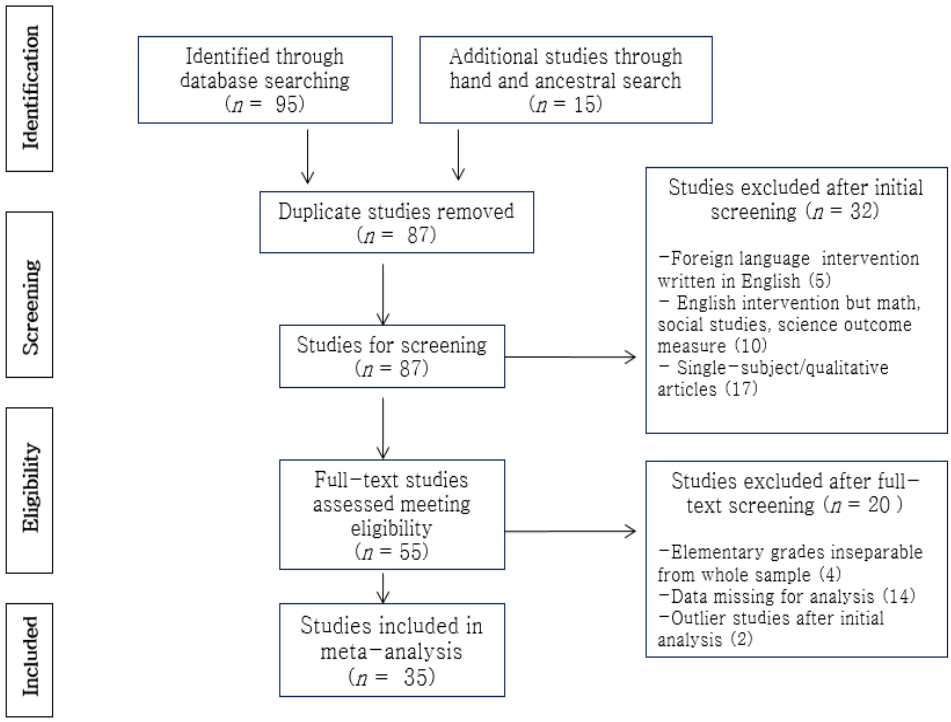


Figure 1. PRISMA flow diagram for meta-analysis

After full screening and initial analysis to screen for any outliers, a final of 35 studies were left. Of 35 studies, 22 studies were published in Korean and 13 studies were published in English.

Although some research suggest that analyses of pre-posttest gains within a single group tend to produce significant positive biases in results when compared to the other research designs employing matched control-group (Borman et al., 2003), with the nature of difficulty in obtaining experimental-control group studies in some cases (e.g., schools on site not allowing control group assignment for the researchers as evaluating on control group without providing intervention violates their school ethical code of conduct, the nature of treatment-control groups is difficult when learning English instruction is a foreign language, (Lee & Chang, 2017), single group pre and posttest study designs were also included for the analyses to follow.

3.4. Coding Characteristics of Study

The coding process in a meta-analysis is used to determine which relevant information needs to be extracted from each study (Lipsey & Wilson, 2001). Coding should be guided by individual research questions but also should include specific aspects of the studies that needed to be considered to help draw generalizations, such as characteristics of the sample, measurement, design and source (Card, 2012).

Source characteristics included factors such as the number of the study or ID number, author, title, year of the study, and publication type. Lipsey and Wilson (2001) recommend coding variables separately by adding a decimal to the identification number. For this meta-analysis, the studies with additional independent sub-studies were coded separately this way. The author added a decimal to the study identification number. Each included study was coded for categories of (a) participant characteristics, (b) study design characteristics, (c) outcome measure outcomes (dependent variables), (d) intervention type and components (independent variables), and 5) other study contextual characteristics.

Rationales for coding:

Participants' characteristics and demographic information included participant's grade level, classroom type (i.e., special education or general education setting), learner type (i.e., EFL, ELL, low-achievers, underachievers) defined by the authors in the studies. Regarding the intervention, all individual names of interventions provided, their type, interventionist, minutes spent on the intervention, weekly frequency, total session number of interventions for duration were included for coding. Regarding outcome measures as dependent variables, all individual outcomes, outcome measure type as in affective (e.g., motivation, confidence) and in achievement (e.g., fluency, phonics) were included for coding.

Affective and achievement measures were also previously used by meta-analysis (Lee & Chang, 2017; Jung & Choi, 2019) for the reason that the English underachievers be analyzed by cognitive (intelligence), affective, and environmental factors that must be accounted together for their underachievement (Lee & Chang, 2017). Within affective outcome measure, any similarly defined dependent variables within a study that were simply inversely coded (e.g., "fear" and "confidence") only one measure ("confidence") was included to reduce duplicate affective measures.

If the study uses more than one assessment, Lipsey and Wilson (2001) suggest coding each measure separately to allow for a more comprehensive "empirical examination of the relationship between the particular ways in which a construct is operationalized and the nature of findings from different studies" (Lipsey & Wilson, 2001, p. 78). As the coding steps recommended by Vanchu-Orosco (2012), the following steps were included to code and classify studies in the initial stage: 1) create the codebook with initial set of codes, 2) read five articles with the initial codebook and revise as new information comes to attention, 3) code three or more articles with the revised codebook and continue to revise again, 4) create coding forms and coding manual to accompany the codebook, and 5) code all the remaining studies (Vanchu-Orosco, 2012). The code book in in Table 1 and Appendix 1.

Table 1

Codebook with Descriptors

Category	**Definition of Descriptors**
Participants	
Grade	The current grade level in elementary school
Grade level range	Lower (1st and 2nd); Mid-upper (3rd to 6th)
Learner type	Underachiever (UA); Low achiever (LA); English as Foreign Language (EFL); English Language Learners (ELL), English as a Second Language (ESL); Specific

	Language Impairment (SLI)
Classroom type	Special education (inclusion), General education, Both
Country / City / Region	Name of country, city, region of the intervention location if stated
Intervention	
Intervention components	All names of intervention as independent variables (e.g. repeated reading, phonics, direct teaching)
Intervention component type	All intervention types into 1) content-based, 2) learner approach-based, or 3) teacher/instructional method-based
Interventionist	Teacher; Researcher; Bilingual teacher; Teaching Assistant or paraprofessionals (TA); Research Assistant (RA)
Minutes	Actual minutes spent for intervention
Weekly frequency	Sessions per week
Duration	Total number of weeks
Session number	Total sessions of intervention
Outcome measures	
Outcome measure domain	Measured outcome domains (e.g. fluency, motivation)
Outcome measure type	All measures were group into 1) affective (e.g. motivation, attitude), or 2) achievement type (e.g. fluency)
Outcome type	Researcher-developed or standardized
Study level	
Publication information	Journal; Dissertation
Study design information	Pre and posttest control group; posttest only control group; One group (single group) pre and posttest

3.5. Coding reliability

Full texts of 55 articles were inspected by two authors. Initial agreement was 91 percent, and any disagreements occurred

were openly discussed and reached consensus and settled through discussions on the coding suitability of the studies.

For the coding process, the first author served as the primary coder, coding of all studies for participants, methodological, and study characteristics. Then, a doctor with Doctor of Education degree in special education and elementary education was trained in the initial stage and later independently coded a randomly selected 25 percent of the total studies in the final stage. Coding reliability came out to be 94 percent. The calculated average interrater agreement (IRA) using the following formula:

IRA= agreements / (agreements + disagreements) x 100

Initial training for coding to reach interrater agreement progressed through three stages: (1) co-coding, (2) individual practice coding, and (3) formal coding. An iterative process for clarification and rewriting of codes was part of the process at each stage. In stage 1, co-coding, the second coder was provided with a copy of the coding manual created by the first author. Then a training meeting was scheduled whereby the first author clarified any questions about the coding manual and provided examples. Next, those two raters coded 5 studies together. After additional discussion and clarification of the coding manual, each rater independently coded as in stage 2, and IRA was assessed after the second coder finally coded random 25 percent of the studies by the end. Any information that was not provided by the authors of the study was indicated as Not Specified (NA) or Not Reported (NR) on the coding sheet to the right column. Coding reliability result on the second round was 94 percent. Coding results with disagreement outcomes were discussed until full consensus was reached.

3.6. Data Analysis

The effect size for a meta-analysis allows the researcher to be able to estimate the strength of a relationship between the independent variable and the dependent variable (Borenstein et al.,

2009; Gliner, Morgan, & Leech, 2009). Effect size is defined as "an index of the direction and magnitude of association between two variables and may include a correlation between two variables, differences between two groups, and contingencies between two dichotomies. An important criterion for effect size is that it must be possible to compute or approximate its standard error. The standard error allows a researcher to give more weight to studies that have small standard errors than those that provide less precise estimates" (Card, 2012, p. 87). The summary statistic for calculating effect sizes for the studies chosen for this meta-anaysis was the standardized mean difference effect size (Lipsey & Wilson, 2001). In fact, standardized mean difference applies to "comparisons between means of outcome measures for experimental and control groups in treatment effectiveness research" (Lipsey & Wilson, 2001, p. 48).

All the analyses were conducted using the software, Comprehensive Meta-Analysis (CMA) version 3.0 software Borenstein, Hedges, Higgins & Rothstein, 2009). Visual graphs from CMA were used and displayed in this paper to visually inspect the data for patterns in effect size magnitude, such as distribution of effect sizes, forest plot with 95% confidence intervals, funnel plot for publication bias, and scatter plot for variability distribution. In meta-analyses, "effect sizes can be viewed as the dependent (or criterion) variables and the features and characteristics of the study as independent (or predictor) variables (Cooper et al., 2009, p. 13). Specifically, hedge's g was calculated to empower low sample numbers in some studies extracted and because Cohen's d "has a slight bias, tending to overestimate the absolute value of δ in small samples" (Cooper et al., 2009, p.226), the Hedge's g has been applied throughout in this meta-analysis to produce an unbiased estimate and the formula for Hedge's g is as follows:

$$d = \frac{\overline{Y}_E - \overline{Y}_c}{s}$$

$$s = \sqrt{\frac{(n_E-1)S_E^2+(n_c-1)S_c^2}{n_E+n_c-2}}$$

$$J = 1 - \frac{3}{4df-1}$$

$$g = d \times J$$

The precision of all effect size estimates was established by calculating the standard error (SE) of the mean (i.e., square root of the sum of the inverse variance weights) and by using it to create a 95% confidence interval around the mean (Cooper et al., 2009):

$$\text{Lower limit of 95\% } CI = \bar{g} - 1.96 * SE$$
$$\text{Upper limit of 95\% } CI = \bar{g} + 1.96 * SE$$

Estimation of heterogeneity

To assess consistency across studies, heterogeneity of the distribution of effect sizes was assessed using the Q statistic and the I^2 statistic. The Q statistic has an approximate chi-square distribution with $k-1$ degrees of freedom, in which k is the number of independent effect sizes (Cooper et al., 2009). "Significant Q statistics indicate the existence of heterogeneity and I^2 describes the percentage of total variation across studies that are considered truly due to heterogeneity rather than by chance" (Higgins, et al., 2003, p.557). I^2 was calculated using the formula below:

$$I^2 = 100\% * \frac{Q-(k-1)}{Q}$$

Suggested guidelines for interpreting the value of I^2 is as follows: I^2 = 25% suggests a small amount of heterogeneity, I^2 = 50% suggests medium heterogeneity and I^2 = 75% suggests large heterogeneity (Cooper et al., 2009; Higgins et al., 2003).

For this meta-analysis study, a random-effects model was used as this model is considered more appropriate than a fixed-effect model. Within a random-effects model, "studies under

synthesis can be viewed as representative of a larger population or universe of implementations of a treatment" (Cooper et al., 2009, p. 306). The primary assumption of the fixed-effect model is that "one true effect size underlies all the studies in the analysis, and that all differences in observed effects are due to sampling error" (Borenstein, Hedges, Higgins, & Rothstein, 2010, p.97). Therefore, a random-effects model is deemed more suitable for the studies in the research areas of social sciences education as the research findings tend to vary along multiple dimensions (Borman et al., 2003; Borenstein, et al., 2010; Cooper et al, 2009). That is, given that the included studies represent different types of samples, different methodological features, and varied study characteristics, the expectation is that a distribution of effect sizes that is due to more than sampling error alone will be revealed. As such, the total variance of the distribution of effect sizes consists of both within- and between-studies error variance (Borenstein et al., 2010). Hence the random-effects model was used for better generalizability to include unexplained heterogeneity.

Sub-group/Moderator analysis

The presence of heterogeneity in the overall mean effect size estimate indicates that the possible existence of variables (e.g., study features, publication type) exist and serve to moderate the effects (Borenstein et al., 2010; Huedo- Medina et al., 2006). Moderator analysis was done by disaggregating study effect size estimates and grouping them into appropriately designated sub-groups or categories. The chi-square test of homogeneity (Q statistic) of effect sizes and the associated I^2 statistics were used to determine the significance and magnitude of between (Q_B) and within (Q_W) group differences in the mean effect size for each potential moderator variable. Because these tests traditionally tended to have low power to detect departures from homogeneity (Cooper et al., 2009; Mittlböck & Heinzl, 2006), some researchers (e.g., Petitti, 2001) have recommended selecting a statistical

significance level of α = .10. However, for this study, the traditional significance level of α = .05 was chosen so as to avoid increasing Type I error for taking a risk of a false positive as Higgins et al. (2003) stated.

Potential outlier/Influential points

Harwell and Maeda (2008) recommend identifying potential outliers in meta-analysis and "performing key analyses after temporarily excluding suspect studies and examining the similarity of findings with and without these studies" (p. 424). Zheng et al. (2013) reported removing outliers, which they defined as "effect sizes lying beyond the first gap of at least one standard deviation between adjacent effect size values in a positive direction" (p.101). Extreme outliers are further defined as three interquartile ranges below the first and above the third interquartile ranges (Tukey. 1977). If four times larger than the overall mean effect size, then WWC states that it is considered too large for the experimental group. As such, in the present meta-analysis, the identified potential outliers, or influential points, two studies were eliminated because their individual effect sizes were found to be larger than g=3.6, four times larger in the positive direction than the overall mean effect size (g=0.903) when all 37 studies were initially included for the first analysis.

Chapter 4. Results

4.1. Descriptive Statistics

A total of k=150 final independent effect sizes were extracted from 35 studies. Table 2 below and Appendix 2 illustrate the summary of included studies. The final set of studies selected were published from 2003 to 2019, of which 11 studies were from

2003 to 2009 and 24 studies were from 2010 to 2019, making that 68.6% of the total number of studies included in this study were published after 2010. Among those studies, 19 studies were dissertations, and the remaining 16 studies were published in peer-reviewed journals, making that 45.7% of the total number of studies were journal articles. Publication type was analyzed as a potential moderator variable and the results are reported later in this chapter.

In this chapter, the results of the meta-analyses for the research questions are described, including the summary of the studies, overall mean effect size, moderator analysis and meta-regression analysis. As a supplemental means of resolving dependency issues within studies, shifting unit-of-analysis was implemented for sub-construct grouping of the variables (Cooper, 1998; Graham & Hebert, 2011). Later, a funnel plot was utilized in order to assess possible publication bias and adjustment. Findings are discussed later in this chapter.

The summary of the included studies is in Table 2 below.

Table 2

Summary of Studies Included

	Study Name	Participants	Learner type	Intervention	Interventionist	Intervention frequency (per week; total session #)	Outcome measures	Affective measure descriptions	Study Design
1	Kang, B. (2008)*	N=8; Gr.6;	UA (EFL)	Phonics	Researcher	3/wk; 30 sessions	PA, FL, Affective	Learning attitude	One group pre-posttest
2	Kang, S. (2010)*	N=15, Gr.4(5), 5(5), 6(5)	UA (EFL)	Phonics/Eng song	Researcher	2/wk; 8-15 sessions	LC, RC, Affective	Motivation, confidence	One group pre-posttest
3	Kim, Y.J. (2013)*	N=5, Gr.5(5)	UA (EFL)	Phonics	Researcher	3/wk; 12 session	SPK, WR, Word ID,		One group pre-posttest
4	Kim, Y.K. (2014)*	N=5, Gr.6(5)	UA (EFL)	Word list/ Phonics	Researcher	2/wk; 46 sessions	LT, FL, PA, VOC, RC, FL		One group pre-posttest
5	Kim & Jeong (2010)	N=5, Gr.6(5)	UA (EFL)	Multisensory/ Repetitive	Researcher	3/wk; 6 sessions	CGN	Cog abilities	One group pre-posttest
6	Kim, J. (2013)*	N=11, Gr.6(11)	UA (EFL)	Phonics	Researcher	1/wk; 15 sessions	LT, PA, SPL	Letter order, Upper & lower case	One group pre-posttest
7	Kim & Park (2015)	N=43, Gr.4(22), Gr.5(21)	UA (EFL)	Feeling-based	Researcher	2/wk; 28 sessions	RC, Affective	learning time, satisfaction rating, self-driven, self-esteem, motivation,	One group pre-posttest

								competitiveness, attitude	
8	Kim, H. (2014)*	N=5, Gr.5(5)	UA (EFL)	Eng song	Researcher	5/wk; 28 sessions	FL		One group pre-posttest
9	Nam, H. (2010)*	N=10, Gr.6(10)	UA (EFL)	Shared reading	Researcher	5 wk; 40 sessions	FL		One group pre-posttest
10	Moon (2015)*	N=12, Gr.5(6), Gr.6(6)	UA (EFL)	Role play/Repetitive	Researcher	2 wk; 48-60 sessions	SPK		One group pre-posttest
11	Song (2015)	N=9, Gr.4(3), Gr.5(3), Gr.6(3)	UA (EFL)	Phonics	Researcher	1 wk; 12 sessions	PA		One group pre-posttest
12	Shin (2017)*	N=22, Gr.5(22)	UA (EFL)	Phonics/Strategies	T & R	2/wk; 23 sessions	CGN	cog. flexibility	One group pre-posttest
13	An (2016)*	N=8, Gr.4(8)	UA (EFL)	Supplementary WB	Researcher	2/wk; 8 sessions	Affective	Changes in mindset	One group pre-posttest
14	Uhm & Kim (2014)	N=8, Gr.5(8)	UA (EFL)	Phonics/Voc	Teacher	1/wk; 16 sessions	FL, WR, Affective	Motivation, confidence, participation	One group pre-posttest
15	Cho* (2015)	N=4, Gr.3(4)	UA (EFL)	Shared reading	T & R	3/wk; 42 sessions	FL, VOC, SPK, Affective	Motivation, confidence, participation	One group pre-posttest
16	Seaman (2015)*	N=117, Gr.3-4(117)	ELL	Lab-instruction	Teacher	5/wk; 260 sessions	FL		One group pre-posttest

17	Mason et al. (2006)	N=9, Gr.5, LA (5) + LD (1), LD/SLI (1), EBD (2)	ELL	Strategy	Researcher	3/wk; 15 sessions	SPK, WR	Oral retell	One group pre-posttest
18	Cheon (2003)*	N=64, Gr.5(64)	UA (EFL)	Shared reading	Researcher	1/wk; 9 session	FL, Affective	Motivation, confidence	Pre-posttest control group
19	Heo (2006)*	N=45, Gr.5(45)	UA (EFL)	Phonics	Researcher	2/wk; 20 sessions	FL, Affective	Motivation, confidence	Pre-posttest control group
20	Kang, A. (2016)*	N=20, Gr.4(20)	UA (EFL)	Phonics	Researcher	3/wk; 36 sessions	RC, LT, VOC, PA, Affective	Motivation, confidence	Pre-posttest control group
21	Kim, W.J. (2019)*	N=16, Gr.5(16)	UA (EFL)	Growth mindset	Researcher	1/wk; 8 sessions	Affective	Learning belief, determination	Pre-posttest control group
22	NamGoong. S. (2012)*	N=18, Gr.5(18)	UA (EFL)	Eng song	Researcher	1/wk; 14 sessions	LC, Affective	Motivation, confidence	Pre-posttest control group
23	Bae (2015)*	N=56, Gr.5(56)	UA (EFL)	Small group collab	Teacher	2/wk; 12 sessions	SPK, Affective	Motivation, learning attitude	Pre-posttest control group
24	O'Connor et al. (2007)	N=37, Gr.2-4(37)	LA	Repeated/Continuous	Teacher	3/wk; 42 sessions	FL, RC, Word ID, PA		Pre-posttest control group
25	Oostdam et al. (2015)	N=126, Gr.2-4(126)	UA	Repeated	Other	4/wk; 48 sessions	FL, RC, VOC, Affective		Pre-posttest control group
26	Vaughn et al. (2006)	N=46, Gr.1(46)	ELL	Systematic/Explicit	Bilingual Teacher	5/wk; 100 sessions	LT, PA, LC, FL, RC		Pre-posttest control group

27	Guthrie et al. (2009)	N=62, Gr.5(62)	LA	Concept-orientedrd instruction	Teacher	3/wk; 36 sessions	RC, VOC, FL		Pre-posttest control group
28	Harn et al. (2008)	N=54, Gr.1(54)	LA	Intensive reading	Other	5/wk; 120 sessions	RC, PA, FL		Pre-posttest control group
29	Hatcher et al. (2006)	N=77, Gr.1(77)	LA	Small group collab	Teacher	2-3/wk 50 sessions	PA, LT, FL		Pre-posttest control group
30	Martens et al. (2007)	N=30, Gr.2(30)	LA	FL program	Other	3 wk; 18 sessions	FL		Pre-posttest control group
31	Mason et al. (2012)	N=9, Gr.4(9), LA(5), LD(1)	LA	TWA (Thinking strat)	Other	1-2 wk; 10 sessions	WR, RC, STN, VOC		Pre-posttest control group
32	Vadsay & Sanders (2008)	N=119, Gr.4-5(119)	ELL	FL program	Teacher	4 wk; 72 sessions	VOC, FL, RC		Pre-posttest control group
33	Jeong and Kim (2017)	N=12, Gr.5(12)	UA (EFL)	Multiple Intelligence	T & R	2/wk; 26 sessions	RC, FL, Affective	Motivation, confidence, participation	Posttest only control group
34	Wagner (2011)*	N=29, Gr.5-6(29)	LA	Immediate instrc.	Researcher	2/wk; 6 sessions	VOC, FL, RC		Posttest only control group
35	Kaniuka (2010)	N=85, Gr.3-5(85); LD(9)	LA	Remedial reading program	Researcher	NR	Affective	Attitude, self-esteem	Posttest only control group

Note. *Unpublished dissertation; UA = Underachiever; LA = Low achiever; EFL = English as a foreign language, ELL = English language learners; LD = (Specific) learning disabilities; T & R = teacher and researcher; PA = phonics; FL = fluency; VOC = vocabulary; RC = reading comprehension; LC = listening comprehension; SPK = speaking, SPL = spelling; WR = writing; Word ID = word identification; LT = letter detection, letter recognition; CGN = cognitive achievement; T & R = teacher and researcher, Affective = affective components

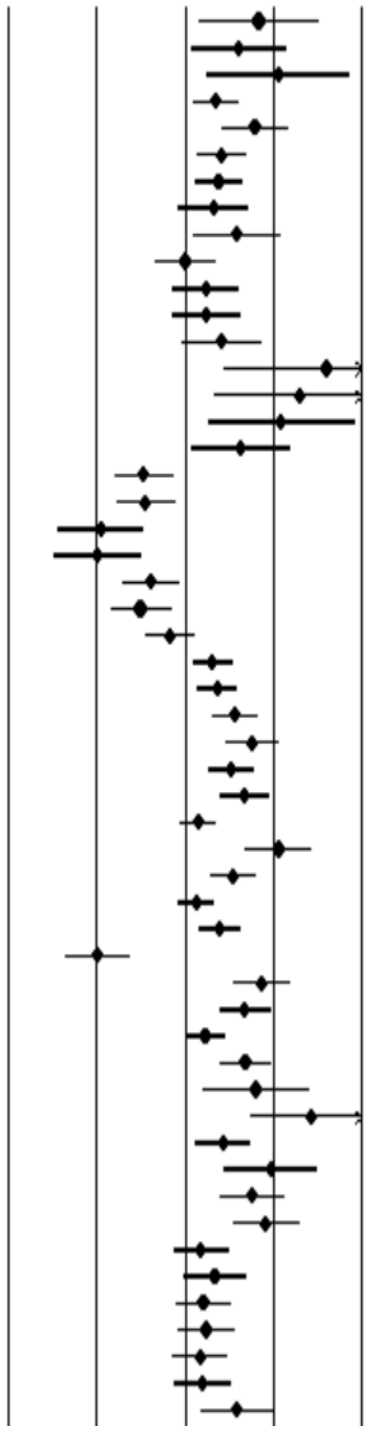
4.2. Meta-analysis: Overall Mean Effect of English Intervention

If the effect sizes are four times larger than the overall mean effect, then WWC states that it is considered too large of an effect for the experimental group. When these effect sizes were included in the analyses, the critical assumption of normality was violated, and there was significant heterogeneity within several of the potential moderator variables, both of which threaten the validity of conclusions (Cooper et al., 2009). As such, those studies were eliminated after the first meta-analysis as described in the earlier section.

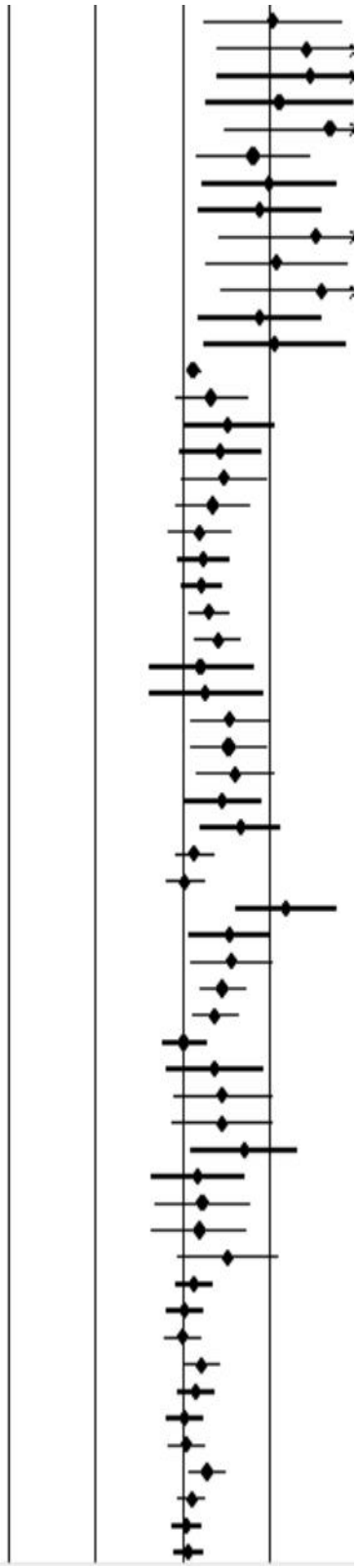
Within one group pre-posttest studies, 73 effect sizes were yielded. Within pre-posttest treatment control group studies, 67 effect sizes were yielded. Within posttest only treatment control group studies, 10 effect sizes were yields, totaling k=150.

Figure 2 presents a distribution of the 150 independent Hedges' g effect sizes extracted from the 35 included studies with their 95% confidence intervals. The mid-point of the circles represents the point estimate of each effect size, and the width of the line shows the 95% chance that the true effect will be within that range. Variation in the width of the confidence intervals represents variation in the precision of effect size estimates. The mean effect size was Hedge's g=0.655 [SE=0.054; CI_{95}=0.550, 0.761].

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Kang, B. 1.1	1.665	0.692	0.479	0.309	3.021	2.406	0.016
Kang, B. 1.2	1.204	0.560	0.314	0.107	2.302	2.151	0.031
Kang, B. 1.3	2.111	0.830	0.689	0.484	3.737	2.543	0.011
Kang, S. 1.2	0.687	0.274	0.075	0.149	1.225	2.503	0.012
Kang, S. 1.3	1.578	0.378	0.143	0.838	2.318	4.178	0.000
Kang, S. 1.4	0.838	0.288	0.083	0.273	1.403	2.909	0.004
Kang, S. 1.5	0.758	0.281	0.079	0.208	1.308	2.700	0.007
Kim, Y.J. 1.2	0.636	0.410	0.168	-0.169	1.440	1.549	0.121
Kim, Y.J. 1.3	1.156	0.511	0.262	0.153	2.158	2.260	0.024
Kim, Y.J. 1.4	0.000	0.358	0.128	-0.701	0.701	0.000	1.000
Kim, Y.J. 1.5	0.462	0.386	0.149	-0.296	1.219	1.195	0.232
Kim, Y.K. 1.1	0.476	0.401	0.161	-0.309	1.262	1.188	0.235
Kim, Y.K. 1.2	0.825	0.466	0.217	-0.089	1.738	1.770	0.077
Kim, Y.K. 1.3	3.216	1.194	1.425	0.876	5.556	2.694	0.007
Kim, Y.K. 1.4	2.585	0.984	0.968	0.657	4.514	2.628	0.009
Kim, Y.K. 1.5	2.180	0.852	0.726	0.510	3.850	2.558	0.011
Kim, Jeong	1.244	0.571	0.326	0.126	2.363	2.180	0.029
Kim, J. 1.1	-0.944	0.343	0.118	-1.617	-0.271	-2.748	0.006
Kim, J. 1.2	-0.903	0.338	0.114	-1.566	-0.239	-2.668	0.008
Kim, J. 1.3	-1.917	0.494	0.244	-2.885	-0.948	-3.877	0.000
Kim, J. 1.4	-1.992	0.508	0.258	-2.987	-0.997	-3.923	0.000
Kim, J. 1.5	-0.787	0.325	0.106	-1.424	-0.150	-2.422	0.015
Kim, J. 1.6	-1.018	0.353	0.125	-1.710	-0.326	-2.885	0.004
Kim, J. 1.7	-0.349	0.288	0.083	-0.914	0.216	-1.211	0.226
Kim, Park 1.1	0.621	0.226	0.051	0.179	1.064	2.752	0.006
Kim, Park 1.2	0.728	0.233	0.054	0.271	1.185	3.125	0.002
Kim, Park 1.3	1.110	0.265	0.070	0.590	1.629	4.188	0.000
Kim, Park 1.4	1.511	0.307	0.094	0.910	2.113	4.926	0.000
Kim, Park 1.5	1.033	0.258	0.066	0.528	1.538	4.006	0.000
Kim, Park 1.6	1.327	0.287	0.082	0.765	1.889	4.627	0.000
Kim, Park 1.7	0.283	0.210	0.044	-0.129	0.694	1.347	0.178
Kim, Park 1.8	2.098	0.377	0.142	1.359	2.837	5.562	0.000
Kim, Park 2.1	1.071	0.267	0.071	0.547	1.594	4.008	0.000
Kim, Park 2.2	0.250	0.213	0.046	-0.169	0.668	1.170	0.242
Kim, Park 2.3	0.780	0.242	0.059	0.305	1.254	3.222	0.001
Kim, Park 2.4	-2.007	0.374	0.140	-2.741	-1.274	-5.365	0.000
Kim, Park 2.5	1.730	0.340	0.115	1.065	2.396	5.094	0.000
Kim, Park 2.6	1.348	0.295	0.087	0.769	1.927	4.561	0.000
Kim, Park 2.7	0.451	0.221	0.049	0.017	0.884	2.038	0.042
Kim, Park 2.8	1.362	0.297	0.088	0.780	1.944	4.585	0.000
Kim, H.	1.583	0.615	0.379	0.377	2.789	2.573	0.010
Nam H.	2.845	0.699	0.488	1.475	4.215	4.071	0.000
Moon	0.846	0.319	0.102	0.220	1.472	2.650	0.008
Song	1.927	0.545	0.297	0.859	2.995	3.536	0.000
Shin 1.1	1.499	0.379	0.144	0.756	2.242	3.957	0.000
Shin 1.2	1.818	0.388	0.151	1.058	2.579	4.686	0.000
An 1.1	0.361	0.327	0.107	-0.280	1.001	1.103	0.270
An 1.2	0.668	0.356	0.127	-0.029	1.366	1.878	0.060
An 1.3	0.408	0.330	0.109	-0.240	1.055	1.234	0.217
An 1.4	0.469	0.335	0.113	-0.188	1.127	1.399	0.162
An 1.5	0.329	0.325	0.106	-0.307	0.966	1.014	0.311
An 1.6	0.404	0.330	0.109	-0.243	1.051	1.225	0.221
An 1.7	1.182	0.431	0.186	0.336	2.027	2.740	0.006



Uhm, Kim 1.3	2.067	0.816	0.666	0.467	3.666	2.532	0.011
Uhm, Kim 1.4	2.865	1.076	1.158	0.756	4.974	2.662	0.008
Uhm, Kim 1.5	2.951	1.105	1.221	0.786	5.117	2.671	0.008
Uhm, Kim 1.6	2.224	0.866	0.750	0.526	3.921	2.567	0.010
Uhm, Kim 1.7	3.394	1.254	1.572	0.936	5.851	2.707	0.007
Uhm, Kim 1.8	1.622	0.679	0.461	0.291	2.953	2.389	0.017
Uhm, Kim 1.9	1.990	0.792	0.627	0.438	3.542	2.513	0.012
Cho 1.1	1.766	0.722	0.522	0.350	3.182	2.444	0.015
Cho 1.2	3.078	1.147	1.317	0.829	5.327	2.683	0.007
Cho 1.3	2.153	0.844	0.712	0.500	3.807	2.552	0.011
Cho 1.4	3.200	1.188	1.412	0.871	5.529	2.693	0.007
Cho 1.5	1.762	0.721	0.521	0.348	3.176	2.443	0.015
Cho 1.6	2.134	0.838	0.702	0.492	3.776	2.548	0.011
Seaman 1.1	0.236	0.089	0.008	0.061	0.411	2.648	0.008
Mason et al. (2006) 1.1	0.659	0.426	0.181	-0.175	1.494	1.548	0.122
Mason et al. (2006) 1.1.2	1.053	0.542	0.294	-0.009	2.116	1.943	0.052
Mason et al. (2006) 1.2.1	0.864	0.483	0.233	-0.083	1.810	1.789	0.074
Mason et al. (2006) 1.2.2	0.935	0.505	0.255	-0.054	1.924	1.853	0.064
Mason et al. (2006) 1.3.1	0.692	0.434	0.189	-0.159	1.543	1.593	0.111
Mason et al. (2006) 1.3.2	0.387	0.366	0.134	-0.330	1.104	1.058	0.290
Cheon 1.1	0.490	0.310	0.096	-0.118	1.098	1.579	0.114
Cheon 1.2	0.429	0.241	0.058	-0.044	0.902	1.778	0.075
Cheon 1.3	0.611	0.248	0.062	0.125	1.097	2.465	0.014
Heo 1.1	0.805	0.270	0.073	0.276	1.333	2.983	0.003
Heo 1.2	0.411	0.618	0.382	-0.801	1.623	0.665	0.506
Heo 1.3	0.537	0.666	0.443	-0.768	1.841	0.806	0.420
Kang 1.1	1.086	0.461	0.213	0.181	1.990	2.353	0.019
Kang 1.3	1.057	0.460	0.211	0.156	1.958	2.299	0.022
Kang 1.4	1.199	0.468	0.219	0.281	2.117	2.560	0.010
Kang 1.5	0.908	0.452	0.204	0.023	1.794	2.011	0.044
Kang 1.6	1.322	0.477	0.227	0.387	2.256	2.773	0.006
Kim, W.J. 1.1	0.276	0.236	0.056	-0.187	0.738	1.168	0.243
Kim, W.J. 1.2	0.061	0.235	0.055	-0.399	0.522	0.261	0.794
Namgung 1.1	2.365	0.597	0.357	1.194	3.537	3.959	0.000
Namgung 1.2	1.087	0.484	0.234	0.139	2.036	2.246	0.025
Namgung 1.3	1.111	0.486	0.236	0.159	2.063	2.287	0.022
Bae 1.1	0.920	0.278	0.077	0.376	1.464	3.314	0.001
Bae 1.2	0.745	0.273	0.074	0.210	1.280	2.728	0.006
Bae 1.3	0.021	0.264	0.070	-0.496	0.538	0.079	0.937
Oconnor et al. 1.1	0.749	0.576	0.332	-0.381	1.878	1.299	0.194
Oconnor et al. 1.2	0.916	0.587	0.345	-0.235	2.066	1.560	0.119
Oconnor et al. 1.3	0.905	0.586	0.344	-0.244	2.054	1.543	0.123
Oconnor et al. 1.4	1.416	0.631	0.398	0.180	2.652	2.245	0.025
Oconnor et al. 1.5	0.339	0.558	0.312	-0.755	1.434	0.608	0.543
Oconnor et al. 1.6	0.452	0.562	0.316	-0.650	1.553	0.804	0.421
Oconnor et al. 1.7	0.376	0.559	0.313	-0.720	1.473	0.673	0.501
Oconnor et al. 1.8	1.042	0.597	0.356	-0.127	2.211	1.747	0.081
Oostdam et al. 1.1	0.268	0.219	0.048	-0.160	0.697	1.228	0.220
Oostdam et al. 1.2	0.047	0.218	0.047	-0.380	0.473	0.215	0.830
Oostdam et al. 1.3	0.002	0.218	0.047	-0.424	0.429	0.011	0.991
Oostdam et al. 1.4	0.433	0.220	0.048	0.002	0.865	1.968	0.049
Oostdam et al. 1.5	0.293	0.219	0.048	-0.136	0.722	1.338	0.181
Oostdam et al. 1.6	0.051	0.218	0.047	-0.376	0.477	0.233	0.816
Oostdam et al. 1.7	0.074	0.218	0.047	-0.353	0.501	0.340	0.734
Oostdam et al. 1.8	0.541	0.222	0.049	0.106	0.975	2.440	0.015
Oostdam et al. 1.9	0.197	0.173	0.030	-0.142	0.536	1.140	0.254
Oostdam et al. 1.10	0.069	0.173	0.030	-0.269	0.407	0.400	0.689
Oostdam et al. 1.11	0.134	0.173	0.030	-0.205	0.472	0.774	0.439



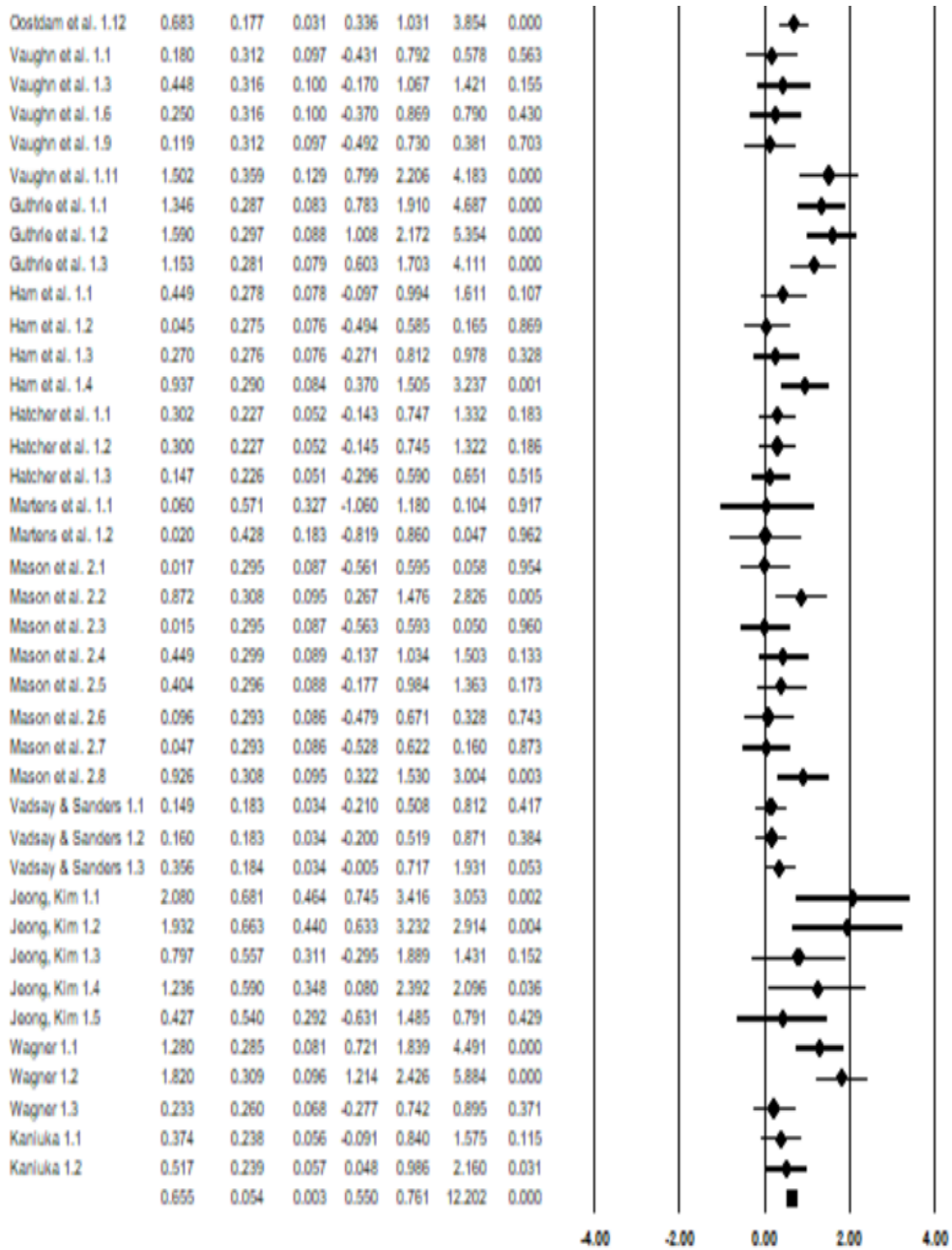


Figure 2. Hedges's *g* Effect Sizes in Forest Plot after Influential Points Removed

Next, homogeneity statistics were calculated to determine whether there are true differences underlying the results of the studies (i.e. heterogeneity) or whether the variation in the results is compatible with chance alone (i.e. homogeneity) (Higgins et al., 2003). When all 150 independent effect sizes were included, the chi-square test of homogeneity was significant (Q_B = 587.12, p < 0.001), and the magnitude of heterogeneity was I^2 = 74.62%, illustrating a large percentage of the total variation across, which would mean that those studies in this meta-analysis can be accounted for by genuine heterogeneity as opposed to chance alone (Higgins & Thompson, 2002; Higgins et al., 2003). The results are in Table 3 below.

Table 3

Summary of Heterogeneity of Overall Mean Effect Sizes

	Effect size and 95% confidence interval					Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
All	150	0.655	0.054	0.550	0.761	0.000
	Heterogeneity results					
	Q	df(Q)	*p*	I^2	T^2	*SE*
All	587.120	149	0.000	74.622	0.277	0.055

4.3. Meta-Analysis of sub-groups: Potential Moderators to Effect Size

The presence of heterogeneity in the overall mean effect size estimate holds moderator analyses to account for that variance (Borenstein et al.,2010; Huedo-Medina et al., 2006). Based on the guidelines by Cooper et al. (2009), moderator analyses were conducted by disaggregating study effect size estimates and

grouping them into appropriate categories. The chi-square test of homogeneity (Q statistic) of effect sizes and the associated I^2 statistic were used to determine the significance and magnitude of between (Q_B) and within (Q_W) group differences in mean effect sizes for each potential moderator variable (e.g., participant characteristics, study design characteristics, outcome measure characteristics, and contextual characteristics of intervention).

In the moderator analysis, "mixed-effects" analysis refers to a random-effects model and combine studies within each subgroup. As a random-effects model better suits the studies collected in this study, results of mixed-effects analysis were interpreted in this section.

In order to resolve dependency issue within studies, the "shifting-unit-of-analysis" approach (Cooper, 1998) by separating measures by construct (e.g., reading fluency, reading comprehension, phonics) was thoroughly conducted and meta-analyzed separately in the moderator analysis section. This approach attempts to group particular outcome measure domains together into categories while preserving statistical independence, allowing the analysis process to maintain all of the information from each study (Scammacca, Roberts & Stuebing, 2014). However, this approach also tends to produce small k number in some constructs. Such results are explained under each moderator analysis.

1. Publication type

Tables 4 illustrates the results of publication type moderator analysis. For the type of studies, published peer-reviewed articles and unpublished dissertation studies had point estimate of 0.687 and 0.637 in Hedge's g with the Q-value of 0.191 (df=1, p=0.662), indicating that the effects are not differed by publication type; publication type did not moderate the effect. The variance across subgroups was 0.277, and the combined estimate of I^2 was 74.622, explaining the proportion of the variance in observed effects which

is due to variance in true effects (Borenstein, Higgins, Hedges, & Rothstein, 2009).

Table 4
Effects of Publication Type

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Dissertation	64	0.687	0.097	0.498	0.877	0.000
Journal	86	0.637	0.064	0.512	0.762	0.000
Overall	150	0.652	0.053	0.548	0.756	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	0.191	1	0.662	74.622	0.277	0.055

2. Participant Characteristics

2-1. By Grade Level

Tables 5 presents the results of moderator analysis by grade level range on lower and mid-upper grades. Because primary grades within the elementary school years are particularly important for the prevention of many reading difficulties and disabilities (Fuchs, et al., 2008; Partanen & Siegel, 2014) the analysis separated lower and mid-upper graders.

The lower grades include first and second grade, mid-upper grades include third, fourth, fifth and sixth grade. As explained in the earlier section, when the studies clearly indicated first and second grade students as participant samples, they were included in "lower." And, because many studies have mixed grades of third to sixth graders as their study participants without further separation, any participants ranging from third to sixth grade were grouped

together and referred as "mid-upper" grades for this part of analysis. Lower grade participants had the effect size of g=0.558 and mid-upper grade participants had g=0.665 with the Q-value of 0.578 (df=1, p=0.447), indicating that grade level difference did not moderate the effect.

Table 5
By Grade Level

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Lower	20	0.558	0.129	0.305	0.810	0.000
Mid-upper	130	0.665	0.059	0.551	0.780	0.000
Overall	150	0.647	0.053	0.542	0.751	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	0.578	1	0.447	74.622	0.277	0.055

2-2. By Classroom Type

Table 6 presents the results of moderator analysis by classroom type. Study participants were grouped according to the classroom type of general education or special education (inclusion) classroom setting as individual studies indicated. If a struggling reader is an EFL or ELL, it was reasonable to assume that the student was placed in the general education setting unless the study indicated that the student had an IEP or receiving special education services in addition. Some studies (e.g. Mason et al., 2006; Vadsay & Sanders, 2008) dealt with participants from both general education and special education setting without dividing the participants into classroom type for their analysis. In this case, those students were grouped under "both" as their classroom setting type.

The effect size of general and special education setting was g=0.691 and g=0.618 respectively. However, the p-value was 0.575, denoting statistical insignificance among the type of classroom. The summed Q-value was 1.988 (df=3, p=0.575), indicating whether students belonged to the general education classroom or to the special education classroom setting or pulled from both settings did not moderate the effect. The classroom type distinction among struggling readers seems hard to play a critical factor for their intervention results. Both participant groups in general and special education classroom had medium effects.

Table 6

By Classroom Type

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
General	93	0.691	0.069	0.556	0.826	0.000
Special	48	0.618	0.095	0.432	0.804	0.000
Both	9	0.531	0.304	-0.129	1.127	0.159
Overall	150	0.656	0.054	0.550	0.762	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	1.988	3	0.575	74.622	0.277	0.055

2-3. By All Learner Type

Table 7 presents the results of moderator analysis by learner type. As defined in the original studies by the authors, learners in the studies were additionally grouped into categories of [learning] English as a Foreign Language (EFL), English as a Second Language (ESL) or English Language Learners (ELL), low achievers (LA) with English as their first language, and underachievers (UA) with English as their first language as well.

The effect size was g=0.827 for EFLs, g=0.465 for ELLs, g=0.221 for both ELL and SPED combined group, g=0.610 for LAs and g=0.222 for UAs. This category of learner type had the summed Q−value of 38.709 (df=5, p=0.000), indicating that the learner type difference moderated the effect, and the results are displayed in Table 7.

Table 7
By All Learner Type

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
EFL	90	0.827	0.246	0.683	1.097	0.000
ELL	12	0.465	0.162	0.147	0.783	0.004
ELL with SPED	3	0.221	0.106	0.013	0.429	0.037
LA	35	0.610	0.075	0.463	0.757	0.000
UA	10	0.222	0.069	0.086	0.357	0.001
Overall	150	0.535	0.045	0.446	0.623	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	38.709	5	0.000	74.622	0.277	0.055

2−4. Learner Type of EFL, ELL, LA

Participants learning English in a non−speaking English environment are technically referred to as learning English as Foreign Language (EFLs), not English Language learners (ELLs) who learn English in an English−surrounding environment. Both groups of participants are also different from those struggling students who learn English as their first language. Therefore, within all struggling learners, a separate moderator analysis was

conducted on EFLs, ELLs and combination group of LA and UA. The effect size was g=0.855 for EFLs, g=0.379 for ELLs, g=0.456 for both LA and UA population combined. The summed Q-value was 17.681 (df=2, p=0.000), indicating that whether the struggling students are learning English in a foreign language environment, in English surrounding environment, or as their first language moderated the effect. The results are displayed in Table 8.

Table 8

By EFL, ELL, LA Type

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
EFL	90	0.855	0.090	0.680	1.031	0.000
ELL	15	0.379	0.084	0.214	0.543	0.000
LA+UA	45	0.456	0.066	0.325	0.586	0.000
Overall	150	0.535	0.045	0.446	0.623	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	17.681	2	0.000	74.622	0.277	0.055

3. Type of Outcome Measure

Tables 9 presents the results of moderator analysis by outcome measure type, whether the outcome measure assessment developed by researcher or standardized or combination of the two moderated the effect. The effect sizes for the combination assessment type was g=0.707, for researcher-developed type was g=0.745, and for the standardized assessment outcome type was g=0.610. The summed Q-value was 1.285 (df=2, p=0.526), indicating that the type of outcome measure did not moderate the effect.

Table 9

Outcome Measure Type

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Combination	6	0.707	0.183	0.349	1.066	0.000
Researcher-developed	61	0.745	0.109	0.532	0.958	0.000
Standardized	83	0.610	0.062	0.488	0.731	0.000
Overall	150	0.648	0.052	0.547	0.749	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	1.285	2	0.526	74.622	0.277	0.055

4. Outcome Measure (dependent variables in the studies)

4-1. All Outcome Measures

Table 10 presents the results of moderator analysis by individual outcome measures. All outcomes were grouped into affective or achievement type according to their outcome characteristics as previous meta-analyses implemented as the important measurement framework (Jung & Choi, 2019; Lee & Chang, 2017). Especially for the struggling elementary school students, affective measures are quite essential in evaluating the students' learning performance and serve as an important set of criteria (Park & Park, 2007).

Affective measures included attitude, belief, competitiveness, confidence, interest, learning time, mindset, motivation, participation and satisfaction for analysis. However, of these, belief, competitiveness, learning time, mindset and satisfaction measures

were not included in the table below for having too little number of effects sizes ("belief" k=2, "competitiveness" k=2, "learning time" k=2, "mindset" k=1, "satisfaction" k=2).

Variables of attitude (g=0.798, SE=0.204), confidence (g=0.971, SE=0.143), interest (g=0.884, SE=0.201) and participation (g=1.668, SE=0.582) had moderate to large effects with statistical significance within the affective measure type.

Achievement measures included cognition (CGN), fluency (FL), listening comprehension (LC), letter identification (LT), phonological awareness (PA), reading comprehension (RC), speaking (SPK), spelling (SPL), syntax (SYN), vocabulary (VOC), word identification (Word ID) and writing (WR). However, of these, cognition (k=2), listening comprehension (k=3), spelling (k=1), syntax (k=2) measures were not included in the table for having too little k as well. Reading comprehension (g=0.658, SE=0.136), speaking (g=0.841, SE=0.149), vocabulary (g=0.537, SE=0.183), word identification (g=0.495, SE=0.197) and writing (g=0.535, SE=0.194) had moderate effects with statistical significance within the achievement measure type.

With the summed Q-value of 76.548 (df=21, p=0.000), outcome measures moderated the effect. Results are displayed in Table 10.

Table 10

All Outcome Measures in Domains of Affective and Achievement

		Effect size and 95% confidence interval					Test of null (2-tail)
	Measure in	*k*	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Affective	Attitude*	8	0.798	0.204	0.399	1.198	0.000
	Confidence*	12	0.971	0.143	0.690	1.251	0.000
	Interest*	10	0.884	0.201	0.490	1.277	0.000
	Motivation	5	0.601	0.316	-0.501	1.704	0.285
	Participation*	4	1.668	0.582	0.528	2.808	0.004
Achievement	FL*	23	0.635	0.114	0.412	0.857	0.000
	LT	7	0.059	0.256	-0.442	0.561	0.816
	PA	12	0.527	0.312	-0.084	1.138	0.091
	RC*	18	0.658	0.136	0.391	0.924	0.000
	SPK*	7	0.841	0.149	0.550	1.133	0.000
	VOC*	15	0.537	0.183	0.178	0.897	0.003
	Word ID*	4	0.495	0.197	0.108	0.881	0.012
	WR*	8	0.535	0.194	0.155	0.915	0.006
		Heterogeneity results					
		Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
	Overall	76.548	21	0.000	74.622	0.277	0.055

4−2. Outcome Measures by Achievement vs. Affective

Table 11 presents outcome measure comparison by affective vs. achievement measure type. Large effect (g=0.845, SE=0.096) was found in the affective measure type and medium effect was found in achievement measure type (g=0.564, SE=0.063) With Q−value of 5.977 (df=1, p=0.014), outcome measure type significantly moderated effect.

Table 11

Outcome Measure by Achievement and Affective

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Achievement	102	0.564	0.063	0.440	0.688	0.000
Affective	48	0.845	0.096	0.657	1.033	0.000
Overall	150	0.649	0.053	0.546	0.753	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	5.977	1	0.014	74.622	0.277	0.055

5. Intervention (independent variables in the studies)

In a meta-analysis examining instructional effects on English underachieving students, Kim, Wi, and Kim (2015) grouped instructions into activity-focused, material-focused, strategy-focused and program-focused instructions in order to compare instruction type for their effectiveness. Similarly, Lee and Change (2017) also grouped the interventions into general reading instruction, technology instruction, program/model instruction and others in an attempt to find each intervention type effectiveness. Since English intervention names written in individual studies tend to vary with no limited vocabulary range, grouping interventions into similar types was also necessary for this meta-analysis as finding intervention effectiveness was one of the main purposes in the study.

The first author and a doctor with an Ed.D. degree in special education and current elementary school special education teacher with 21 years of experience in Connecticut, USA, together grouped all the existing interventions under certain categories and the groups were provided with appropriate intervention type names.

All the intervention names were grouped into three categories:

content-based type (i.e., if contents of English sub-construct names were used), learner approach-based type (i.e., if learner's manipulation and learning behavior embedded in the intervention) and teacher/instructional method-based type (i.e., if it was an actual instructional name carried by teachers). Teacher/instructional method-based instruction included intervention names such as direct instruction , flipped classroom instruction, repeated instruction.

Intervention names that did not belong to any of aforementioned categories or were considered too vague by definition (e.g. "workbook" "eclectic" "remedial" "lab") were classified under "others" and were not included in the analysis. 98% interrater agreement was found among three researchers for this grouping and the minor disagreement was resolved after discussions.

As such, in the *content-based* type, studies with interventions of (a) phonics, (b) vocabulary, (c) fluency program were listed. If any two contents were combined in the yielding of research results, they were categorized under "mixed" group. Secondly, in the *learning approach-based* type, studies with interventions of (a) multisensory/multi-intelligence, (b) feeling-based, (c) role- play, (d) song, (e) strategies, (f) mindset/concepts were listed. Thirdly, within teacher/*instructional method-based* type, studies with interventions of (a) repeated reading, (b) shared reading, (c) collaborative learning, (d) small group instruction, e) systematic-explicit instruction, (f) intensive instruction, and (g) immediate instruction were listed. Full classification table is displayed in Appendix 3. Those three intervention types were analyzed individually and then analyzed for the group heterogeneity significance.

5-1. Content-based intervention

Tables 12 presents the results of moderator analysis by content-based intervention for fluency, phonics and mixed contents. The effect size for fluency intervention was g=0.205 (SE=1.101),

for phonics intervention was g=0.289 (SE=0.234) and for the mixed contents was g=1.733 (SE=0.227). Studies using mixed contents as intervention names had very large effect while fluency and phonics as intervention names had small effects. The summed Q-value was 43.188 (df=3, p=0.000), indicating that the content-based intervention type moderated the effect.

Table 12

Content-based Intervention

Effect size and 95% confidence interval						Test of null (2-tail)
	k	g	SE	Lower Limit	Upper Limit	p-value
Fluency	5	0.205	1.101	0.006	0.403	0.043
Phonics	22	0.289	0.234	-0.171	0.748	0.218
Mixed	14	1.733	0.227	1.287	2.179	0.000
Heterogeneity results						
	Q-value	df(Q)	p-value	I^2	T^2	SE
Overall	43.188	3	0.000	74.622	0.277	0.055

5-2. Learner approach-based intervention

Tables 13 presents the results of moderator analysis by learning approach-based intervention for feeling-based, mindset/concept, multi-intelligence/multi-sensory, role play, song and strategies. All of the learning approach-based interventions had significant effect sizes. Multi-sensory/intelligence intervention had the largest effect (g=1.21, SE=0.314) followed by song (g=1.09, SE=0.171), mindset/concept (g=0.868, SE=0.308) and feeling-based intervention (g=0.853, SE=0.18). All of the learner approach-based interventions showed large effects and strategies had moderate effect (g=0.617, SE=0.133). The Q-value was 17.513 (df=7, p=0.014), indicating that the learning approach-based intervention type moderated the effect. Due to small k

number again, role-play (k=1) was excluded from the table 13.

Table 13

Learner approach-based Intervention

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Feeling-based	16	0.853	0.180	0.500	1.206	0.000
Mindset/concept	5	0.868	0.308	0.264	1.471	0.005
Multi-sensory/intelligence	6	1.217	0.314	0.594	1.825	0.000
Song	8	1.090	0.171	0.755	1.424	0.000
Strategies	16	0.617	0.133	0.356	0.878	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	17.513	7	0.014	74.622	0.277	0.055

5-3. Teacher/Instructional method-based intervention

Tables 14 presents the results of moderator analysis by teacher/instructional method interventions. Small group instruction as intervention was excluded from the table for too little *k* number. The summed *Q*-value was 31.561 (*df*=7, *p*=0.000), indicating that the teacher/instructional method-based intervention type moderated the effect. Of all interventions in this group, shared reading and immediate feedback had large effects (*g*=1.449, *SE*=0.299; *g*=1.10, *SE*=0.085) followed by medium effect of systematic-explicit (*g*=0.479, *SE*=0.237), mixed instructions (*g*=0.426, *SE*=0.0780) and intensive instruction (*g*=0.418, *SE*=0.187). Repeated instruction had a small effect (*g*=0.321, SE=0.067), all interventions with statistical significance.

Table 14

Teacher/Instructional method-based Intervention

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Immediate feedback	3	1.100	0.085	0.585	0.920	0.020
Intensive instr.	4	0.418	0.187	0.052	0.784	0.025
Repeated instr.	22	0.321	0.067	0.190	0.453	0.000
Shared reading	10	1.449	0.299	0.862	2.036	0.000
Syst-explicit	5	0.479	0.237	0.014	0.945	0.044
Mixed	13	0.426	0.078	0.273	0.579	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	31.561	7	0.000	74.622	0.277	0.055

5-4. By Intervention Type

Table 15 presents the results of moderator analysis by different types of interventions. All of the intervention types were statistically significant for the mean. The learner-approach intervention had large effect (g=0.787, SE=0.095) followed by content-based (g=0.69, SE=0.155) and teacher/instructional-based (g=0.557, SE=0.061). The Q-value for the heterogeneity was 4.284 (df=2, p=0.117) indicating that the intervention component type difference did not moderate the effect.

Table 15

Intervention Component Type

Effect size and 95% confidence interval	Test of null (2-tail)

	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Content-based	41	0.690	0.155	0.386	0.994	0.000
Lrner. Approach	51	0.787	0.095	0.600	0.973	0.000
Teacher-Instrc.	57	0.557	0.061	0.437	0.677	0.000

Heterogeneity results

	Q-value	df(Q)	*p*	I^2	T^2	*SE*
Overall	4.28	2	0.117	74.62	0.28	0.056

6. By Interventionist

Table 16 present the results of moderator analysis by interventionist type. The teacher and researcher combination group had the largest effect (g=1.586, SE=0.175) followed by teacher group (g=0.736, SE=0.114) and researcher group (g=0.695, SE=0.087). The others group in which intervention was provided by paraprofessionals or teaching assistants, a small effect was found (d=0.277, SE=0.055). All of the results were statistically significant, and the Q-value was 64.639 (df=4, p=0.000) indicated that this interventionist type difference moderated the effect.

Table 16

By Interventionist Type

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Researcher	78	0.695	0.087	0.524	0.865	0.000
Teacher	28	0.736	0.114	0.514	0.959	0.000
T&R	13	1.586	0.175	0.244	1.928	0.000
Bilingual T	5	0.479	0.237	0.014	0.945	0.044
Others	26	0.277	0.055	0.170	0.384	0.000
Overall	150	0.508	0.041	0.428	0.588	0.000

Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	64.639	4	0.000	74.622	0.277	0.055

7. Frequency of Intervention

Table 17 presents the results of moderator analysis by frequency of the intervention. Of all intervention session numbers per week, 3 sessions per week had the largest effect (g=1.035, SE=0.096) followed by two sessions per week (g=0.846, SE=0.092). Interventions conducted for more than three sessions per week had a small effect (g=0.335, SE=0.06). The Q-value of 49.616 (df=4, p=0.000) indicated that this difference was statistically significant and therefore intervention frequency on a weekly bases measure moderated the effect.

Table 17

Frequency of Intervention

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
Once per week	34	0.410	0.151	0.114	0.706	0.007
Twice per week	49	0.846	0.092	0.666	1.027	0.000
3x per week	33	1.035	0.096	0.847	1.224	0.000
More than 3x	32	0.335	0.060	0.218	0.451	0.000
NR	2	0.445	0.169	0.115	0.775	0.008
Overall	150	0.579	0.041	0.498	0.660	0.000

Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	49.616	4	0.000	74.622	0.277	0.055

8. Total intervention sessions

Table 18 presents the results of moderator analysis by the total number of intervention sessions each study reported. Of all session numbers in its entirety, 20 to 29 session range had the largest effect (g=1.025, SE=0.149). Less than 10 sessions, 10 to 19 sessions, 30 or more sessions had about the medium effect sizes. And, the Q-value of 9.413 (df=4, p=0.052) indicated that this session number difference was marginally insignificant as the moderator to the mean. However, each session range effect size was significant.

Table 18

Total Number of Intervention Sessions

Effect size and 95% confidence interval						Test of null (2-tail)
	k	*g*	*SE*	Lower Limit	Upper Limit	*p*-value
<10	19	0.603	0.105	0.397	0.809	0.000
10-19	46	0.529	0.127	0.279	0.778	0.000
20-29	26	1.025	0.149	0.279	1.316	0.000
30 or more	27	0.577	0.068	0.734	0.709	0.000
NR	2	0.445	0.169	0.444	0.775	0.000
Overall	150	0.610	0.047	0.115	0.703	0.000
Heterogeneity results						
	Q-value	df(Q)	*p*-value	I^2	T^2	*SE*
Overall	9.413	4	0.052	74.622	0.277	0.055

4.4. Meta-Regression

Meta-regression analysis was additionally done in order to explore the amount of variance accountable for the sets of moderators. Meta-regression is a statistical technique used in a meta–analysis to examine how characteristics of studies are related to variation in effect sizes across studies (Borenstein et al., 2009). Generally speaking, meta-regression is analogous to regression analysis but using effect sizes as outcomes, and information are being extracted from studies as moderators or predictors. When there is a heterogeneous set of effect sizes drawn, it can be used to examine the association among characteristics of the study and variation among effect sizes a priori for either categorical and continuous (Borenstein et al., 2009.) In this meta-regression, only covariates without missing values and any covariates not linked to collinearity were included for the analysis. For example, covariate sets linking to possible collinearity (e.g. “sessions per week” and “total number of sessions,” “motivation” and “affective outcome measure”) were screened and excluded to prevent the collinearity problem.

The meta-regression results are summarized in Table 19. While other predictors being held constant, learner type intervention (p=0.023), interventionist type (p=0.0014) type of outcome measure (p=0.0013), sessions per week (p=0.0002) were found to have significant variance like seen in the moderator analyses by meta-ANOVA. On the other hand, grade level, intervention types, publication type were not found to be significant in variance in effect sizes. The regression model explained 27% of the variance. The analysis results are the displayed in hedges g and scatter plots for variability distribution of those predictors are presented in Appendix E.

Table 19

Meta-Regression Results

Set	Covariate	Coefficient	*SE*	95% lower	95% upper	Z-value	2-sided *p*-value	Set
	Intercept	-0.673	0.3215	-1.3032	-0.0429	-2.09	0.0363	
Grade	lower vs. mid-upper	-0.0503	0.1998	-0.4419	0.3413	-0.25	0.8012	
Learner type*	EFL vs. ELL	-1.029	0.3823	-1.7783	-0.2797	-2.69	0.0071	*p*=0.023
	EFL vs. LA	-0.3278	0.2669	-0.8509	0.1953	-1.23	0.2193	
Interventionist type*	Researcher vs. bilingual T.	1.0277	0.5222	0.0042	2.0513	1.97	1.0277	*p* =0.0010
	Researcher vs. Teacher	0.609	0.2542	0.1107	1.1072	2.4	0.0166	
	Researcher vs. T & R	0.9555	0.2481	0.4692	1.4417	3.85	0.0001	
	Researcher vs. others	0.1468	0.3156	-0.4719	0.7654	0.47	0.6419	
Sessions per week*	Once vs. twice	0.2441	0.1671	-0.0835	0.5717	1.46	0.1441	*p* =0.0002
	Once vs. 3x	0.9606	0.2151	0.5391	1.3822	4.47	0.0000	
	Once vs. more than 3x	0.4962	0.2145	0.0758	0.9166	2.31	0.0207	
Publication type	Dissertation vs. journal	0.2556	0.1769	-0.091	0.6023	1.45	0.1483	

Test of Model	T^2	I^2	**Q**	**df**	***p*-value**	R^2
			74.88	17	0.000	
Goodness of fit	0.2765	66.24%	390.95	132	0.000	
Total btw. study variance	0.2765	74.62%	587.12	149	0.000	
Proportion of total btw. study variance explained by the model						0.27

4.5. Publication bias

One of the limitations about meta-analysis may be the existence of publication bias. The term publication bias refers to the likelihood of studies being published or unpublished due to the studies' tendency or directions of the results (Higgins & Green, 2011). Studies with significant results are more likely to be published while studies with insignificant results are not likely to be published, staying in the "file cabinet," a possible reflection of "file drawer" problem. (Borenstein, et al., 2009, p.285)

Publication bias was tested by creating a funnel plot (Cooper et al., 2009) and Egger's test of asymmetry was utilized for any significance (Egger, Smith, Schneider, and Minder, 1997). Figure 3 below shows that larger studies are usually distributed near the top and smaller studies near the bottom of the funnel plot. Smaller studies tend to be more widely spread for wider range of standard error. Although these studies may be less influential on the meta-analytic results because they provide a small weight in the weighted average effect size computation (Pai, Sears, & Maeda, 2015, p. 86) all cases must be considered.

Because of there was some degree of asymmetry, Duval and Tweedie's Trim-and-Fill method was applied to address slight publication bias. Trim-and-fill is part of sensitivity analysis for possible influential effects on overall studies, not having the major focus on modifying the overall effect (Borenstein, et al., 2009). By doing so, this method removed most extreme studies from the positive direction, "re-compute the effects size at each iteration until the funnel plot is symmetric about the new effect size" (Borenstein et al., 2009, p.286). And the "fill" part in actuality has no impact on the point estimate but provides to correct the variance (Duval & Tweedie, 2000). Prior to the readjustment, the point estimate was 0.497, and after the adjustment the point estimate came out to be 0.336, This difference can be interpreted as not a significant difference for a major change in the final results, indicating that the publication bias was not a major concern in this study.

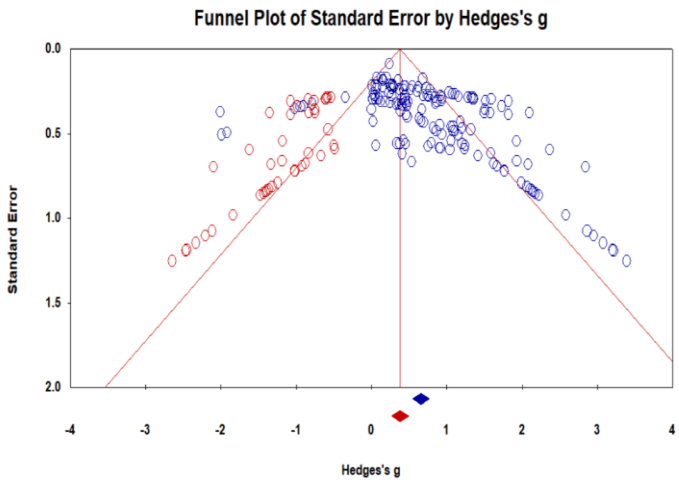


Figure 3. Funnel Plot for Publication Bias, Trim-and-Fill Model

Chapter 5. Discussion

This present meta-analysis examined the collected studies in order to assess effectiveness of English interventions on struggling elementary students in general and special education setting inclusive of various types of struggling students from 2000 to 2019. Findings were drawn from a total of 35 studies with 150 effect sizes of experimental and quasi-experimental study designs. Applying the random-effects model, the overall mean effect was drawn followed by moderator analyses on sub-groups. The variables used for sub-group analyses included (1) participant-related variables (e.g., classroom type, grade level, English as a Foreign Language/English Language Learners/low achievers of English as first language), (2) outcome-related variables (e.g., intervention outcome measures, type of outcome), (3) intervention components (e.g., intervention types), (4) contextual characteristics (e.g., interventionist, frequency, duration). In this chapter, the summary of findings, implication for discussion and

study limitations are discussed.

5.1. What is the overall effectiveness of English language interventions on struggling readers in elementary school?

The results indicated the overall mean effect size g=0.655 [CI_{95}: 0.55, 0.761] with heterogeneity Q−value of 587.12 (df=149, p=0.000) indicating a moderate overall effect of English intervention on struggling readers. According to the Q−value, the variances reflect true effect size differences from study to study rather than likelihood of sampling error.

5.2. To what extent do intervention effects vary as per participant−related moderator? (i.e., classroom type, grade level, LA/ELL /EFL status)

When moderator analyses for the categorized sub−groups were conducted on the characteristics of participants, the effect sizes were sought for (1) classroom type, (2) grade level range, (3) all learner types and (3) grouped learner types. First, classroom type analysis revealed that it was a non−moderator at the significance level. The general education classroom had g=0.691 and special education classroom had g=0.618 while studies that had incorporated both setting had g=0.531, all indicating moderate effects to the mean with the p−value of 0.575. This illustrated that whether the participant belonged to the special education classroom setting or in general education classroom setting in school did not influence the intervention effect. Interestingly, a study done by Hollo and Wehby (2017) attempted to examine teacher talk differences of general and special education classroom setting at the elementary schools, and its results revealed no statistically significant between−groups effects for any of the outcome variables. That is, the teachers observed in general education and special education classrooms were not significantly

different in their classroom talk quantity, complexity, content, or clarity. In fact, no statistically significant differences on any of these sub-structural elements of teacher talk across early or late elementary grade classrooms were found. As this study implies, when the teacher differences are hard to discern, the overall classroom atmosphere caused by the teachers talk cannot be too different between general and special education classroom setting. The classroom type distinction is hard to be a critical factor for academic progress as Zigmond's study (2003) also confirmed in her study. A serious debate on which classroom type benefits particular student type occurred in the mid-1980s and after reflecting on the 35 years of efficacy research on the education settings, "I can say with some certainty that place is not what makes special education 'special' or effective. Effective teaching strategies and an individualized approach are the more critical ingredients in special education, and neither of these is associated solely with one particular environment" (Zigmond, 2003, p. 198). Although educators are very aware that research has shown that typical general education environments are not supportive places to embrace all types of students, especially with disabilities (Zigmond, 1996), it appears that drawing effectiveness of instruction simply based on the classroom setting type is not reliable when other imperative factors are also present.

Second, whether the struggling readers were in the lower grades (first and second graders) or were in mid-upper grades (third, fourth, fifth and sixth) was not a significant factor contributing to the mean effect. The effect size of the lower grades was g=0.59 while the effect size of the mid-upper grades was g=0.67, both having moderate effects though not significant as a moderator. This result aligns with the meta-analysis by Kim et al. (2020) which examined reading intervention studies for struggling readers using improvement rate difference (IRD) index on single-subject research design studies. This study carefully scrutinized grade level differences and found large effects in following orders:

grade 10-12 (ES=0.86, mixed grades (ES=0.81), preschooler (ES=0.79), grade 1-3 (ES=0.76), grade 4-6 (ES=0.73), and grade 7-9 (ES=0.71). Weighted mean effects of all group were higher than 0.70, yet the large effects among all grade levels had no significant differences. This finding is also in line with Shin and McMaster (2019) in that there was no grade range difference on oral reading achievement as part of CBM in predicting reading comprehension. A meta-analysis by Jung and Lee (2019) similarly did not find any significant grade level differences when elementary English underachievers were examined on achievement and affective measures. In sum, it appears that grade level differences for distinguishable effectiveness do not seem to be a strong moderator.

Third, all struggling readers were grouped and analyzed by all learner types, defined by study authors as EFL (LA, UA), ELL, ELL with SPED, low-achievers, underachievers either in general education or special education class setting, having English as their first or second language. A large effect (g=0.827) was found for EFLs, moderate (g=0.465) for ELLs and LA (g=0.610), small (g=0.221) for ELLs receiving special education population combined and small (g=0.222) for UAs. For the subsequent analysis, the goal was to separate groups into smaller chunks minimizing the differences, underachievers and low-achievers were combined for having English as their first language, and ELLs (living in English speaking countries) and EFLs (living in non-English speaking countries) altogether were compared as three separate learner sub-groups. When moderator analyses were conducted comparing only EFLs, ELLs and low-achievers, the effect size was again large (g=0.855) for EFLs, small to moderate (g=0.379) for ELLs, moderate (g=0.456) for both LA and UA population combined. And, both analyses indicated not only individual types that moderated the mean effect, but also whether a struggling student is learning English in a foreign language environment, in English surrounding environment, or as their first

language moderated the effect. Struggling EFLs' having a large effect on the interventions is aligned with a meta-analysis done by Lee and Chang (2017) with 55 experimental studies on struggling English readers which found the overall effect size of d=1.32 (SE=0.071 p=0.000), with sub-categories of English achievement measure (d=1.381, SE=0.095, p=0.000) and affective measure (d=1.241, SE=1.241, p=0.000), concluding that elementary school level interventions, in contrast to the secondary school intervention, for English struggling readers has been working quite well (Lee & Chang, 2017). The difference in effect sizes for low-achievers and ELLs is not quite distinctive (g=0.456, g=0.379) and it is noteworthy, at least according to this study, to interpret that ELL struggling readers whose first language is not English were not much different in performance than the low-achieving struggling readers whose first language was English. In the meta-analysis study by Kim et al. (2020), low achievement population had ES=0.74 and learning disability population had ES=0.72, both having moderate-large intervention effects with a difficult line to distinguish the group differences.

5.3. To what extent do intervention effects vary by outcome type and outcome measures? (i.e., researcher-developed vs. standardized vs. combined, intervention outcome measures, achievement/affective outcomes)

Gersten et al. (2015) stated that certain characteristics and the quality of the measures that are selected and measured for intervention effect evaluation influences internal validity of a study. Therefore, variables related the measures should be examined with caution.

First, when the moderator analysis was conducted to find any significance on the type of outcome measure, whether the outcome measure was developed by the researcher or utilized the standardized or use of both kinds did not moderate the mean effect.

Both researcher-developed and use of both kinds were slightly larger than the standardized type alone (g=0.610); however, these medium effects in all did not significantly moderate the mean effect. Some have argued that standardized measures are more reliable while arguing that treatment effects tended to be stronger when research experimenters developed their own measures than standardized version of the same construct (Gersten, Baker, & Lloyd, 2000). Swanson, Hoskyn, and Lee (1999) also stated that it is very typical for researcher-developed measure to yield higher effect sizes. In the meta-analysis, Wanzek el al. (2010) also found researcher-developed measure on reading comprehension to have shown high effects in their study. Nonetheless, in this study, the differences of the type of outcome measures were not significantly different, being identical to the meta-analysis outcome findings by Lein (2016) on children with mathematical difficulties. In this case, the results can be interpreted as both measures had identical effects to the overall effect size, not having one type of measure more likely to yield higher effects than the other to the mean.

Second, outcome measures were carefully sorted and listed under domains of affective versus achievement. All outcomes were grouped into affective or achievement type according to their outcome characteristics as previous meta-analyses implemented as the important measurement framework (Jung & Choi, 2019; Lee & Chang, 2017). Especially for the underachieving elementary school students, affective measures are considered quite essential in evaluating the students' learning performance and serve as an important set of criteria (Park & Park, 2007).

When all outcomes were analyzed, within the achievement outcome measure, speaking measure had a large effect (g=0.841) followed by reading comprehension (g=0.658) and fluency (g=0.635). When outcome measure k was only between one and three, those variables were eliminated from the tables for too little effect size cases to interpret. Therefore, listening comprehension (k=3), spelling (k=1), syntax/grammar (k=2), spelling (k=1) were

not included in results section although meta-analyzed as moderators. Within affective outcome measure, participation (g=1.688) and confidence (g=0.971) had very large effects followed by self-interest (g=0.884) and attitude (g=0.798). These variables all can be interpreted as being highly effective factors to intervention success. And, all of the outcome measures significantly moderated the effect (p=0.000). Similarly, Chung and Choi (2019) found large effects in self-confidence (g=1.3, SE=0.22), participation rate (g=1.27, SE=0.21), interest (g=1.08, SE=0.18). Within affective measure outcome tpye, mindset (k=1), satisfaction (k=2), learning time (k=2), belief (k=2) and competitiveness (k=2) variables were not included in the table for limited effect size numbers for interpretation.

Lastly, when all the outcome measures were grouped into affective (e.g. confidence, belief) and achievement (e.g., fluency, reading comprehension) type by their outcome definition, whether the outcome measure was an affective or achievement type moderated the effect (p=0.014) with affective outcome measure type having a large effect (k=48, g=0.845, SE=0.096) while the achievement outcome measure type with the moderate effect (k=102, g=0.564, SE=0.063). The large effect in the affective outcome measure type can be explained by similar findings in other meta-analysis studies (Chung & Choi, 2019; Lee & Chang, 2017) and intrinsic motivation drive. Perhaps it is not surprising that the variables such as self-confidence, self-interest and attitude towards [English] learning are clearly considered as intrinsically motivating factors. Graus and Coppen (2017) conducted an intervention study on grammar in the low achieving foreign language classroom and found that the learning was very much mediated by student-teacher perceptions of learner autonomy and motivation. The results of this study implied the importance of uncovering how those affective factors (e.g., motivation) related to the learner's intrinsic motivation were affecting the student's performance in addition to the instructional method alone by the

teacher. Lee and Chang (2017) also found a very large effect on struggling students' English learning in the affective measure domain with d=1.24 regardless of the grade level and region of those elementary schools. Bondie and Zusho (2017) studied ELLs with learning disabilities and the participants' motivation to learn. The following themes emerged through the findings– student feelings were strongly related to perception of understanding, classroom conversation and self–confidence along with curriculum clarity (i.e., goals, activities, quality), and realistic expectations of progress were keys to academic success.

5.4. To what extent do intervention effects vary as per intervention components and their type (i.e., content–based, learning approach–based, teacher/instructional method–based)?

A number of meta–analyses examining instructional effects on English underachieving students tended to group intervention types as Kim, Wi, and Kim (2015) grouped instructions into activity–focused, material–focused, strategy–focused and program–focused instructions in order to compare instruction type for their effectiveness. Similarly, Lee and Change (2017) also grouped the interventions into general reading instruction, technology instruction, program/model instruction and others in an attempt to find each intervention type effectiveness. Likewise, intervention grouping of similar types was also necessary for this meta–analysis as finding intervention effectiveness was one of the main purposes in the study.

Within content–based type, both fluency and phonics alone has small effect (g=0.205, g=0.289) but any studies having mixed contents as intervention had a large effect (g=1.733, SE=0.227). This moderator was also significant (p=0.000). Similarly, O'Connor et al. (2002) found that one–on–one instruction with struggling readers on multi–components of phonemic awareness, fluency and spelling together had a very large effect (ES=1.56)

compared to the classroom matched mean (ES=1.26). Similarly, Therrien et al. (2006) also reported that the incorporation of fluency and comprehension together to have shown higher effects in the treatment group (ES=0.44) compared to the general reading control group (ES=0.37) through their achievement results by the Broad Reading scale of the Woodcock-Johnson Achievement Test III (Wanzek et al., 2010). Perhaps, struggling readers often benefit more from being exposed to multi-component structure of English reading intervention than one component at a time.

Within the learner approach-based, multi-sensory/intelligence interventions had the largest effect (g=1.127); song-based and feeling-based interventions also had a large effect (g=1.09, g=0.853); strategies had the medium effect (g=0.617). Multisensory and use of a variety of intelligence has gained much success throughout many intervention studies. For example, Multisensory Reading Program for First-Grade Students) evaluate the efficacy of the Institute for Multi-Sensory Education's supplementary Orton-Gillingham based reading program across three schools in a single school district. When students were assessed by Dynamic Indicators of Basic Early Literacy Skills (DIBELS) to measure the reading skills of 224 treatment and 476 comparison group first-grade students, the treatment group performed far better than the control group. Jubran (2012) also found multi-sensory instructional approach effective for teaching English. It is not exaggerated when students get involved with their kinesthetic way of learning, the learning effects have proved to be quite high.

Within teacher/instructional method approach, shared reading had the largest effect (g=1.449). It is not surprising shared reading had such a large effect when numerous reading intervention literature accentuated the importance of shared reading. According to Dougherty Stahl (2012) using "Shared Reading to Bridge the Difference," the teacher must assume the responsibility for reading a text for the majority of the students

and what happened after shared reading experience was that the children have emergent literacy skills that were easily recognized. Furthermore, shared reading is defined as assessing the intrinsic value of a literature-based health intervention, and as the community-based spaces in which participants can relate with both literature *and* one another (Longden et al., 2015). In his study, quantitative study data showed that that this shared reading intervention was associated with enhancement of a sense of 'purpose in life' and implied this phenomenon as "intrinsic value of a literature-based health intervention." By definition, shared reading occurs in a paired group of colleagues or teacher, this has proven to have a meaningful and high effect in the learner's performance.

Systematic-explicit instruction and immediate feedback also have been found with medium-large effects (g=0.479, g=1.10) as part of the teacher/instruction-based intervention. According to Allington (2002), the exemplary teachers in the study routinely gave immediate, direct, systematic and explicit demonstrations of the cognitive strategies that good readers use when they read for successful reading episodes. Graus and Coppen (2017) also studied the effective intervention components for EFL students, and the study findings showed that the participants considered direct, explicit, systematic in grammar instruction a necessary condition not only for linguistic correctness but also for advanced communicative competence (Scheffel, Shaw & Shaw, 2008). All teacher/instruction-based interventions were statistically significant except small group instruction with a small effect (g=0.25, p=0.056), and repeated reading showed small effect (g=0.321, p=0.000) with statistical significance.

When interventions were compared via component types of content, learner approach and teacher/instructional-based, the grouping differences of these three component types did not moderate the mean effect (p=0.117), indicating that a particular intervention component type did not necessarily matter more than

other two to the mean. Rather, it appears that individual interventions proven to have large effects are what is needed to be implemented on to the struggling readers. Again, effective teaching strategies and an individualized approach meeting the student's needs and academic goals are the more critical ingredients in education (Zigmond, 2003). Lee and Chang (2017) also asserted that the educators should not be fixated in particular type of intervention they design; rather, considering different cognitive and affective characteristics of the learners should be serious taken for the development of intervention and apply.

5.5. To what extent do intervention effects vary by other contextual characteristics of interventions (i.e., interventionist, intervention frequency)?

In this meta-analysis, who administered the actual intervention played as a significant moderator. There were true differences among all interventionists with Q-value=64.64 (df=4, p=0.00), and the teacher and researcher combination group as the intervention provider had the large effect of g=1.586, followed by teacher group only with g=0.736 and researcher group only with g=0.695. When interventions were provided by others (e.g. paraprofessionals, teacher assistants, research assistants), the effects were small (g=0.28) and the interventions given by bilingual teachers was also moderate (g=0.479). This finding was unique in that conducting intervention by both researcher and teacher was a rare combination in other meta-analyses and effectiveness studies. The teachers' group with the effect size of g=0.736 is also considered nearly large. It contrasts with some other studies in which all interventionist type showed no statistical difference as in the study done by Lein (2016) on intervention on children with mathematical difficulties. Also, Kim et al. (2020) on the interventionist effects found reading interventions implemented by the teacher having the large effect (ES=0.83) and the interventions

delivered by the researchers also resulted in large effects (ES= 0.79); however, it showed no significant difference between these two groups.

Nonetheless, a study by Hirschstein, Edstrom, Frey, Snell, & MacKenzie (2007) found the important role of a teacher and underscored that the teachers have been identified as key agents in reducing bullying in schools. The study further argued that the teachers tended to intervene at both the universal classroom level and at the individual level; they are the leaders for instruction and activities and also often provide extra guidance or coaching, accentuating an essential role both at the instruction and social level. As noted earlier, according to Allington (2002), the exemplary teachers in the study tended to routinely gave direct, systematic and explicit demonstrations of the cognitive strategies that good readers use when they read for successful reading episodes. This can equate to the notion that the teacher's role as a great instruction implementer cannot easily be overlooked for the quality provided in class. In fact, as this meta-analysis indicates, teacher and researcher as instructional professionals that carry on the actual instruction mattered very much contributing to the mean effect.

Insofar as other characteristics of the intervention were concerned, intervention session frequency per week also played as a significant moderator to the effect. Through the analysis, intervention frequency of three times per week had the large effect (g=1.035), followed by two times per week (g=0.846). However, more than three times per week had a small effect size (g=0.335) and having once per week also had small-medium effect size (g=0.41). It may appear ironic how providing intervention more than three times a week may not have a large effect, if not bigger than three times, but providing instruction sessions too frequently may not always be the ideal for the struggling readers. Jung and Choi (2019) also found identical results- providing English intervention to underachievers three times a week had the largest

effect in affective measures (g=3.12, SE=0.79) and achievement measure (g=1.52, SE=0.23) in the study both with statistical significance.

With the significant heterogeneity but a marginal p-value of p=0.052, intervention session numbers between 20-29 had the largest effect (g=1.025, SE=0.149) while other number of sessions as in less than 10 sessions (g=0.603), between 10-19 sessions (g=0.529), sessions of 30 or more (g=0.577) all showed moderate effect sizes. Interestingly, providing intervention sessions between 25 to 63 hours in the meta-analysis study by Wanzek et al. (2018) was found with a moderate effect as well. Identically, in the study by Kim et al. (2020), the mean weighted effect size for 21-30 sessions showed the largest effect (ES=0.82) in their study. In contrast, Chung and Choi (2019) found more than 33 sessions with the largest effect size in both affective and achievement measures (g=1.93, SE=0.36; g=1.56, SE=0.22). It appears that in regard to the number of intervention sessions provided to the students, there may not be one most effective session range; however, providing more than 20 sessions is what seems effective to many struggling readers.

Chapter 6. Limitations and Future Directions

Some of the limitations of this study include the following. First, as aligned with the intention of the author, all of the struggling learner types and classroom types were included along with other moderators for the analyses. However, this may be considered too broad for drawing overall generalization. Especially, when there exists an extensive level of heterogeneity within each struggling reader population group for English (e.g., EFLs, ELLs, LAs), such an attempt to examine the struggling readers by their grouping may yield dissimilarities beyond the findings of this study. For instance, in order to better assess the EFLs' intervention effects, speakers of other languages than Korean as the students' first language must,

too, be examined; the ELL students in English speaking countries other than the U.S. and Canada (e.g., England) should also be looked into for better generalization of the entire ELL population. Low achievers are also largely heterogeneous by nature, and various definitions are already used by previous researchers. Therefore, more careful generalization is necessary for each learner population.

Second, the sub-groups were carefully scrutinized as part of shifting-unit-of-analysis approach in order to reduce existing dependency for the effect sizes collected on all the outcomes and interventions. According to Scammacca, Roberts, Stuebing (2014), one of the ways in handling dependence in meta-analysis is to use shifting-unit-of-analysis approach suggested by Cooper (1998), separating measures for each construct (e.g., reading comprehension, fluency) and meta-analyze effect sizes for each construct individually. In turn, "this approach tends to minimize violations of the assumption of independence of the effect sizes while preserving as much as the data as possible" (Scammacca, Roberts, & Stuebing, 2014 p.7). Although used for this meta-analysis, this approach can result in having multiple meta-analyses for outcome types with a very small number for individual effect sizes (k) and therefore having little power (Scammacca, Roberts, & Stuebing, 2014). In this meta-analysis, Hedge's g, too, was used instead of Cohen's d in order to supplement small numbers in some sub-groups as d tends to expand the effect size results in small sample size (Cooper, Hedges, & Valentine, 2009). Even with an effort to supplement, the small independent effect size numbers (k) in outcome measures indeed occurred in this study. For instance, in outcome measures, variables of learning time, mindset, satisfaction, syntax, spelling, listening comprehension and other variables all had small effect size numbers (k=1-3). Also, it is important to note that the mixed results of the moderator analyses for contextual characteristics on intervention tend to align with the notion that these variables are intricately intertwined, and variation in one may

lead to variation in others, which should be given attention (Richards-Tutor, et al., 2016).

Third, other influential factors that previous research have proved on having significant effects on a student's performance, such as the student's socioeconomic status, academic and family background, cultural background and their learning style were not examined in this study. Those factors should give sensitive interpretations in addition to what this meta-analysis could provide, especially with the types of learners. In addition, instructional group size was not included as a moderator in this meta-analysis. In numerous occasions, many meta-analyses and effectiveness studies include instructional group size for examining if pertinent group sizes contribute more to having large effects than other group sizes. As Wanzek et al. (2016) concluded, effect sizes were larger when the intervention was administered at the smallest group size level possible among kindergarten and first grade children. And, this type of intervention or instructional group size might have been additionally helpful.

In sum, future research is warranted for (a) population-specific criteria and detailed demographic information with EFL, ELL, LA as struggling readers for better generalizability, (b) stronger methods for overcoming dependency and small k numbers other than shift-of-analysis approach and Hedge's g application, (c) detailed contextual characteristics including SES, cultural background, and (d) more comprehensive moderators to be included such as instructional group size.

Chapter 7. Conclusion and Implications

Although this meta-analysis has certain limitations, the following conclusion and implications for education can be drawn from it. First, classroom type and grade level performance differences within the elementary school level are hard to discern. As supporting evidence previous research articles also confirmed in

the Discussions chapter earlier, educators should be very careful about placing weight in classroom type and grade level differences. Second, differentiating learner types and prior assessment and evaluation of an individual learner is quite essential. Assessing the learner's background information, social and academic development in other subjects, responses to intervention both at the first language and second language level of English, observational assessment, along with other demographic information to best address the learner's attributes and needs must be seriously taken. Jung and Choi (2019), in fact, asserted that for the struggling readers, individual attention and modification are especially important regardless of the intervention group size. Kim et al. (2020) also stated that the effect size was significantly smaller when students with learning difficulties and students with other disabilities received interventions together in a mixed group, suggesting that perhaps reading intervention should carefully be tailored and implemented to students with similar academic and social needs. Individually customized student support has recently gained emphasis and collecting direct data through various paths is what is truly important for providing adequate education (Lee et al., 2016). Such quality assessment and intervention by educators must be critical for different learners when the learner type has been proven to be a significant moderator through the sub-group analysis.

Third, affective outcome measures proven to be higher in effectiveness than achievement outcome measures may imply that educators should consider successful intervention factors that tend to motivate learners as to encourage their intrinsic motivation. Those variables were found as motivation, attitude, confidence and self-interest. Teachers' tailored instructions were directly linked to increasing self-esteem that leads to intrinsic motivation in many students (Lee & Chang, 2017); therefore, the instructional components and constructs of the right measure must also be carefully planned and carried out to learners. When affective

outcome measures serve as an important set of criteria especially for the elementary struggling learners (Park & Park, 2007), those affective components should be included frequently and qualitatively along with other English achievement measures (e.g., English contents). Fourth, mixed-content teaching, shared reading, multi-sensory/multi-intelligence, feeling-based instruction and immediate feedback by teachers were found as highly effective instructions for elementary school struggling readers. Those interventions involve student-student, student-teacher interactions, which seems to be of importance for successful learning experience. Effectiveness further depends not only on the characteristics and needs of a particular student but also on the quality of the program's implementation (Zigmond, 2003). Intervention sessions being most effective with three sessions per week, quite different from other intervals within a week, may suggest the most effective learning interval for struggling readers. In line with the findings by Lee and Chang (2017), educators should remember that 'the longer the better' does not necessarily work. Finally, when the results also indicate that interventions done by teachers and researchers produced higher effects than by teaching assistants (TA) and paraprofessionals in class, more attention should also be given to training qualified professionals as educators to the struggling readers.

References

Al Otaiba, S., Connor, C. M., Foorman, B., Schatschneider, C., Greulich, L., & Sidler, J. F. (2009). Identifying and intervening with beginning readers who are at-risk for dyslexia: Advances in individualized classroom instruction. *Perspectives on language and literacy, 35*(4), 13-19.

Allington, R. L. (2002). What I've learned about effective reading instruction: From a decade of studying exemplary elementary classroom teachers. *Phi Delta Kappan, 83*(10), 740-747. https://doi.org/10.1177/003172170208301007

American Psychological Association (2008). *Reporting Standards for Research in Psychology: Why Do We Need Them? What Might They Be?* American Psychological Association. https://doi.org/10.1037/0003-066X.63.9.839

An, S., & Lim, H. W. (2013). Exploring the characteristics and causes of English-underachieving learners. *English Language Teaching, 25*(4), 167-186

*An, Y. J. (2016). Development and Application of Supplementary Materials Model for Underachievers in Elementary English Education. Master's thesis, Incheon National University.

Andersson, U., & Lyxell, B. (2007). Working memory deficit in children with mathematical difficulties: A general or specific deficit? *Journal of experimental child psychology, 96*(3), 197-228. https://doi.org/10.1016/j.jecp.2006.10.001

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1),

50-57. https://doi.org/10.1111/j.1540-5826.2005.00120.x

*Bae, M. J. (2015). Effects of Different Grouping Methods on Low-achieving Students' Speaking Ability, Attitude and Motivation. Master's thesis, Yonsei University. UCI: G901:A-0005929521

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of educational psychology, 93*(3), 498-520. https://doi.org/10.1037/0022-0663.93.3.498

Bondie, R., & Zusho, A. (2017). Racing against yourself: high stakes for adolescent English Language Learners with disabilities. *Teachers College Record, 119*(9), 1-42. Retrieved from http://www.tcrecord.org

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods, 1*(2), 97-111. https://doi.org/10.1002/jrsm.12

Borenstein, Michael & Hedges, Larry & Higgins, Julian & Rothstein, Hannah. (2009). *An Introduction to Meta-Analysis.* John Wiley & Sons. https://doi.org/10.1002/9780470743386

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of educational research, 73*(2), 125-230. https://doi.org/10.3102/00346543073002125

Card, N. A. (2012). *Applied Meta-Analysis for Social Science Research*. New York, NY: Guilford Press. https://doi.org/10.1111/insr.12011_17

Chall, J. S. (1983). Literacy: Trends and explanations. Educational Researcher, 12(9), 3-8.

https://doi.org/10.3102/0013189x012009003

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The Reading Crisis: Why Poor Children Fall Behind*. Harvard University Press.

*Cheon, S. Y. (2003). Cooperative Reading Activities for Children with Mild Reading Problems in the Elementary English Classroom. Master's thesis, Seoul National University of Education. UCI: G901:A-0005064515

*Cho, H. J. (2015). A Study on Improvement of Elementary School Relative Underachievers in Elementary English through a Shared Reading Activity. Master ' s thesis, Gyeongin National University of Education. UCI: G901:A-0005869803

Choi, S. H., & Ahn, K. J. (2015). A case study of using smartphone applications for elementary English underachievers vocabulary learning. *Modern Studies in English Language & Literature, 59*(1), 409-439. http://dx.doi.org/10.17754/MESK.59.1.409

* Jeong and Kim (2017). Early reading instruction effects on elementary underachievers using multiple intelligence theory. *Elementary English Education 23*(2), 69-90.

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin. https://doi.org/10.1198/jasa.2005.s22

Cooper, H. M. (1998). *Synthesizing research: A guide for literature reviews* (Vol. 2). Sage.

Cooper, H., Hedges, L., & Valentine, J. (2009). *The Handbook of Research Synthesis and Meta-Analysis* (2nd Edition). Russell Sage Foundation.

Daane, M. C., Campbell, J. R., Grigg, W. S., & Goodman, M. J. (2005). *Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading*. *The Nation's Report Card*. National Center for Education Statistics.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: a meta-analysis of 137 personality traits and subjective well-being. *Psychological bulletin, 124*(2), 197-229.

Denton, C., Fletcher, J., Anthony, J., & Francis, D. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities, 39*, 447-466. http://dx.doi.org/10.1037/0033-2909.124.2.197

DiCerbo, P. A., Anstrom, K. A., Baker, L. L., & Rivera, C. (2014). A review of the literature on teaching academic English to English language learners. *Review of Educational Research, 84*(3), 446-482. https://doi.org/10.3102/0034654314532695

Donovan, M. S., & Cross, C. T. (2002). *Minority Students in Special and Gifted Education.* Committee on Minority Representation in Special Education. https://doi.org/10.17226/10128

Dougherty Stahl, K. A. (2012). Complex text or frustration-level text: Using shared reading to bridge the difference. *The Reading Teacher*, *66*(1), 47-51.

Duffy, A. M. (2000). Balance, literacy acceleration, and responsive teaching in a Summer school literacy program for elementary school struggling readers. *Reading Research Instruction, 40*(2), 67-100. https://doi.org/10.1080/19388070109558336

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629-634.

Fien, H., Santoro, L., Baker, S. K., Park, Y., Chard, D. J., Williams, S., & Haria, P. (2011). Enhancing teacher read alouds with small-group vocabulary instruction for students with low vocabulary in first-grade classrooms. *School Psychology Review, 40*(2), 307-318. https://doi.org/10.1080/02796015.2011.12087720

Foorman, B. R., & Moats, L. C. (2004). Conditions for sustaining research-based practices in early reading instruction. *Remedial and Special Education, 25*(1), 51-60. https://doi.org/10.1177/07419325040250010601

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. https://doi.org/10.1111/0938-8982.00020

Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Based Recommendations for Instruction and Academic Interventions. Practical Guidelines for the Education of English Language Learners*. Center on Instruction. Retrieved April 11,2008, from http://www.centeroninstruction.org/files/ELL1-Interventions.pdf

Frost, R. (1992). *Orthography and Phonology: The Psychological Reality of Orthographic Depth*. John Benjamins Publishing. https://doi.org/10.1075/tsl.21.19fro

Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, J., & Davis, G. N. (2008). Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing*, *21*(4), 413-436.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93–99. https://doi.org/10.1598/RRQ.41.1.4

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., Bryant, J. D., & Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities?. *Developmental psychology, 46*(6), 1731-1746. https://doi.org/10.1037/a0020662

Gersten, R., Baker, S., & Lloyd, J. W. (2000). Designing high-quality research in special education: Group experimental design. *The Journal of Special Education, 34*(1), 2-18. https://doi.org/10.1177/002246690003400101

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*(3), 1202-1242. https://doi.org/10.3102/0034654309334431

Gersten, R., Rolfhus, E., Clarke, B., Decker, L. E., Wilkins, C., & Dimino, J. (2015). Intervention for first graders with limited number knowledge: Large-scale replication of a randomized controlled trial. *American Educational Research Journal, 52*(3), 516-546. https://doi.org/10.3102/0002831214565787

Gliner, J. A., Morgan, G. A., & Leech, N. L. (2009). *Research Methods in Applied Settings: An Integrated Approach to Design and Analysis* (2nd edition). NY: Routledge. https://doi.org/10.4324/9780203843109

Goldenberg, C. (2008). Teaching English language learners: What the research does-and does not-say. *American Educator*, 8-44. Retrieved from

http://www.aft.org/pdfs/americaneducator/summer2008/goldenberg.pdf

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard educational review, 81*(4), 710-744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Graus, J., & Coppen, P. A. (2017). The interface between student teacher grammar cognitions and learner-oriented cognitions. *The Modern Language Journal, 101*(4), 643-668. https://doi.org/10.1111/modl.12427

Guthrie, J. T., & Davis, M. H. (2003). Motivating struggling readers in middle school through an engagement model of classroom practice. *Reading & Writing Quarterly, 19*(1), 59-85. https://doi.org/10.1080/10573560308203

*Guthrie, J. T., McRae, A., Coddington, C. S., Lutz Klauda, S., Wigfield, A., & Barbosa, P. (2009). Impacts of comprehensive reading instruction on diverse outcomes of low- and high-achieving readers. *Journal of Learning Disabilities, 42*(3), 195–214. https://doi.org/10.1177/0022219408331039

Hall, M. S., & Burns, M. K. (2018). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology, 66*, 54-66. https://doi.org/10.1016/j.jsp.2017.11.002

Hamayan, E. V., Marler, B., Lopez, C. S., & Damico, J. (2013). *Special Education Considerations for English Language Learners: Delivering a Continuum of Services.* Philadelphia: Caslon Publishing.

Han, D. (2009). Case study of the effects on reading ability and affective factors by shared reading in elementary school English underachiever. Master's thesis, Pusan University of Education.

*Harn, B. A., Linan-Thompson, S., & Roberts, G. (2008). Does additional instructional time make a difference for the most at-risk first graders?. *Journal of Learning Disabilities, 41*(2), 115-125. https://doi.org/10.1177/0022219407313586

Harwell, M., & Maeda, Y. (2008). Deficiencies of reporting in meta-analyses and some remedies. *The Journal of Experimental Education, 76*(4), 403-430. https://doi.org/10.3200/JEXE.76.4.403-430

*Hatcher, P. J., Hulme, C., Miles, J. N., Carroll, J. M., Hatcher, J., Gibbs, S., Smith, G., Bowyer-Crane, C., & Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: a randomised controlled trial. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 47*(8), 820–827. https://doi.org/10.1111/j.1469-7610.2005.01559.x.

*Heo, J. S. (2006). A Study on the Effect of Phonics for Students with Weak English Reading Skills. Master's thesis, Jeju National University of Education. UCI: G901:A-0005313455

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560. https://doi.org/10.1136/bmj.327.7414.557

Higgins, J.P., & Green, S. (2011) *Cochrane Handbook for Systematic Reviews of Interventions.* The Cochrane Collaboration. Retrieved from https://handbook-5-1.cochrane.org/

Hirschstein, M. K., van Schoiack Edstrom, L., Frey, K. S., Snell, J. L., & MacKenzie, E. P. (2007). Walking the talk in bullying

prevention: Teacher implementation variables related to initial impact of the Steps to Respect program. *School Psychology Review, 36*(1), 3-21. https://doi.org/10.1080/02796015.2007.12087949

Hollo, A., & Wehby, J. H. (2017). Teacher talk in general and special education elementary classrooms. *The Elementary School Journal, 117*(4), 616-641. https://doi.org/10.1086/691605

Huedo-Medina, T. B., Sánchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index?. *Psychological methods, 11*(2), 193. https://doi.org/10.1037/1082-989X.11.2.193

Hwang, S. (2015). Understanding Meta-Analysis. Hakjisa.

Im, T. B. (2015). A Study of the Effects of Metacognitive Strategy Instruction on English Reading: With Regard to the High School Low Achievers. Master's thesis, Yonsei University.

Jung, P. G. & Choi, N. (2019). The effect of phonics instruction for English underachiever in elementary school: A meta-analysis. *Journal of Learner-Centered Curriculum and Instruction, 19*(1), 623-649https://doi.org/10.22251/jlcci.2019.19.1.623

*Kang, A. R. (2015). A Study of Effect of English Reading Programs for Underachieving Elementary School Students. Master's thesis, Kookmin University. UCI: I804:11014-000002291539

*Kang, B. K. (2008). A Study on the Improvement of Reading Abilities of Children with Reading Problems through English Phonological Awareness Instructions. Master's thesis, Seoul National University of Education. UCI: G901:A-0005442990

*Kang, S. H. (2010). A Study on Teaching English to Elementary School Underachievers through Analyzing the Actual English Learning Environment. Master’s thesis, Cheongju National University of Education. UCI: G901:A-0005526061

Kangas, S. (2017). "That's Where the Rubber Meets the Road": The intersection of special education and dual language education. *Teachers College Record, 119*(7), 1-36. Retrieved from https://www.researchgate.net/publication/315478019_That's_Where_the_Rubber_Meets_the_Road_The_Intersection_of_Special_Education_and_Dual_Language_Education

*Kaniuka, T. S. (2010). Reading achievement attitude toward reading, and reading self-esteem of historically low achieving students. *Journal of Instructional Psychology, 37*(2), 184-188. Retrieved from https://link.gale.com/apps/doc/A231807638/AONE?u=anon~e8cdd1b3&sid=googleScholar&xid=780caa57

Kavale, K. A. (2001). Meta-analysis: A primer. *Exceptionality, 9*(4), 177-183. https://doi.org/10.1207/S15327035EX0904_2

Kavale, K. A., & Forness, S. R. (2000). What definitions of learning disability say and don't say: A critical analysis. *Journal of Learning Disabilities, 33*(3), 239-256. https://doi.org/10.1177/002221940003300303

*Kim & Jeong (2010). Development of a blended learning model for teaching lower leveled students in elementary English education. Korean Elementary English Education 16(1), 7-30.

Kim, D., An, Y., Shin, H. G., Lee, J., & Park, S. (2020). A meta-analysis of single-subject reading intervention studies for struggling readers: using Improvement Rate Difference (IRD). *Heliyon*, *6*(11), e05024.

Kim D., Shin, H. G., Kim, H., & Cho, E. (2018) An Analysis of the

Effects of the Synthetic Phonics Program on Phonological Awareness Skills of Underachieving Children, *Special Children Education, 20*(2), 25-50.

Kim, D., Kim, H. J., Ahn, S. J., Lee, M. J., Jang, S. Y., An, Y. J., & Shin, H. G. (2017). A Study on the effect of program for the dyslexia and reading difficulties. *The SNU Journal of Education Research, 26*(3), 1-25. UCI: I410-ECN-0102-2018-300-000690266

*Kim, H. R. (2014). Effects of Children's Nursery Rhyme on Vocabulary Ability and Affective Domains of Underachivers at a Korean Elementary School. Master's thesis, Pukyong National University. UCI: G901:A-0005826195

Kim, J. G. (2013). Features and practical ways of underachievers in elementary school English. *Education Theory and Practice, 23*, 3-28. UCI: I410-ECN-0102-2014-300-001642179

*Kim, J. G., & Park, E. S. (2015). Development and application of emotion-based English instructional model for teaching lower-level elementary students. *Journal of Research in Curriculum & Instruction, 19*(2), 523-542. https://doi.org/10.24231/rici.2015.19.2.523

*Kim, J. J. (2012). A Study of the Effects on English Reading Abilities of Underachieved Children Using Phonics. Master's thesis, Wonkwang University. UCI: G901:A-0005773137

Kim, J., Wi, S., & Kim, Y. H. (2015). A meta-analysis of different instructional effects on English underachievers. *Linguistic Research, 32,* 125-150. https://doi.org/10.17250/khisli.32..201507.007

Kim, D., Shin, J., & Lee, D. (2016). Understanding Children with Specific Learning Disabilities. Hakjisa.

*Kim, W. J. (2019). A study on the development and

effectiveness of a growth-mindset intervention program for English language learners in elementary schools. Doctor's thesis, Seoul Women's University. UCI: I804:11036-200000177712

*Kim, Y. J. (2013). A Study on the Improvement of English Reading Abilities Of Underachievers through Phonics Instruction and Theme-Based Teaching. Master's thesis, Pukyong National University. UCI: G901:A-0005768418

*Kim, Y. K. (2014). A Study on Teaching Reading for the Reading Ability Improvement of Slow Learners in Elementary English. Master's thesis, Gyeongin National University. UCI: G901:A-0005800642

Klingner, J. K., & Edwards, P. A. (2006). Cultural considerations with response to intervention models. *Reading Research Quarterly, 41*(1), 108-117. https://doi.org/10.1598/RRQ.41.1.6

Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *Elementary School Journal, 96*, 275-293. https://doi.org/10.1086/461828

Korea Institute for Curriculum and Evaluation (2021). Retrieved from https://www.kice.re.kr/sub/info.do?m=010302&s=kice

Korean Ministry of Education (2019). Retrieved from https://www.moe.go.kr/boardCnts/view.do?boardID=294&boardSeq=77172&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=020402&opType=N

Lee, Y. (2015). A study of teachers' awareness and demand regarding 2009 modified elementary English curriculum. *Primary English Education, 23*(2), 91-117.

Lee, J. Y., & Chang, K. S. (2017). Effects of English language

instruction for underachievers in the elementary school: A meta-analysis. *Primary English Education, 23*(2), 91-118. https://doi.org/10.25231/pee.2017.23.2.91

Lee, J. Y., Chang, K. S., & Jeon, Y. J. (2017). Effects of English language instruction for underachievers in the secondary school: A meta-analysis. *Studies in English Education, 22*(4), 251-278. https://doi.org/10.22275/SEE.22.4.10

Lee, K. Y., Lee, W. K., & Bae, H. S. (2016). A meta-analysis on the effects of phonological awareness instruction. *The Journal of Korea Elementary Education, 27*(4), 359-374. https://doi.org/10.20972/kjee.27.4.201612.359

Lein, A. (2016). Effectiveness of Mathematical Word Problem Solving Interventions for Students with Learning Disabilities and Mathematics Difficulties: A Meta-Analysis. University of Minnesota Ph.D. dissertation. Retrieved from the University of Minnesota Digital Conservancy, https://hdl.handle.net/11299/181764.

Lightbown, P. M., & Spada, N. (2013). *How Languages are Learned (4th edition)*. Oxford Handbooks for Language Teachers.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. SAGE publications, Inc.

Longden, E., Davis, P., Billington, J., Lampropoulou, S., Farrington, G., Magee, F., ... & Corcoran, R. (2015). Shared Reading: assessing the intrinsic value of a literature-based health intervention. *Medical Humanities*, *41*(2), 113-120.

Lovett, M. W., Steinbach, K. A., & Frijters, J. C. (2000). Remediating the core deficits of developmental reading disability: A double-deficit perspective. *Journal of Learning Disabilities, 33*(4), 334–358. https://doi.org/10.1177/002221940003300406

Lovett, M., Meyer, O., & Thille, C. (2008). The open learning initiative: Measuring the effectiveness of the OLI statistics course in accelerating student learning. *Journal of Interactive Media in Education, 14*, 1-16. https://doi.org/10.5334/2008-14

Marchand-Martella, N. E., Klingner, J. K., & Martella, R. C. (2013). *Effective Reading Intervention Practices for English Language Learners*. McGraw Hill Education.

*Martens, B. K., Eckert, T. L., Begeny, J. C., Lewandowski, L. J., DiGennaro, F. D., Montarello, S. A., Arbolino, L. A., Reed, D. D., & Fiese, B. H. (2007). Effects of a fluency-building program on the reading performance of low-achieving second and third grade students. *Journal of Behavioral Education, 16*, 38–53. https://doi.org/10.1007/s10864-006-9022-x

*Mason, L. H., Davison, M. D., Hammer, C. S., Miller, C. A., & Glutting, J. J. (2013). Knowledge, writing, and language outcomes for a reading comprehension and writing intervention. *Reading and Writing, 26*, 1133-1158. https://doi.org/10.1007/s11145-012-9409-0

*Mason, L. H., Snyder, K. H., Sukhram, D. P., & Kedem, Y. (2006). TWA + PLANS strategies for expository reading and writing: Effects for nine fourth-grade students. *Exceptional Children, 73*(1), 69–89. https://doi.org/10.1177/001440290607300104

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*(2), 148-182. https://doi.org/10.1598/HRQ.40.2.2

McCoach, D. B., & Siegle, D. (2003). Factors that differentiate underachieving gifted students from high-achieving gifted

students. *Gifted Child Quarterly, 47*(2), 144-154. https://doi.org/10.1177/001698620304700205

Mittlböck, M., & Heinzl, H. (2006). A simulation study comparing properties of heterogeneity measures in meta-analyses. *Statistics in Medicine, 25*(24), 4321-4333. https://doi.org/10.1002/sim.2692

*Moon, Y. H. (2015). A Study on Improving Communicative Ability for Educationally Retarded Children in Primary School by Using the Conversational English Program. Master's thesis, Chosun University. UCI: G901:A-0005909117

Morrison, F. J., Bachman, H. J., & Connor, C. M. (2008). *Improving literacy in America*. Yale University Press.

*Nam, H. W. (2010). The Effects of English Story Book on the Reading Ability of Elementary School English Underachievers. Master's thesis, Korea Nazarene University.

*Namgung, S. (2012). A Study of the Development of English Teaching Model through Song Activities and its Application to English Underachiever in Elementary School. Master's thesis, Gyeongin National University of Education.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356. https://doi.org/10.1111/j.1467-9817.2004.00238.x

National Center on Intensive Intervention (2012). *Assessment Practices Within a Multi-Tiered System of Supports.* National Center on Intensive Intervention. Retrieved from http://ceedar.education.ufl.edu/tools/innovationconfigurations/

National Reading Panel (US), National Institute of Child Health, &Human Development (US). (2000). Teaching children to read: An evidence-based assessment of the scientific

research literature on reading and its implications for reading instruction. National Institute of Child Health and Human Development, National Institutes of Health. Washington, DC: U.S. Government Printing Office. Retrieved from https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

No, W. K., Park, S. H., & Chang, K. S. (2016). *Research on Educational Support for General High School Low Achievers with Reference to Math and English Subjects (I)*. Seoul: Korea Institute for Curriculum and Evaluation.

O'Connor, R. E., Bell, K. M., Harty, K. R., Larkin, L. K., Sackor, S. M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology, 94*(3), 474–485. https://doi.org/10.1037/0022-0663.94.3.474

*O'Connor, R. E., White, A., & Swanson, N. L. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*(1), 31-46. https://doi.org/10.1177/001440290707400102

*Oostdam, R., Blok, H., & Boendermaker, C. (2015). Effects of individualised and small-group guided oral reading interventions on reading skills and reading attitude of poor readers in grades 2–4. *Research Papers in Education, 30*(4), 427-450. https://doi.org/10.1080/02671522.2014.953195

Orosco, M. J., & Klingner, J. (2010). One school's implementation of RTI with English language learners: "Referring into RTI". *Journal of Learning Disabilities, 43*(3), 269-288. https://doi.org/10.1177/0022219409355474

Ortiz, A. A. (1997). Learning disabilities occurring concomitantly with linguistic differences. *Journal of Learning Disabilities, 30*(3), 321–332.

https://doi.org/10.1177/002221949703000307

Pai, H.H., Sears, D. A., & Maeda, Y. (2015). Effects of small-group learning on transfer: A meta-analysis. *Educational Psychology Review, 27*(1), 79-102. https://doi.org/10.1007/s10648-014-9260-8

Park, S. S., & Park, Y. (2007). A study of using children's literature for underachievers in elementary school English class. *Journal of Learner-Centered Curriculum and Instruction, 7*(2), 95-117. UCI: G704-001586.2007.7.2.001

Partanen, M., & Siegel, L. S. (2014). Long-term outcome of the early identification and intervention of reading disabilities. *Reading and Writing, 27*(4), 665-684. https://doi.org/10.1007/s11145-013-9472-1

Petitti, D. B. (2001). Approaches to heterogeneity in meta-analysis. *Statistics in medicine, 20*(23), 3625-3633. https://doi.org/10.1002/sim.1091

Reynolds, C. R., & Shaywitz, S. E. (2009). Response to Intervention: Ready or not? Or, from wait-to-fail to watch-them-fail. *School Psychology Quarterly, 24*(2), 130. https://doi.org/10.1037/a0016158

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children, 82*(2), 144-169. https://doi.org/10.1177/0014402915585483

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84*(3), 328-364. https://doi.org/10.3102/0034654313500826

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84*(3), 328-364. https://doi.org/10.3102/0034654313500826

Scheffel, D. L., Shaw, J. C., & Shaw, R. (2008). The efficacy of a supplementary multisensory reading program for first-grade students. *Journal of Reading Improvement, 45*(3), 139-152. Retrieved from https://kb.gcsu.edu/eds/12.

*Seaman, J. (2015). The Effects of Reading Lab Intervention on Elementary English Language Learner Dibels Scores. Doctor's thesis, Northern Arizona University. Retrieved from https://www.proquest.com/dissertations-theses/effects-reading-lab-intervention-on-elementary/docview/1694579829/se-2?accountid=6802

Seo, H. & Bae, J. (2015). Effects of early word reading skills from English literacy based on Self-Determined Learning Model Intervention (SDLMI). *Journal of Curriculum and Instruction, 19*(3), 765-790.

Shin, J., & McMaster, K. (2019). Relations between CBM (oral reading and maze) and reading comprehension on state achievement tests: A meta-analysis. *Journal of School Psychology, 73*, 131-149. https://doi.org/10.1016/j.jsp.2019.03.005

*Shin, S. J. (2017). Applying Cognitive Science into English Reading Practice. Doctoral thesis, Yonsei University.

Slavin, R. E., & Cheung, A. (2003). *Effective Reading Programs for English Language Learners. A Best-Evidence Synthesis.* Association for Supervision and Curriculum Development. https://doi.org/10.1080/15235882.2005.10162835

Snow, C. E., Griffin, P. E., & Burns, M. (2005). *Knowledge to*

support the teaching of reading: Preparing teachers for a changing world. Jossey-Bass. ISBN: 978-0-787-97465-7

*Song, J. M. (2015). The effectiveness of using decodable books to teach elementary school students with limited English reading abilities. *Studies in English Language & Literature, 57*(1), 369-390. https://doi.org/10.18853/jjell.2015.57.1.018

Swanson, H. L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcomes*. NY: Guilford Press. https://doi.org/10.5860/choice.37-2302

National Institute of Child Health and Human Development. (2000). *Report of the national reading panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for instruction. Reports of the subgroup.* NIH Publication No. 00-4754. Washington, DC: Government Printing Office. Retrieved February 10, 2009, from http//nichd.nih.gov/publicationa/nrp/report.htm

Torgesen, J. K., & Wagner, R. K. (1998). *Alternative diagnostic approaches for specific developmental reading disabilities.* Learning Disabilities Research & Practice.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Publishing Company Reading, Mass. https://doi.org/10.2307/2529486

*Uhm, W. Y., & Kim, H. R. (2014). The effect of strategies based leveled English reading programs on reading proficiency and affective domains of underachieving elementary school students. *Foreign Languages Education, 21*(4), 173-200http://dx.doi.org/10.15334/FLE.2014.21.4.173

Vacca, R. T., & Vacca. J. A. (1999). *Content Area Reading: Literacy and Learning Across the Curriculum.* (6th Edition).

Pearson. ISBN: 978-0321003638

*Vadasy, P. F., & Sanders, E. A. (2008). Repeated reading intervention: Outcomes and interactions with readers' skills and classroom instruction. *Journal of Educational Psychology, 100*(2), 272-290. https://doi.org/10.1037/0022-0663.100.2.272

Vadsay, P. F., & Sanders, E. A. (2008). Benefits of repeated reading intervention for low-achieving fourth- and fifth-grade students. *Remedial and Special Education, 29*(4), 235-249. https://doi.org/10.1177/0741932507312013

Vanchu-Orosco, M. (2012). A Meta-Analysis of Testing Accommodations for Students with Disabilities: Implications for High-Stakes Testing. Doctoral dissertation, University of Denver. Retrieved from https://www.researchgate.net/publication/301232647_A_meta-analysis_of_testing_accommodations_for_students_with_disabilities_Implications_for_high-stakes_testing

*Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagan, E., & Francis, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal, 107*(2), 153-180. https://doi.org/10.1086/510653

Verhallen, M., & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied linguistics, 14*(4), 344-363. https://doi.org/10.1093/applin/14.4.344

Vernon-Feagans, L., Gallagher, K., Ginsberg, M. C., Amendum, S., Kainz, K., Rose, J., & Burchinal, M. (2010). A diagnostic teaching intervention for classroom teachers: Helping struggling readers in early elementary school. *Learning*

Disabilities Research & Practice, 25(4), 183-193. https://doi.org/10.1111/j.1540-5826.2010.00316.x

*Wagner, D. L. (2011). The Relative Effects of Various Intervention Approaches on the Reading Fluency and Comprehension of Fifth and Sixth Grade Struggling Readers. Doctor's thesis, University of Minnesota. https://doi.org/10.1080/02702711.2014.927812

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of tier 2 type reading interventions in grades K-3. *Educational psychology review*, *28*(3), 551-576.

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. *Journal of Learning Disabilities, 51*(6), 612-624. https://doi.org/10.1177/0022219418775110

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing, 23*(8), 889-912. https://doi.org/10.1007/s11145-009-9179-5

Wilson, S. J., Tanner-Smith, E. E., Lipsey, M. W., Steinka-Fry, K., & Morrison, J. (2011). Dropout prevention and intervention programs: Effects on school completion and dropout among school-aged children and youth. *Campbell Systematic Reviews*, *7*(1), 1-61.

Wolf, M. A. (2012). Culture Shift: Teaching in a Learner-Centered Environment Powered by Digital Learning. Digital Learning Series. *Alliance for Excellent Education*.

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal*

of Learning Disabilities, 33(4), 387-407. https://doi.org/10.1177/002221940003300409

Yeo, S. (2010). Longitudinal trajectories of self-efficacy among students with learning disabilities, *Journal of Special Education, 11*(4), 497-514.

Yoon. H. (2002). A study on the applications of storytelling for expansion and remediation of English education in elementary school. (mater's thesis). Korea National University of Education.

Yoon, K. J., & Lee, J. K. (2017). An analysis of teacher's recognition for phonics improvement in elementary English education. *Journal of Learner-Centered Curriculum and Instruction, 17*(9), 515-540. https://doi.org/10.22251/jlcci.2017.17.9.515

Ysseldyke, J. E., Algozzine, B., Shinn, M. R., & McGue, M. (1982). Similarities and differences between low achievers and students classified learning disabled. *The Journal of Special Education, 16*(1), 73-85. https://doi.org/10.1177/002246698201600108

Zheng, X., Flynn, L. J., & Swanson, H. L. (2013). Experimental intervention studies on word problem solving and math disabilities: A selective analysis of the literature. *Learning Disability Quarterly*, *36*(2), 97-111.

Zigmond, N. (2003). Where should students with disabilities receive special education services? Is one place better than another?. *The Journal of Special Education, 37*(3), 193-199. https://doi.org/10.1177/00224669030370030901

Appendix A

Codebook with Descriptors

Category	Definition of Descriptors
Participants	
Grade	The current grade level in elementary school
Grade level	Lower (1st and 2nd); Mid-upper (3rd to 6th)
Learner type	Underachiever (UA); Low achiever (LA); English as Foreign Language (EFL); English Language Learners (ELL), English as a Second Language (ESL); Specific Language Impairment (SLI)
Classroom type	Inclusion (special education), General
Country / City / Region	Name of country, city, region if stated
Intervention	
Intervention components	All names of intervention as independent variables (e.g. repeated reading, phonics, direct teaching)
Intervention component type	All interventions were grouped into 1) content-based, 2) learner approach-based, or 3) teacher/instructional method-based
Interventionist	Teacher; Researcher; Bilingual teacher; Teaching Assistant or paraprofessionals (TA); Research Assistant (RA)
Minutes	Actual minutes spent for intervention
Weekly frequency	Sessions per week
Duration	Total number of weeks
Session number	Total sessions of intervention
Classroom type	Whole class; Inclusion class (defined by researcher)
Outcome measures	

Outcome measure domain	Measured outcome domains (e.g. fluency, motivation, etc.)
Outcome measure type	All measures were group into 1) affective (e.g. motivation, attitude), or 2) achievement type (e.g. fluency)
Assessment tool	Researcher-developed or standardized
Study level	
Publication information	Journal publisher; Dissertation
Study design information	Pre and posttest control group; posttest only control group; One group (single group) pre and posttest

Appendix B

Summary of Studies Included

	Study Name	Participants	Learner type	Intervention	Interventionist	Intervention frequency (per week; total session #)	Outcome measures	Affective measure descriptions	Study Design
1	Kang, B. (2008)*	N=8; Gr.6;	UA (EFL)	Phonics	Researcher	3/wk; 30 sessions	PA, FL, Affective	Learning attitude	One group pre-posttest
2	Kang, S. (2010)*	N=15, Gr.4(5), 5(5), 6(5)	UA (EFL)	Phonics/Eng song	Researcher	2/wk; 8-15 sessions	LC, RC, Affective	Motivation, confidence	One group pre-posttest
3	Kim, Y.J. (2013)*	N=5, Gr.5(5)	UA (EFL)	Phonics	Researcher	3/wk; 12 session	SPK, WR, Word ID,		One group pre-posttest
4	Kim, Y.K. (2014)*	N=5, Gr.6(5)	UA (EFL)	Word list/ Phonics	Researcher	2/wk; 46 sessions	LT, FL, PA, VOC, RC, FL		One group pre-posttest
5	Kim & Jeong (2010)	N=5, Gr.6(5)	UA (EFL)	Multisensory/ Repetitive	Researcher	3/wk; 6 sessions	CGN	Cog abilities	One group pre-posttest

6	Kim, J. (2013)*	N=11, Gr.6(11)	UA (EFL)	Phonics	Researcher	1/wk; 15 sessions	LT, PA, SPL	Letter order, Upper & lower case	One group pre-posttest
7	Kim & Park (2015)	N=43, Gr.4(22), Gr.5(21)	UA (EFL)	Feeling-based	Researcher	2/wk; 28 sessions	RC, Affective	learning time, satisfaction rating, self-driven, self-esteem, motivation, competitiveness, attitude	One group pre-posttest
8	Kim, H. (2014)*	N=5, Gr.5(5)	UA (EFL)	Eng song	Researcher	5/wk; 28 sessions	FL		One group pre-posttest
9	Nam, H. (2010)*	N=10, Gr.6(10)	UA (EFL)	Shared reading	Researcher	5 wk; 40 sessions	FL		One group pre-posttest
10	Moon (2015)*	N=12, Gr.5(6), Gr.6(6)	UA (EFL)	Role play/Repetitive	Researcher	2 wk; 48-60 sessions	SPK		One group pre-posttest
11	Song (2015)	N=9, Gr.4(3), Gr.5(3), Gr.6(3)	UA (EFL)	Phonics	Researcher	1 wk; 12 sessions	PA		One group pre-posttest

12	Shin (2017)*	N=22, Gr.5(22)	UA (EFL)	Phonics/Strategies	T & R	2/wk; 23 sessions	CGN	cog. flexibility	One group pre-posttest
13	An (2016)*	N=8, Gr.4(8)	UA (EFL)	Supplementary WB	Researcher	2/wk; 8 sessions	Affective	Changes in mindset	One group pre-posttest
14	Uhm & Kim (2014)	N=8, Gr.5(8)	UA (EFL)	Phonics/Voc	Teacher	1/wk; 16 sessions	FL, WR, Affective	Motivation, confidence, participation	One group pre-posttest
15	Cho* (2015)	N=4, Gr.3(4)	UA (EFL)	Shared reading	T & R	3/wk; 42 sessions	FL, VOC, SPK, Affective	Motivation, confidence, participation	One group pre-posttest
16	Seaman (2015)*	N=117, Gr.3-4(117)	ELL	Lab-instruction	Teacher	5/wk; 260 sessions	FL		One group pre-posttest
17	Mason et al. (2006)	N=9, Gr.5, LA (5) + LD (1), LD/SLI (1), EBD (2)	ELL	Strategy	Researcher	3/wk; 15 sessions	SPK, WR	Oral retell	One group pre-posttest
18	Cheon (2003)*	N=64, Gr.5(64)	UA (EFL)	Shared reading	Researcher	1/wk; 9 session	FL, Affective	Motivation, confidence	Pre-posttest control group
19	Heo (2006)*	N=45, Gr.5(45)	UA (EFL)	Phonics	Researcher	2/wk; 20 sessions	FL, Affective	Motivation, confidence	Pre-posttest control group

20	Kang, A. (2016)*	N=20, Gr.4(20)	UA (EFL)	Phonics	Researcher	3/wk; 36 sessions	RC, LT, VOC, PA, Affective	Motivation, confidence	Pre-posttest control group
21	Kim, W.J. (2019)*	N=16, Gr.5(16)	UA (EFL)	Growth mindset	Researcher	1/wk; 8 sessions	Affective	Learning belief, determination	Pre-posttest control group
22	NamGoong. S. (2012)*	N=18, Gr.5(18)	UA (EFL)	Eng song	Researcher	1/wk; 14 sessions	LC, Affective	Motivation, confidence	Pre-posttest control group
23	Bae (2015)*	N=56, Gr.5(56)	UA (EFL)	Small group collab	Teacher	2/wk; 12 sessions	SPK, Affective	Motivation, learning attitude	Pre-posttest control group
24	O'Connor et al. (2007)	N=37, Gr.2-4(37)	LA	Repeated/Continuous	Teacher	3/wk; 42 sessions	FL, RC, Word ID, PA		Pre-posttest control group
25	Oostdam et al. (2015)	N=126, Gr.2-4(126)	UA	Repeated	Other	4/wk; 48 sessions	FL, RC, VOC, Affective		Pre-posttest control group
26	Vaughn et al. (2006)	N=46, Gr.1(46)	ELL	Systematic/Explicit	Bilingual Teacher	5/wk; 100 sessions	LT, PA, LC, FL, RC		Pre-posttest control group
27	Guthrie et al. (2009)	N=62, Gr.5(62)	LA	Concept-orientedrd instruction	Teacher	3/wk; 36 sessions	RC, VOC, FL		Pre-posttest control group

28	Harn et al. (2008)	N=54, Gr.1(54)	LA	Intensive reading	Other	5/wk; 120 sessions	RC, PA, FL		Pre−posttest control group
29	Hatcher et al. (2006)	N=77, Gr.1(77)	LA	Small group collab	Teacher	2−3/wk 50 sessions	PA, LT, FL		Pre−posttest control group
30	Martens et al. (2007)	N=30, Gr.2(30)	LA	FL program	Other	3 wk; 18 sessions	FL		Pre−posttest control group
31	Mason et al. (2012)	N=9, Gr.4(9), LA(5), LD(1)	LA	TWA (Thinking strat)	Other	1−2 wk; 10 sessions	WR, RC, STN, VOC		Pre−posttest control group
32	Vadsay & Sanders (2008)	N=119, Gr.4−5(119)	ELL	FL program	Teacher	4 wk; 72 sessions	VOC, FL, RC		Pre−posttest control group
33	Jeong and Kim (2017)	N=12, Gr.5(12)	UA (EFL)	Multiple Intelligence	T & R	2/wk; 26 sessions	RC, FL, Affective	Motivation, confidence, participation	Posttest only control group
34	Wagner (2011)*	N=29, Gr.5−6(29)	LA	Immediate instrc.	Researcher	2/wk; 6 sessions	VOC, FL, RC		Posttest only control group
35	Kaniuka (2010)	N=85, Gr.3−5(85); LD(9)	LA	Remedial reading program	Researcher	NR	Affective	Attitude, self−esteem	Posttest only control group

Appendix C

Study participants' learner and classroom type

	Study Name	Participants	Learner type	Classroom type
1	Kang, B. (2008)*	N=8; Gr.6;	UA (EFL)	Inclusion
2	Kang, S. (2010)*	N=15, Gr.4(5), 5(5), 6(5)	UA (EFL)	Inclusion
3	Kim, Y.J. (2013)*	N=5, Gr.5(5)	UA (EFL)	Inclusion
4	Kim, Y.K. (2014)*	N=5, Gr.6(5)	UA (EFL)	Inclusion
5	Kim & Chung (2010)	N=5, Gr.6(5)	UA (EFL)	Inclusion
6	Kim, J. (2013)*	N=11, Gr.6(11)	UA (EFL)	Inclusion
7	Kim & Park (2015)	N=43, Gr.4(22), Gr.5(21)	UA (EFL)	Inclusion
8	Kim, H. (2014)*	N=5, Gr.5(5)	UA (EFL)	General
9	Nam (2010)*	N=10, Gr.6(10)	UA (EFL)	General
10	Moon (2015)*	N=12, Gr.5(6), Gr.6(6)	UA (EFL)	General
11	Song (2015)	N=9, Gr.4(3), Gr.5(3), Gr.6(3)	UA (EFL)	General
12	Shin (2017)*	N=22, Gr.5(22)	UA (EFL)	General

13	Ahn (2016)*	N=8, Gr.4(8)	UA (EFL)	General
14	Uhm & Kim (2014)	N=8, Gr.5(8)	UA (EFL)	General
15	Cho* (2015)	N=4, Gr.3(4)	UA (EFL)	General
16	Seaman (2015)*	N=117, Gr.3-4(117)	ELL	General
17	Mason et al. (2006)	N=9, Gr.5, LA (5) + LD (1), LD/SLI (1), EBD (2)	ELL	General
18	Cheon (2003)*	N=64, Gr.5(64)	UA (EFL)	General
19	Hur (2006)*	N=45, Gr.5(45)	UA (EFL)	General
20	Kang (2016)*	N=20, Gr.4(20)	UA (EFL)	General
21	Kim, W.J. (2019)*	N=16, Gr.5(16)	UA (EFL)	Inclusion
22	Nam, G.S. (2012)*	N=18, Gr.5(18)	UA (EFL)	General
23	Bae (2015)*	N=56, Gr.5(56)	UA (EFL)	General
24	O'Connor et al. (2007)	N=37, Gr.2-4(37)	LA	General
25	Oostdam et al. (2015)	N=126, Gr.2-4(126)	UA	General
26	Vaughn et al. (2006)	N=46, Gr.1(46)	ELL	General

27	Guthrie et al. (2009)	N=62, Gr.5(62)	LA	General
28	Harn et al. (2008)	N=54, Gr.1(54)	LA	General
29	Hatcher et al. (2006)	N=77, Gr.1(77)	LA	General
30	Martens et al. (2007)	N=30, Gr.2(30)	LA	General
31	Mason et al. (2012)	N=9, Gr.4(9), LA(5), LD(1)	LA	General, Inclusion
32	Vadsay & Sanders (2008)	N=119, Gr.4-5(119)	ELL	General, Inclusion
33	Chung and Kim (2017)	N=12, Gr.5(12)	UA (EFL)	General
34	Wagner (2011)*	N=29, Gr.5-6(29)	LA	General
35	Kaniuka (2010)	N=85, Gr.3-5(85); LD(9)	LA	General

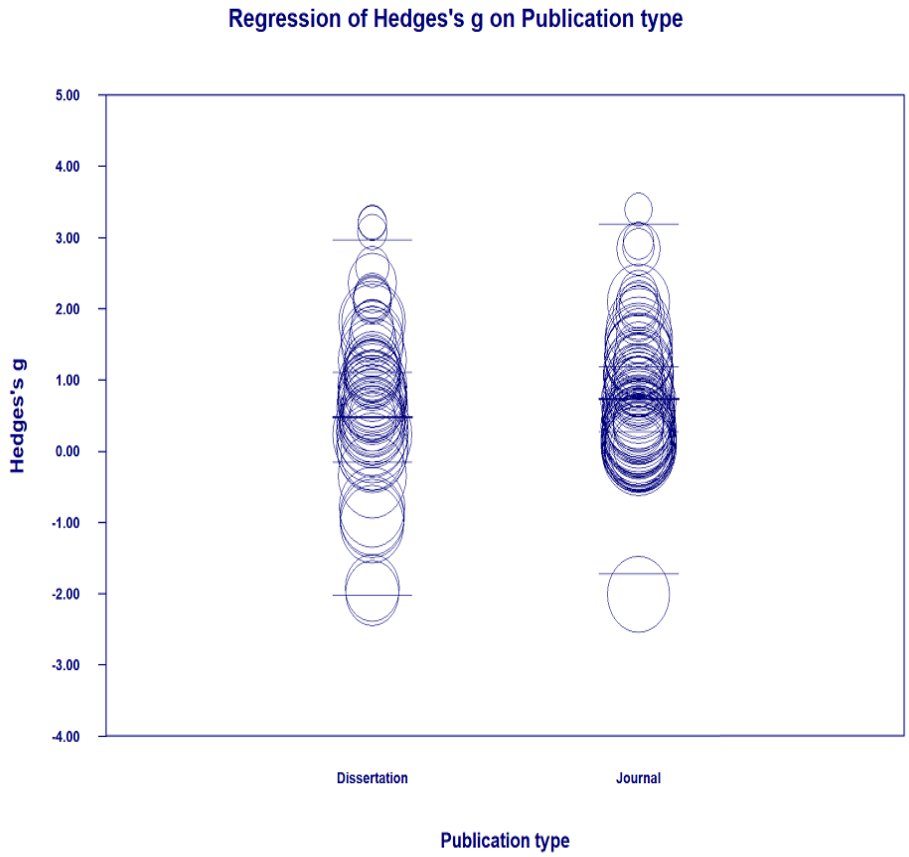
Appendix D

Intervention component type

Intervention Component Type			
Content-based	Learner approach-based	Teacher/Instructional-based	Others
Phonics Fluency program Mixed (any combination)	Multisensory Feeling-based Role- play Song Multi-intelligence Strategies Mindset Concepts approach	Repeated instruction Shared reading Collaborative learning Small group Explicit Intensive Immediate	Eclectic Remedial Workbook Lab-based

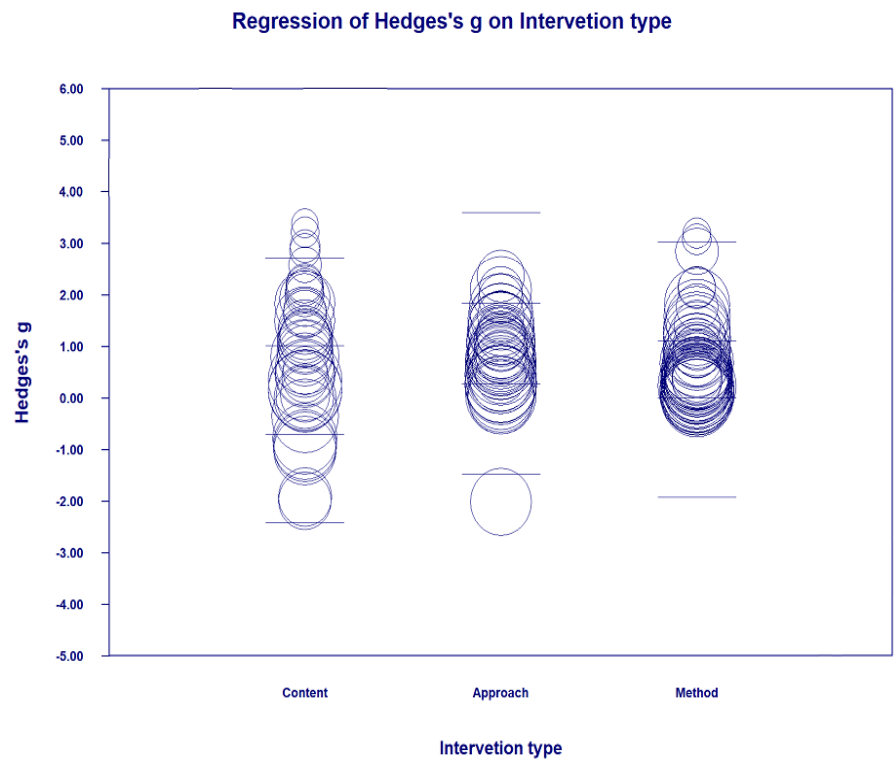
Appendix E (1)

Scatter plot of publication type



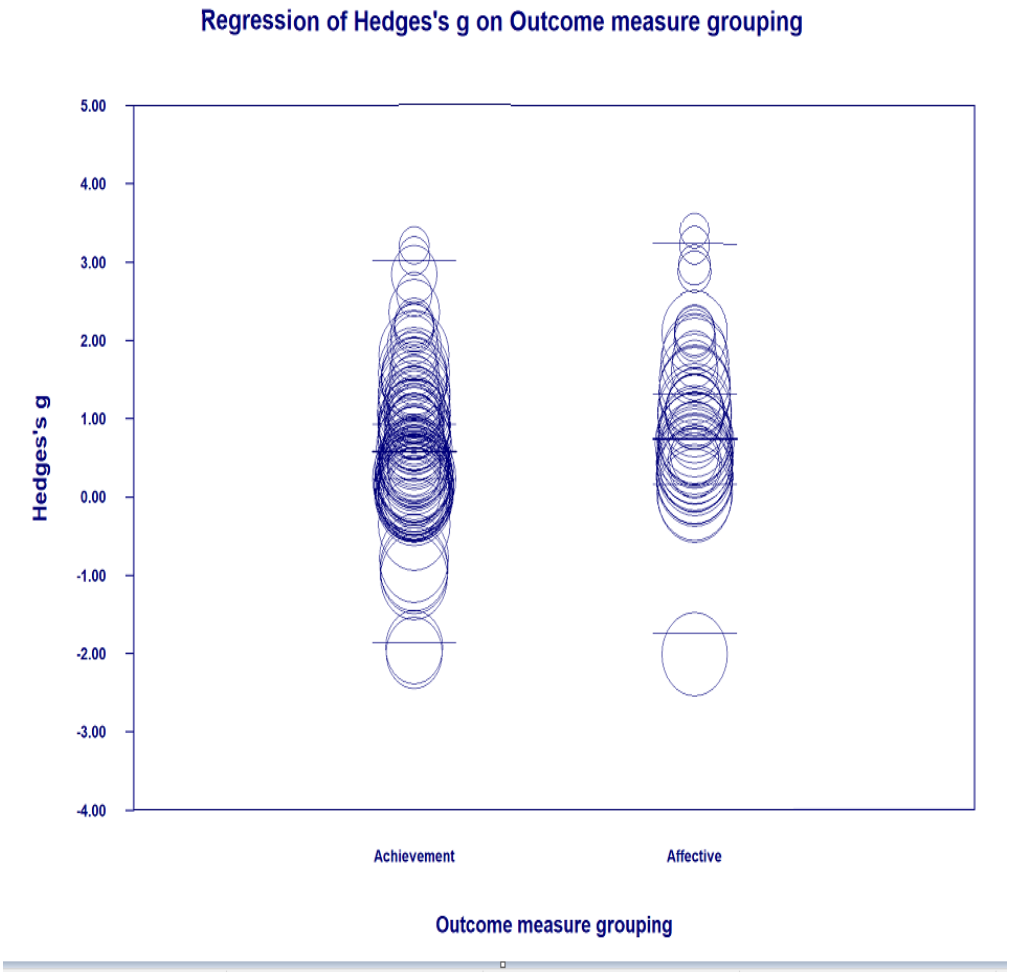
Appendix E (2)

Scatter plot of intervention type



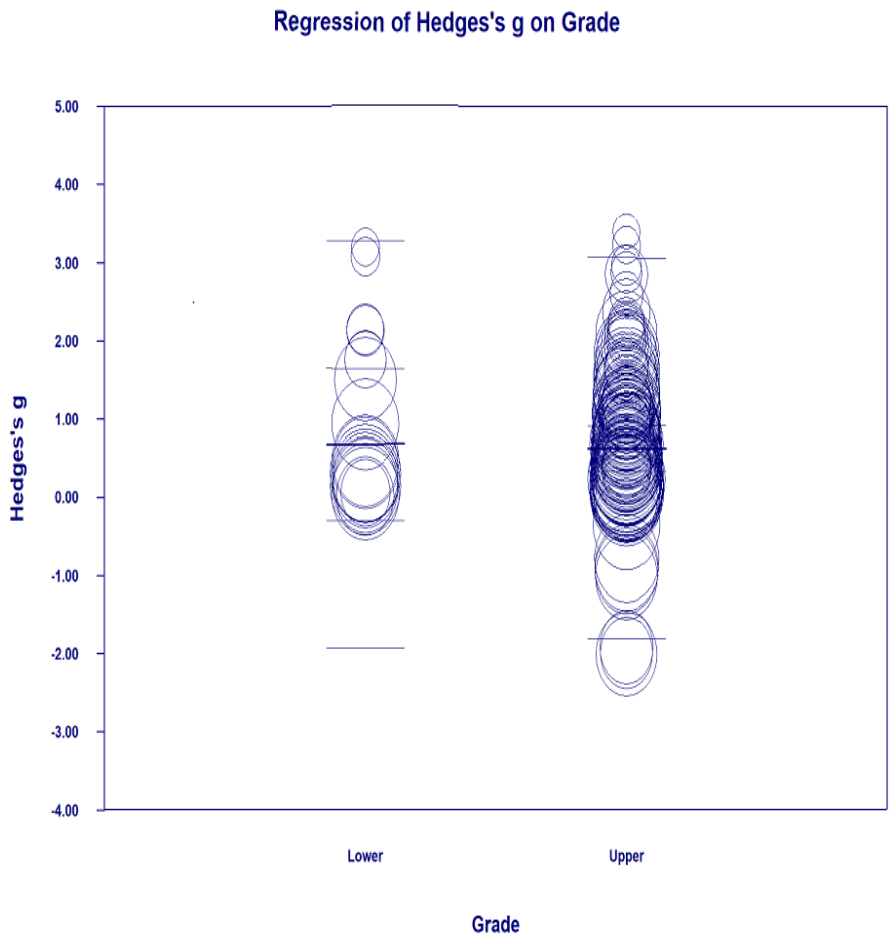
Appendix E (3)

Scatter plot of outcome measure type



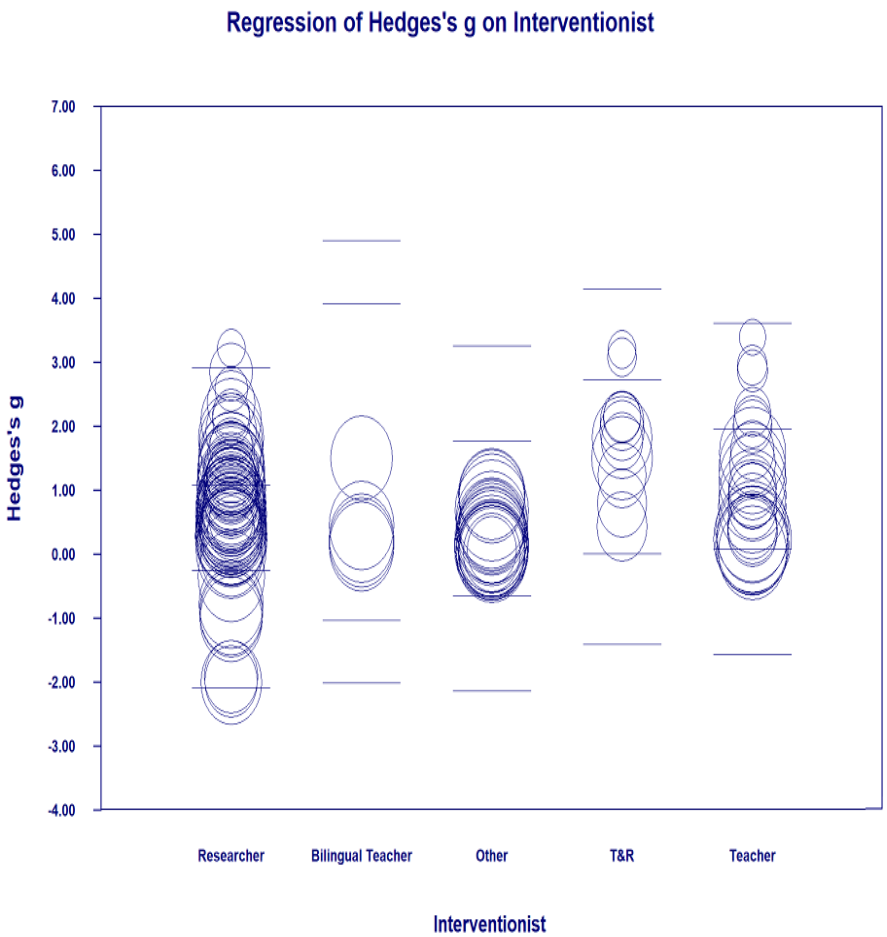
Appendix E (4)

Scatter plot of grade level range



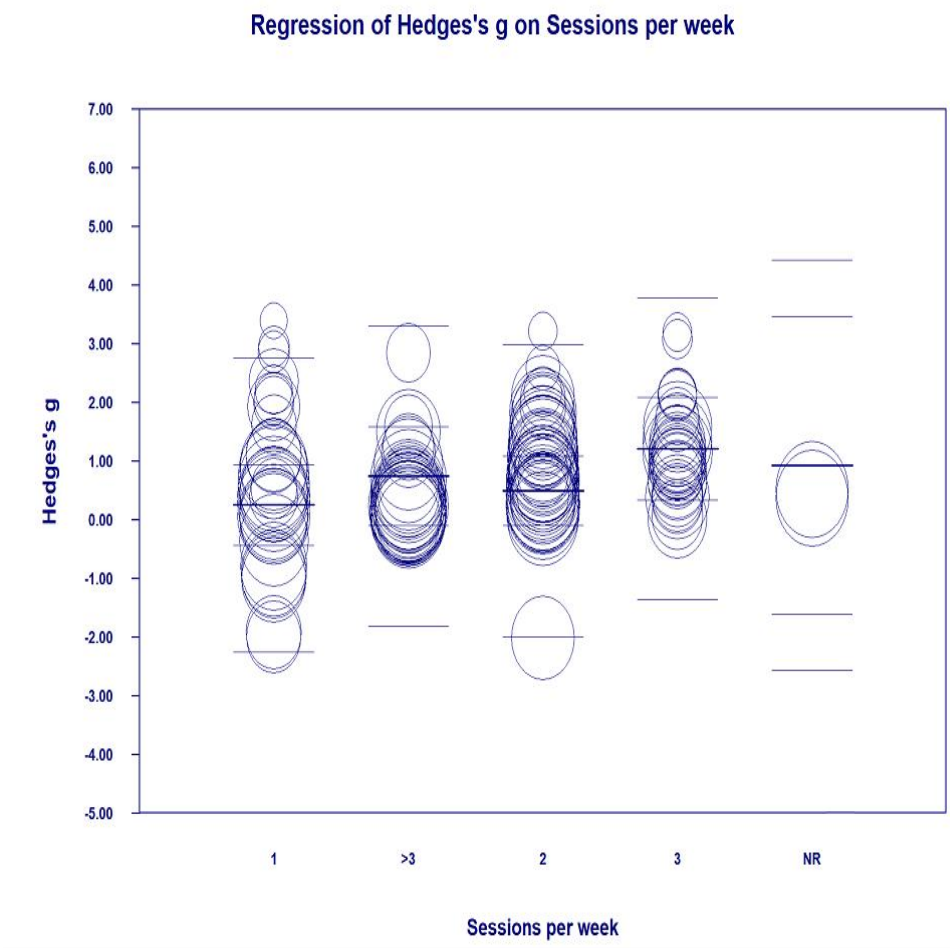
Appendix E (5)

Scatter plot of interventionist



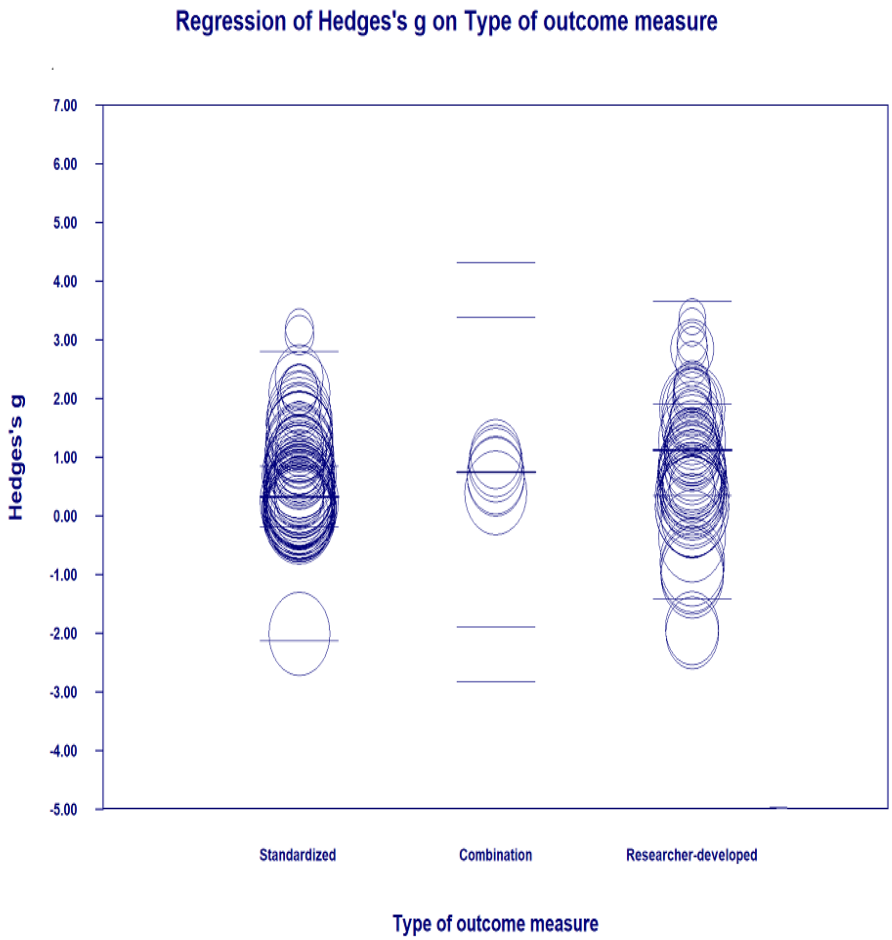
Appendix E (6)

Scatter plot of number of sessions per week



Appendix E (7)

Scatter plot of outcome measure type



국문 초록

본 메타분석 논문에서는 영어 읽기에 어려움이 있는 초등학생들을 대상으로 영어 중재가 미치는 효과를 확인하였다. 실험 및 준 실험 연구 설계를 사용한 총 35편의 분석대상논문과 150개의 효과크기로부터 연구결과가 도출되었다. 무선효과모형 분석 결과에 따르면, 이상치를 제외한 후 전체 평균 효과크기는 Hedge's g=0.655 (SE=0.054)로 나타났고, 다양한 하위집단으로부터 유의한 조절변수의 효과를 면밀히 탐색하기 위하여 조절효과 분석을 실시하였다. 하위집단 분석에 활용된 변수는 다음과 같다: (1) 연구참여자 관련 변수(학급 유형, 학년 수준, EFL(English as a Foreign Language) 학습자 / ELL(English Language Learner) 학습자 / 영어가 모국어인 학습자 중 저성취 학습자 등), (2) 중재결과 관련 변수(중재결과 측정, 결과 유형 등), (3) 중재 구성요소(중재 유형 등), (4) 상황적 특성(중재자, 중재 빈도, 중재 기간 등). 더불어, 공변량에 미치는 효과 정도를 탐색하고 산점도를 확인하기 위하여 메타회귀분석도 실시하였다. 연구결과에 따르면, 학습자 유형, 중재결과 측정 유형(정의적 평가 등), 중재자 유형(교사 등), 특정한 중재(다감각/지능 중재, Shared-reading 중재, 노래 활용 중재, 감정 기반 중재 등), 중재 빈도의 경우 중재 효과에 있어 유의한 조절변수로 나타난 반면, 학년 수준 차이, 학급 유형, 중재 유형의 경우 유의하지 않은 것으로 나타났다. 이를 바탕으로 다양한 유형의 학습자를 위한 교수의 중요성, 보다 효과적인 중재 접근과 관련한 교육적 시사점과 연구의 제한점을 논의하였다.

Keyword: 메타분석, 영어 읽기 중재, 특수교육, 일반교육 환경, 조절변수 분석, 저성취 영어학습자, 영어부진아, 초등영어교육, EFL, ELL, ESL.