THE EFFECTS OF CODE-BASED LITERACY INTERVENTIONS ON SPELLING ACHIEVEMENT: A META-ANALYSIS

by

JOY LIN

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2013

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Dr. Patricia J. Brooks

Date

Chair of Examining Committee

Dr. Maureen O'Connor

Date

Executive Officer

Dr. Joseph Glick

Dr. Bruce Homer

Dr. Sarah Berger

Dr. Jay Verkuilen

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE EFFECTS OF CODE-BASED LITERACY INTERVENTIONS ON SPELLING ACHIEVEMENT: A META-ANALYSIS

by

Joy Lin

Adviser: Dr. Patricia J. Brooks

Poor spelling is a pervasive problem among children and adults alike. Yet despite an abundance of research on reading development, there is a surprising lack of emphasis on spelling, a related and arguably equally important skill. Given that research in the past two decades has established the importance of code-based knowledge such as phonological and alphabetic knowledge in reading, it is reasonable to suspect that code-based knowledge would also impact spelling ability. However, few intervention studies have directly addressed spelling. To investigate this issue, this study utilizes meta-analysis to quantitatively assess the effects of systematic code-based literacy interventions on spelling achievement. Studies included in the meta-analysis were published in English, involved a code-based literacy intervention in a school setting, included a control or comparison group, measured spelling as an outcome at post-test, and reported sufficient statistics to compute effect sizes. Random effects analysis models were based on 91 studies and 153 computed effect sizes. The mean effect size for all studies was moderate, d = 0.58, 95% CI [.49, .67], indicating that systematic code-based literacy instruction is more effective at improving spelling outcomes than non-code-based or less systematic code-based instruction. The total sample size was 9,341 participants in pre-school to Grade 11.

These findings directly oppose the claim that learning to spell is a passive process that occurs in all literacy contexts, as well as lay assumptions that knowledge about spelling does not

need to be taught. Rather, the findings from this meta-analysis suggest that students who had received interventions that incorporated explicit instruction in phonological, orthographic, morphological knowledge fared better than their control-group peers on spelling outcomes. However, given the limitations of meta-analysis, further research is needed to substantiate these findings. Overall, evidence from this meta-analysis suggests that spelling, like reading, is best improved by explicit instruction in linguistic knowledge.

ACKNOWLEDGEMENTS

Gratitude is the best (yet still inadequate) summation of what I wish to express to the people who supported me through this endeavor. To my best friend and husband Jerry, Christiana, and the Chung family…thank you.

Neither would this work have been possible without the support and insight of the knowledgeable scholar-professionals who comprised my committee. To my advisor Dr. Patricia J. Brooks, Dr. Joseph Glick, Dr. Bruce Homer, Dr. Sarah Berger, and Dr. Jay Verkuilen…Thank you very much. I hope to someday pay it forward.

Finally, thank you, Kasey Powers, for your tremendous help in coding studies for reliability.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: Introduction

Statement of the Problem

It is neither surprising nor challenging to find spelling errors in the grade school classroom and beyond-- in advertisements, official documents, emails, and in practically any context that involves writing. These spelling errors include the inevitable typo or misprint, which no amount of spelling knowledge can preclude, but this is certainly not all. In many cases these spelling errors stem from confusion or lack of awareness about the English writing system (i.e., orthography). Thus, grade school children may struggle to grasp the fact that the suffixes such as *-sion, -tion,* and *-cian* can all sound like *shun* (or /ʃ ə n/ rendered here in IPA format, 1999), and even educated adults may pause over words with multiple doubled consonants such as *accommodation*.

While some may easily regard spelling as little more than an anachronism, few can argue against the increasingly significant role that written communication plays in society today. In this mobile, pervasively text-rich world, written communication skills are now more essential than ever. Many people communicate as they travel from place to place, which means that they write (as well as speak and read) almost as frequently as they move. In many scenarios, misspellings can be more than an eye sore. Far from being a trivial nuisance, poor spelling can cause businesses to lose revenue, job candidates to be rejected for employment, and can sometimes give the impression of ignorance or low intelligence.

Still, one might question the usefulness of spelling knowledge in light of the availability of spell-check, grammar-check, and auto-complete functions on word processing software. While such tools serve their function well, it is clear that the availability of these tools does not preclude widespread spelling errors beyond typos. Although spell-checkers almost unfailingly

detect misspellings of words like *photosynthesis*, they cannot reliably take the place of a competent English speller who knows, for instance, the semantic difference between homonyms such as *base* and *bass*. Similarly, a student writing a term paper on Australian aboriginals' rites of passage would gain little help from either the spell- or grammar-check if she had confused *rites* with *rights*. Of course, there is another disadvantage of relying on such software applications, which is that they are not always accessible or convenient. Spontaneous, handwritten notes continue to be a common and useful way to communicate, and anyone would be well served to be able to write a memo without having to consult a dictionary or acquaintance.

Yet there is a more pressing issue, which is that while adults continue to stumble over words like *embarrass* and *misspell*, children are apparently resigned to the same fate, as many children in the United States typically receive little explicit spelling instruction as part of their regular language arts curriculum. Few children receive any spelling instruction at all after formal literacy instruction ends, which typically occurs by Grade four. The assumption behind this instructional practice is that children will naturally learn how to spell by reading and writing. While it is true that children learn about some aspects of spelling words by exposure to print, there is little evidence that incidental and informal methods of spelling instruction have a lasting impact on spelling achievement (Graham, 2000). On the contrary, many aspects of English orthography (i.e., the structure and patterns of the written language) are not easily learned without explicit instruction.

Given the complexity of the English writing system (Ziegler & Goswami, 2005), it is more likely that at least some explicit instruction is required in order to become a competent speller. Relative to other alphabetic languages with more transparent letter-sound correspondence (e.g., Greek and Italian), English is considered an *opaque* orthography (Frost,

1992). Thereby, each phoneme (i.e., the smallest unit of spoken language) can be represented by multiple graphemes (i.e. the letters, or units of written language), and each grapheme can represent multiple phonemes. For example, /a/ can be written in two different ways, such as in the words *main* and *mane*. 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.*

In fact, there are approximately twice as many ways to translate from sound to spelling than there are from spelling to sound (Venezky, 1970; Ziegler, Stone, & Jacobs, 1997). As such, English spelling patterns comprise a minefield of irregular spellings and conditional rules, all of which require years of practice to master. Even fifth-grade students who are competent readers still have difficulty with relatively basic spelling patterns, such as the convention of adding an "s" or "es" to make a regular noun plural (Nunes & Bryant, 2009). Similarly, some college-educated adults may lack the morphological (i.e., related to morphemes, the smallest written unit of meaning, such as prefixes and suffixes) knowledge required to be able to recognize this same basic pattern, in which case they may instead rely on word-specific knowledge to spell words (Mitchell, Kemp, & Bryant, 2011).

Thus, it is evident that becoming a proficient speller of English is no small feat, even for otherwise typically achieving students and adults. Rather than being a problem associated with only certain subgroups, spelling difficulty may manifest in varying degrees in many individuals, not simply those who already struggle to read. Yet spelling has been and continues to be relegated to unimportance, such that poor spelling is often overlooked or simply excused, and spelling instruction is deemed an anachronism. Indeed, despite the ubiquity of spelling errors made by children and adults alike, there is a surprising neglect of the topic of spelling in society.

The widespread disregard of spelling is at once problematic and puzzling, particularly given the moderate to strong correlation between spelling and reading ability (Shankweiler, Lundquist, Dreyer, & Dickinson, 1996; Swanson, Trainin, Necoechea, & Hammill, 2003). Reading and spelling are interrelated processes, such that development in one area strengthens the other (Ehri, 1997). In the context of this relationship, poor spelling may often be an indicator of underlying deficits, such as in phonological awareness (i.e., the knowledge that words are comprised of units of sound, such as phonemes and syllables) and in word decoding skills (Shankweiler et al., 1996). For other individuals who have developed a superficial level of competency in decoding, weakness in spelling may suggest a lack of orthographic knowledge and linguistic sophistication. In many ways, spelling is a more sensitive measure of linguistic ability than reading, and to ignore spelling in cases such as these would be to possibly miss an opportunity for remediation.

Strangely, the importance of spelling is not necessarily reflected in the educational research community. Spelling is not commonly viewed as a topic worthy of investigation, as it is often addressed as an ancillary issue, if at all, in studies of reading. For instance, in the congressionally charged review of best practices for reading instruction conducted by the National Reading Panel, spelling is omitted as a one of five focal points of research (i.e., alphabetics, fluency, comprehension, teacher education, and technology) (NICHD, 2000). Neither is spelling achievement sufficiently addressed in the English Language Arts Common Core State Standards, a nationwide curricular initiative representing the state of the art in educational research. While some aspects related to spelling, such as knowledge of spelling-sound correspondences, are addressed in standards listed under *Phonics and Word Recognition*,

none of the standards more than minimally address spelling as an outcome itself (NGA Center & CCSSO, 2010).

Compared to the prodigious research base in reading development, the number of studies that address spelling as a primary objective is disproportionately small. This discrepancy calls for attention from the educational community, as both reading and spelling are undoubtedly important indicators of overall literacy development.

The Myths of Spelling

In large part, the lack of interest in spelling seems to be driven by rampant "myths" or misconceptions about spelling itself. While there have been slight shifts in educational practice in recent decades in regards to spelling instruction, relatively little has changed in the general conception of spelling. In some ways, the persistence of these myths stems from age-old philosophical traditions on the relation between spoken and written language.

From one perspective, written language is viewed as a direct transcription of spoken language. The argument for the primacy of spoken over written language has mainly been attributed to Aristotle, and then to the Swiss linguist Saussure, who posited that 'the sole reason for the existence of [writing] is to represent [speech]' (Olson, 1996). This philosophical position has one critical implication for the topic of spelling, and literacy in general: If written language is merely a copy or representation of spoken language, then all activities of modern literacy, including reading, spelling, and writing, are separate and distinct from the ideas they are meant to represent. Put another way, writing is not an intellectual endeavor; instead, words are tokens of thought, in much the same way that paper bills are tokens of monetary value. Based on this view, then, the pedagogy of literacy need not be anything more than rote and mechanical.

In fact, this seems to be the prevailing philosophy of spelling pedagogy. The goal here, however, is not to debate philosophy. Rather, in highlighting the philosophical underpinnings of these myths about spelling, we emphasize that these myths are more a function of history and tradition than empirical evidence (Schlagal, 2007). For the purpose of advancing knowledge on the topic of spelling, then, it is important to delineate these myths and to discuss their implications.

Myth #1: Spelling is easy. The first myth is that spelling does not need to be taught. The premise of this myth is that spelling ability is automatically acquired, or "caught," by experience in reading and writing. Even common anecdotal evidence demonstrates that this myth is not completely accurate; if it were, there would be little observed discrepancy between reading and spelling ability, and most children receiving ordinary literacy instruction (barring those with language or cognitive impairments who would already be predisposed to general academic difficulty), would spell just as well as they read. But, as discussed above, this is not the case. For most individuals spelling is typically more difficult than reading, as plenty of apparently competent readers have difficulty spelling (Bosman & Orden, 1997).

While there is evidence that some children may learn some aspects of spelling by reading, children vary in the amount of spelling knowledge they learn by independent reading experience. Some children who appear to be competent readers and spellers may fail to learn other more sophisticated aspects of spelling knowledge, such as morphological and orthographic knowledge, which are more readily learned with explicit instruction. Other populations of children, such as struggling readers who have difficulty identifying sound-letter correspondences, may need more explicit instruction in both reading and spelling than their typically achieving peers.

For most learners, incidental learning may have a transient impact, at best, on spelling. Therefore, "while it is true that what children learn is mediated by their concepts of phonology and their knowledge of the writing system, and a few children appear to teach themselves how to spell, the majority learn faster if they are taught directly" (p. 43, Moats, 1995).

Myth #2: Spelling is too hard. The second myth is that the English spelling system is too inconsistent to learn; thus, if spelling is taught at all, it is best taught with rote methods such as memorization and repetition. Although this myth has roots in historical practices linking spelling with penmanship instruction, it is perpetuated by the popular misconception that English spelling has no teachable patterns. A more accurate statement is that English is *phonetically* inconsistent, but that other meaningful features of the spelling system, such as morphemes (i.e., the smallest unit of meaning in written language, such as affixes), are wholly discernible-- that is, if taught.

In fact, claims of the pervasive irregularity of English spelling patterns have been grossly exaggerated (Dewey, 1971). But relatively recent developments in computerized language analyses have made it possible to quantitatively refute these claims. Since the 1960s, several groups of researchers have conducted large-scale analyses of English spelling to measure the degree of consistency of sound-spelling patterns. One of the first studies of English spelling showed that 74% of phonemes in text would be spelled correctly based on an algorithm for the most frequently used phoneme-spelling patterns (Hanna, Hanna, Hodges, & Rudorf, 1966; See Fry, 2004 for a more succinct summary). In another similarly conducted analysis, Venezky (1970) estimated that 97% of English words are predictable to some extent by phonetic, orthographic, and morphological patterns, while only 3% are truly unpredictable. Although

slightly different methodologies have been utilized in each of these analyses, the consensus has been same: English spelling is far more regular than irregular.

Scrutiny of English orthography beyond the level of the phoneme is especially illuminating and provides more evidence for regularity in English at the onset-rime (i.e., in the word *cat*, the onset is *c* and the rime is *at*) and syllable levels (Kessler & Treiman, 2001; Stanback, 1992). Based on a rigorous analysis of spelling consistency across onsets and rimes, Kessler and Treiman (2001) estimated that, in English, "every sound has two spellings, with one of those spellings being used 86% of the time." Stanback's (1992) analysis of 17,602 words at the rime and syllable levels showed that rime families are the "building blocks" of words in English and account for much of the patterns of the orthography. For example, /i/ is more often spelled *ie* in the rime *–ief* (e.g., brief) and /ee/ in the rime *–eep* (e.g., sleep). An accurate assessment of English spelling consistency, then, must account for how sounds correspond to units of letters, in addition to individual letters.

There is yet another layer of sense in the array of ostensibly senseless spellings in English. One reason that some spellings may appear inconsistent is that, often, phonetic consistency is sacrificed for the sake of semantic clarity, which is often conveyed by morphemes. For example, the word *health* is spelled irregularly (i.e., not *helth)* because it derives from the word *heal*. In other words, the pronunciation changes to preserve spelling, such as in the word *sign.* Knowledge of the spelling of *sign* then serves as a useful reference for some of its derivatives, *signal, signature, signet, signify*, and allows for an efficient way to build vocabulary. In this way, English may be more aptly classified as a morphophonemic language, in that the orthography balances the requirements of both phonetic and semantic demands (Cummings, 1988). Often, meaning trumps phonetics.

Those who have failed to account for the regularities underlying English spelling have sometimes lamented what they view as unnecessarily irksome spelling patterns. For this reason they have called for language reforms to make sound-spelling patterns more transparent (Baron, 1982). Although phonetic transparency may initially seem appealing, there is a cost to a truly transparent language, as meaning is lost (Moats, 1995-96). For example, it would be much more challenging to semantically link the word *magician* to its root form, *magic*¸ if it were spelled *majishun*. As a result, learning new words would be far less efficient.

Kessler & Treiman (2003) outline three overarching principles behind the patterns in the English spelling system: conservatism (i.e., once a spelling is established as convention, the convention sticks); preservation of borrowed foreign words (i.e., *gnocchi* remains spelled as *gnocchi*, not *nokee*); and representation of non-phonemic information (i.e, morphemes such as the past tense suffix *–ed*). All of these principles help to elucidate the case for systematic, rather than unsystematic spelling contingencies in the English language.

These myths have clearly had a significant effect on the both the lay conception of the spelling process, as well as on how spelling is taught. At least in part, these myths seem to have hindered progress in the area of spelling research and instruction.

Spelling Instruction in the United States

It appears that there are few if any thoroughly established principles upon which educators can confidently rely in organizing and conducting the teaching of spelling. (Watson, 1935, p. 5).

Currently, there is a good deal of confusion among teachers about the English spelling system and methods for teaching it. (Schlagal, 2007, p. 180).

As the above statements suggest, nearly a century has passed with little advancement in knowledge about spelling development and instruction. In short, spelling has simply not been an

educational priority. Teachers now seem to be just as confused about what constitutes good spelling instruction as they were in 1935. Considering the breadth of research that has been conducted on reading development and instruction, the lack of advancement in the area of spelling is somewhat startling.

Since the publication of Noah Webster's "Blue-backed Speller," in 1783, spelling instruction in the United States has undergone several currents of change (Schlagal, 2002, 2007; Venezky, 1980). During the 19th century, Webster's text served as the backbone of instruction in not only spelling, but also in reading, grammar, pronunciation, penmanship, and moral improvement (Schlagal, 2002). Whereas the earliest spelling lists included long lists of esoteric words with no apparent organizational scheme, the 1930s marked the beginning of spelling lists organized by word frequency. This development was followed by a novel attention to orthography and word patterns in the 1960s, which was partly influenced by the impact of the burgeoning computer age. In spite of these slight advancements in spelling pedagogy, these and other subsequent improvements in spelling instruction seem to have had only a slight impact on subsequent educational practice.

While there have been some modifications in the spelling curriculum in the United States since the early 19th and 20th centuries, much of spelling instruction today bears a concerning resemblance to those early instructional models. Spelling instruction often primarily involves lists of words to be memorized, with little to no emphasis on phonetic or orthographic patterns in words (Johnston, 2001). Instead, these lists are often organized thematically around school subjects and are taught as "vocabulary" instruction. In other cases, spelling is addressed in word lists and other materials associated with basal reading programs; however, in these programs spelling may only be incidentally addressed (Morris, Blanton, Blanton, & Perney, 1995).

Indeed, recent surveys of nationally representative samples of elementary school teachers have shown that spelling is not a primary concern in the language arts curriculum in many classrooms (Fresch, 2003; Graham, 2007; Johnston, 2001). One sample of grade school teachers reported teaching spelling for an average of 90 minutes per week, though there was considerable variability (SD = 70 minutes) in weekly instruction time (Graham, 2007). The same survey indicated that very few teachers reported making any instructional adjustments for struggling students, even though up to half of the students had difficulty spelling.

Perhaps the lack of attention given to spelling in part stems from teachers' lack of preparation themselves in this area. Teachers often lack the requisite linguistic knowledge to effectively teach students, especially struggling students who require more in-depth instruction (Moats & Lyon, 1996; Moats, 2009). They also report a lack of knowledge in instructional issues such as how to appropriately build word lists, and how to assess individual students' spelling abilities (Fresch, 2003; Templeton & Morris, 1999). This lack of training may partially explain why teachers continue to use outdated methods of spelling instruction. In many classrooms, students receive a list of spelling words on Monday, take a spelling test on Friday, and then forget the words shortly afterward (Gill & Sharer, 1996).

The premise of these weekly spelling tests is essentially no different from that of whole-word approaches to spelling, such as the "cover, copy, compare" (CCC) approach, which, as its name implies, involves having a student read a word, hold the word in memory while covering the word, copy the spelling of the word, and then correct any errors. In effect, CCC and other whole-word methods of spelling instruction are characterized primarily by word exposure and rote practice in transcribing words, rather than by explicit instruction in linguistic knowledge.

But studies comparing the effects of whole word methods to a variety of control conditions on spelling achievement have failed to provide compelling evidence for the benefit of these methods over other types of explicit spelling or reading instruction. Rather, these studies merely suggest that providing opportunities to spell words may sometimes lead to better immediate spelling performance than providing fewer opportunities to do so (Gettinger, 1993; Nies & Belfiore, 2006; Omrod & Spivey, 1990; Schermerhorn & McLaughlin, 1997), though in some cases there may no advantage at all (Hilte, Bos, & Reitsma, 2005). Other whole-word methods utilizing interactive computer technology (ICT) may involve similar instruction to CCC with the added benefit of computer software programs that create individualized lists of words for students, based on their spelling errors (McClurg & Kasakow, 1989). However, a systematic review of seven randomized-control-trial studies of ICT interventions showed that ICT interventions were no better than control conditions in benefiting spelling outcomes (Torgerson & Elbourne, 2002).

Despite the prevalence of whole-word methods in schools, there is little evidence to support the view that whole-word methods do anything more than yield training effects in spelling performance due to repetition in spelling words. Although it is expected that exposure to and practice transcribing spellings of words would have at least an immediate effect on spelling performance, the likelihood of these methods having a long-term impact on spelling knowledge is dubious (Cooke, 1997). Likewise, it is questionable whether these methods benefit students' overall literacy development.

As history has shown, spelling instruction in the United States has generally not been an area of educational innovation. Although contemporary educational practice as a whole has

become considerably more diverse and sophisticated over a century and more of progress, strangely the practice of spelling instructional has not followed suit.

Rationale for the Present Study

Despite its importance in the context of literacy development, spelling has been a relatively neglected topic in both research and practice. Compared to research in reading, little is known about what kinds of instructional methods improve spelling ability, as spelling is not often a focal point in empirical studies. Moreover, pervasive myths about spelling have reinforced the view that spelling is merely an act of memorization, not a representation of linguistic knowledge. The purpose of the present study is to clarify these issues, fill the knowledge gap on what helps spelling, and to provide a stepping stone for the advancement of further research in spelling development and pedagogy. To this end, meta-analysis is used as a tool for synthesizing the existing literature and quantitatively assessing the impact of various school-based literacy interventions on spelling outcomes across a wide population of students.

CHAPTER 2: Literature Review

Reading and Spelling as Knowledge of the Alphabetic Code

Mastery of written language, unlike spoken language, requires an extensive amount of explicit instruction in the respective symbol system. Specifically, the process of learning to read and spell in an alphabetic language (i.e., languages that use symbols, rather than pictures, to represent sound) is founded on knowledge of the alphabetic code, or how the sounds in the language correspond to letters or groups of letters. Knowledge of the alphabetic code "entails an awareness of the internal phonological structure of the words of the language…, which must be more explicit than is ever demanded in the ordinary course of listening and responding to speech" (Liberman, Shankweiler, & Liberman, 1989, p. 5).

Although some researchers may define the alphabetic code (or alphabetic knowledge) as knowledge of sound-letter correspondences (Ehri, 1992; Liberman et al., 1989; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001), in the present study this definition has been broadened to include other types of lexical knowledge, such as orthographic and morphological knowledge. The "code" may represent an individual letter representing a sound, groups of letters representing commonly occurring patterns characteristic of the orthography (e.g., digraphs, such as *ng*, *ea*), or letters that together convey meaning (e.g., morphemes, such as the prefix *pre-*) (Berninger & Fayol, 2008). Respectively, phonological, orthographic, and morphological knowledge each represents a distinct aspect of the alphabetic code, and each plays a role in the process of reading and spelling words.

Fundamental to knowledge of the alphabetic code is phonological awareness, or the awareness of how phonemes (i.e, the smallest unit of spoken sound) map onto graphemes (i.e., the smallest written unit), and vice versa. Literacy research from the past three decades has

established a strong case for the correlation between phonological awareness and reading, particularly in the early years of formal literacy instruction (Blachman, Ball, Black, & Tangel, 1994; Bradley & Bryant, 1983; Bus & van IJendoorn, 1999; Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, & Shanahan, 2001; Ehri, Nunes, Stahl, & Willows, 2001; Juel, 1988; Lundberg, Frost, & Petersen, 1988; Melby-Lervåg, Lyster, & Hulme, 2012; Stanovich, Cunningham, & Cramer, 1984; Wagner & Torgesen, 1987). Although not as widely addressed, the relationship between phonological awareness and spelling seems to be of a similar magnitude (Arra & Aaron, 2001; Caravolas, Hulme, & Snowling, 2001; Foorman, Francis, Novy, & Liberman, 1991; Morris & Perney, 1984).

This evidence adds to the case that the development of spelling ability is intertwined with the development of reading ability in children (Cataldo & Ellis, 1988; Ehri, 1997). In this vein, studies have shown that explicit instruction in spelling can also improve reading outcomes. For example, first graders who received explicit instruction in phoneme segmentation and spelling showed gains in word-level reading (Uhry & Shepherd, 1993). In another study, researchers found that spelling ability at the beginning of first grade predicted reading achievement at the end of the year (Morris, Blanton, Blanton, & Perney, 1995). Spelling has also been found to explain 40% to 60% of the variance in oral reading ability in third- and fifth-graders (Zutell & Rasinski, 1989). While reading and spelling rely on knowledge of the alphabetic code, it is important to acknowledge that they are still distinct processes, and each skill may manifest itself in different individuals in a variety of ways.

Findings from a longitudinal study of 1,342 students in Grades 1 to 4 provide further evidence that reading and spelling stem from a common skill set (Mehta, Foorman, Branum-Martin & Taylor, 2005). In this study, a confirmatory factor analysis supported a “unitary

literacy factor" for reading and spelling, such that word reading, passage comprehension, and spelling were explained by this common literacy factor. Literacy was narrowly defined as "reading and writing," as opposed to a broader definition of literacy that includes non-pedagogical activities. These researchers also accounted for the multi-level nature of classroom research and confirmed the one-factor model for reading and spelling at both the individual student and classroom levels.

The implication of this study, as argued by Mehta et al., is that literacy is a multicomponent construct with multiple indicators. For instance, both reading and spelling rely on a foundation of phonological knowledge. Other aspects of linguistic knowledge, such as morphological knowledge, may also be linked with multiple indicators of word reading, spelling, comprehension, and other aspects of literacy (Bowers, Kirby, & Deacon, 2010; Carlisle, 2010). Thus, spelling ability may be considered one facet of the multi-faceted construct of literacy.

The Role of Phonological Awareness in Spelling

To understand the importance of phonological awareness in spelling, it is helpful to first consider its relationship to decoding, an early reading strategy that involves segmenting words into phonemes. When a child decodes the word *cat* she must visually perceive the letters C, A, and T and derive the appropriate phonemes, which she might say aloud as /k/, /æ/, and /t/ (with guidance, if she is a beginner). Thereafter the child is also likely to write these letters as C-A-T, and read her spelling of the word as /kæt/. All of these activities blur the line between spelling and reading (Ehri, 1997), largely because at this point in development both spelling and reading are reciprocal processes based on the child's attempts to map sounds onto letters, and vice versa.

In fact, there is a growing consensus that knowledge of letter-sound correspondences is a prerequisite for spelling development, and that other forms of linguistic knowledge build on this

foundation (Ehri, 1992; Roberts & Meiring, 2006). Given the complexity and apparent irregularities of the English language, it seems reasonable to posit that a strong foundation in basic sound-letter correspondences is required in order for children to learn other more complex aspects of the language. As such, beginning readers and spellers spend much of their time learning about and reinforcing their knowledge of sounds in words, typically by doing activities involving phoneme manipulation, either with tiles or letters (Liberman, Shankweiler, Fischer, & Carter, 1974; Uhry & Shepherd, 1993). Pre-literate children may play rhyming games to first learn how to discern sounds in words. In effect, spelling begins with the process of encoding sounds as letters and vice versa.

Research suggests that once a word has been decoded a few times, children will come to recognize it as a whole on future occasions (Ehri, 1997; Nunes, 2009). Nevertheless, even if the representation of a word's spelling supports whole-word recognition when the word is read in text, the representation may be insufficient to allow the child to retrieve the word's spelling in the context of writing. Thus, while phonological knowledge serves as a foundation for both spelling and decoding skills, spelling requires a more robust representation of words than decoding, and this representation is often bolstered by phonological knowledge.

On this note, the case of the good reader/poor speller deserves attention. Whereas some individuals may appear to be competent readers in spite of their poor spelling ability, research suggests that poor spelling may reflect underlying impairments in reading ability or linguistic knowledge (Kamhi & Hinton, 2000). For some, these underlying impairments may be subtle deficiencies in orthographic knowledge, in spite of relatively good phonological knowledge. However, for some poor spellers, these impairments may indicate deficiencies in phonological knowledge. For instance, in a study of 12-year olds, Frith (1980) showed that good readers/poor

spellers made mostly phonetically plausible errors, like good spellers, but that they were poorer at reading aloud, which requires more attention to phonetics. This suggests that some good readers/poor spellers may have weaker phonological knowledge than good readers who are also good spellers.

Beyond Letter-Sound Correspondences

In light of the many phonetic inconsistencies in written English, phonological knowledge is necessary but not sufficient to spell a multitude of words. A growing number of studies have shown that at a certain point in development, knowledge of orthographic patterns (i.e., language-specific common and uncommon letter sequences), morphology (see Bowers et al., 2010 for a review), and etymology (word origin) uniquely contribute to the development of spelling skills (Nunes & Bryant, 2009). It seems reasonable to assume, then, that just as children are taught phonological skills, they can also be taught other skills and strategies to facilitate the development of spelling ability.

For example, the spelling of the word *magician* can be demystified with knowledge of the agentive suffix *-ian* (i.e., in this case, an agent of magic) and the word stem *magic*. Research suggests that morphological instruction benefits both elementary and middle school students, as well as students who are less capable spellers (Bowers et al., 2010). Likewise, having knowledge of the word *solemnity*, which includes a vocalization of /n/, provides the learner with a guide for how to spell *solemn*, which, with its silent *n*, may otherwise confuse the learner (Templeton & Morris, 1999).

In addition to morphology, children also need to learn orthographic patterns, such as the distributions of the “silent e” and doubled consonants, and how these orthographic patterns provide clues to the pronunciations of vowels (e.g., *file* vs. *fill*). These kinds of orthographic

patterns can be troublesome for students who are typically taught to identify the pattern (i.e., that only certain letters double), but are not taught the harder task, which is to recognize that double letters change the pronunciation of the preceding vowel (Nunes & Bryant, 2009). In addition, children will inevitably encounter difficulties when learning irregular words like *ache* and *aisle*, or words borrowed from foreign languages such as *karaoke* and *tsunami.* In such cases, etymological and word-specific knowledge are required.

In lights of this complexity, it is not surprising that it takes children an extended period of time to become proficient spellers, and that many may fail to grasp morphological, orthographic, and etymological cues to spelling. Although many children become familiar with some of these patterns through their experiences with text, systematic instruction on a broader range of these patterns would likely improve spelling ability in most students, particularly in older students beyond grade four. Systematic instruction in spelling could also likely provide scaffolding for those who are likely to have difficulties with reading and spelling due to an underlying language impairment.

Even so, there exist certain vagaries of the English language which make particular aspects of spelling amenable to rote methods of instruction. The argument here is not that spelling involves *no* rote learning, or that there is anything inherently wrong about rote learning. For example, no one would argue that developing some degree of automaticity in arithmetic, such as memorizing multiplication tables, is a prerequisite for learning algebra. Rather, it is possible that learning to spell, like learning to do mathematics, is characterized by an iterative process of both conceptual understanding and procedural skill (Rittle-Johnson, Siegler, & Alibali, 2001). Students may benefit from explicit instruction in code-based knowledge, as well

as increased exposure to print, though either guided or independent reading experiences; both facets of learning should be addressed in instruction.

Furthermore, it is likely that even relatively good readers and spellers can benefit from a deeper understanding of language afforded by more systematic spelling instruction. The argument here is not that all children should become linguistic experts. Rather, to the extent that word knowledge can help make connections between words more clear, this type of instruction should be incorporated into the curriculum.

Theoretical Models of Spelling

Any consideration of the impact of literacy interventions on spelling achievement requires consideration of the nature of spelling itself. Issues of how researchers have conceptualized the process of spelling are addressed here.

Dual-route models. Although dual-route models have a long tradition in reading research and computational theories (Coltheart, 2005, 2006), these models have also been applied to spelling (Brown & Loosmore, 1994; Houghton & Zorzi, 2003). The premise of the dual-route model is that reading and spelling are achieved by way of two separate pathways, the lexical and non-lexical routes. The lexical route is the primary route through which a reader discerns irregularly spelled words that do not follow sound-correspondence rules, such as *yacht*, whereas the non-lexical route allows the reader to read an unfamiliar word by making use of sound-letter correspondences.

Cases of brain damage are often interpreted as support for this theory. As support for dual-route models of spelling, Barry (1994) cites Shallice's (1981) case study of a neurological patient who exhibited signs of phonological dysgraphia. This individual was able to successfully spell real words, but at the same time was starkly deficient in his ability to spell non-words.

Evidence from such neurological cases and other non-neurological cases in which individuals exhibit a divergence in phonological and non-phonological competencies support the notion that there are multiple distinct pathways involved in reading and spelling. The strength of the dual-route model is that it accounts for these distinct pathways.

Although the dual-route model can accommodate the existence of two distinct pathways in spelling words, it is less useful in explaining how these two pathways interact with each other simultaneously. Spelling a word is a complex process and requires the use of multiple cognitive processes. For instance, Ehri (1992) explains how retrieving a spelling for a difficult word may involve knowledge of both phonology and orthography, in that knowledge of plausible sound-letter correspondences for the spelling of a word can constrain the available choices of orthographic patterns in the speller's lexical repertoire. In cases like this, it may not be possible to determine whether one cognitive task is mutually exclusive from the other.

Connectionist models. In some ways connectionist theories serve as a response to the inadequacies of dual-route theory, though connectionist theories are also related to computational models of learning. While there is certainly evidence that more than one pathway may be involved in the process of spelling, these pathways are not necessarily independent from each other. With the exception of a few cases, there is evidence that multiple pathways are at work in the process of spelling. Therefore, unlike dual-route theories, which emphasize the distinct and separable pathways related to spelling, connectionist theories highlight the idea that spelling pathways may often be integrated with each other.

The most recent developments in connectionist theories have emphasized the role of statistical learning, which has been observed in children as young as 8-months old (Saffran, Aslin, & Newport, 1996). That is, "learning is statistical in that it goes beyond all-or-none

patterns to encompass probabilistic patterns" (Treiman & Kessler, 2006, p. 643), and learning is characterized by the strength of connections between stimuli. Connectionist theories draw attention to the role of experience, such as a speller's exposure to print, as well as highlight the role of the interaction among multiple variables at a fine grain.

In spelling research, statistical learning is evidenced in studies that have shown that children learn about sounds in words through exposure to word spellings, without explicit instruction (Thompson, Fletcher, Flinn, & Cottrell, 1999). Other studies have shown that children's spellings tend to over-represent letters from their own names (Treiman, Kessler, & Bourassa, 2001), and that Brazilian-speaking children tend to over-represent vowels in their spellings compared to English-speaking children, due to the fact that Portuguese words have about twice as many vowels as English words (Pollo, Kessler, & Treiman, 2005). Children also tend to have more difficulty spelling vowels than consonants, as the sounds of vowels are highly contingent on the surrounding letters in the word (Treiman, 1993); the greater number of encoding possibilities may make it difficult for children to learn this pattern (implicitly or otherwise).

Connectionist views on spelling have drawn attention to the notion that, through statistical learning, children may develop an early awareness of more linguistic aspects of words deemed advanced, such as orthographic and even morphological knowledge. In this vein, some researchers argue that children may develop orthographic knowledge from the very beginning of literacy acquisition (Martinet, Valdois, & Fayol, 2004). Such accounts are based on the notion that although young children are typically learning how to navigate sound-letter correspondences, they have some amount of awareness about other sources of word knowledge beyond the phoneme level. Training studies may be useful in substantiating this view.

Developmental models. If there has been any recent shift in developmental conceptions of spelling, it can arguably be attributed to Read's (1971), Chomsky's (1979) and Bissex's (1980) seminal studies on pre-literate children's invented spellings. (This is not to be confused with the same term used by advocates of a particular kind of spelling instruction based on the idea that children should be encouraged to "invent" their own spellings of words in the process of learning conventional spellings.) The most important finding from this group of studies was that pre-literate children did not come to spell words with a blank slate, filled in only by formal literacy instruction. Rather, children in these studies seemed to produce spellings of words in a more systematic way than had been previously observed or expected.

Contrary to the belief that pre-literate children largely produced random strings of letters and scribbles, a close inspection of the children's writing provided evidence that spellings were not completely random. Instead, patterns in these children's spellings suggested that these children had applied their knowledge of oral speech sounds in their attempts at mapping letters onto sounds. For example, to represent the sound /s/, the children in Read's (1971) studies seemed able to recognize that /z/ was more closely related to /s/ than to /t/. Thus, these studies suggested that even children who had not begun to read words were able to apply phonological knowledge to spell words. What is perhaps the most salient implication of these studies is that spelling, like reading, is a developmental rather than all-or-none production of knowledge. Further, these studies offered evidence that spelling development may at least partially and at most significantly depend on code-based as opposed to meaning-based knowledge.

Subsequent to the publication of Read's and Chomsky's work, several groups of researchers developed various models of spelling development (Bear, Invernizzi, Templeton, & Johnston, 2000; Gentry, 1978, 1982; Henderson, 1981; Henderson & Beers, 1980). Much of the

underlying patterns in these models are similar, as all of the models describe children as progressing from non-phonetic spelling, to phonetic spelling, and then eventually to a consolidated phase in which children are able to spell phonetically but also incorporate orthographic and morphological patterns in words. Table 1 describes some of these stages and aligns them with the use of broad developmental age categories used in the present study.

Although these models of spelling development do much to elucidate common patterns of how children come to spell words, the exact mechanisms of how children make use of different types of word knowledge is not entirely clear from the works of these researchers (Pollo, Treiman, & Kessler, 2008). For instance, it is not clear whether children at different stages of spelling development are qualitatively different from children in another stage, or whether they are simply making use of different strategies. A related issue is how children move from one stage to another.

Some researchers have criticized developmental models of spelling, arguing that these models are too restrictive and pay short shrift to other types of word knowledge other than phonological knowledge (Treiman & Kessler, 2006). With respect to her related word recognition model, Ehri (1998) has acknowledged that alphabetic knowledge is the "key capability that distinguishes among the levels and underlies development" (p. 253). However, Ehri (1997) has also emphasized the use of the term "phase" as opposed to "stage" to convey that the activities and competencies of each phase are not completely independent from those of other phases. Rather, each phase is more aptly described as a characterization of the *dominant* competencies that tend to occur in that age group, rather than a universal prescription for how spelling develops in all individuals.

Table 1

Developmental Phases of Word Recognition and Spelling.

<table>
<tr><th>Type of Reader</th><th>Ehri (1997)</th><th>Gentry (1982)</th><th>Bear et al. (2000)</th></tr>
<tr><td>Pre-readers</td><td>Pre-alphabetic</td><td>Precommunicative</td><td>Emergent</td></tr>
<tr><td rowspan="3">Developing Readers</td><td rowspan="2">Partial-alphabetic</td><td rowspan="2">Semi-phonetic</td><td>Within-word pattern</td></tr>
<tr><td>Letter-name</td></tr>
<tr><td rowspan="2">Fully alphabetic</td><td>Phonetic</td><td rowspan="2">Syllables and affixes</td></tr>
<tr><td rowspan="2">Proficient Readers</td><td>Transitional</td></tr>
<tr><td>Consolidated alphabetic</td><td>Correct</td><td>Derviational</td></tr>
</table>

An integration of models. As generally seems to be the case in theoretical debate in research, researchers tend to "throw the baby out with the bathwater" by opposing one aspect of a particular theory at the expense of negating the entire theoretical model. However, the issue is not whether a theory in its entirety is right or wrong, but to what extent a theory helps to explain the phenomenon in question, based on empirical evidence.

In spite of their limitations, each theoretical model offers valuable insight into the process of spelling. Dual-route theory highlights the fact that the process of spelling involves more than one pathway or strategy, and helps to explain why, for instance, adults may be able to read words

that they cannot spell. Connectionist theories emphasize the ways in which multiple kinds of lexical information are used to spell, and that the way in which this information is used varies depending on the interaction between the speller and the characteristics of the language. Finally, developmental theories help to address the affordances and constraints associated with different aspects of linguistic knowledge, as well as how these affordances and constraints intersect with the developmental capacities of the speller.

Relationship Between the Present Study and Prior Research

For the past two decades, the "great debate" over skills-oriented phonics instruction versus meaning-oriented whole language instruction has taken center stage in the educational arena (Chall, 1996). A large body of research over the past several decades has provided robust evidence for phonological awareness as the cornerstone of literacy development. Based on this view, phonics instruction has been prevalent in literacy curricula in schools throughout the country. On the other end of the educational spectrum is the whole-language instructional approach, whose advocates argue that children best develop reading and spelling skills by self-motivated experiences with literature and with little explicit instruction. That is, the premise of the whole-language approach is that exposure to sufficient text will increase children's vocabulary, decoding skills, and spelling ability without explicit instruction about letter-sound correspondences.

An accumulation of evidence over this same time period has essentially made the case for the integral role systematic phonics instruction plays in early reading and spelling development. Moreover, pure whole language approaches with little explicit instruction in linguistic skills has been shown to be less effective in benefiting literacy outcomes. For example, the recent meta-analysis conducted by the National Reading Panel (Ehri, Nunes, Stahl, & Willows, 2001) on

early reading instruction showed that systematic phonics instruction was superior to whole language approaches to reading. Another research synthesis showed that whole language approaches may be less effective for disadvantaged or learning disabled populations, for whom explicit instruction is crucial (Stahl & Miller, 1989).

In fact, research suggests that neither the whole language nor phonics approach on its own may be sufficient to improve students' spelling skills. Rather, increasing evidence supports the view that students use multiple skills and strategies to spell at different developmental time points, and that explicit instruction plays an important role in this process (Apel, Masterson, & Niessen, 2004).

In order to determine what kinds of instructional approaches are most effective in teaching students to spell properly, intervention studies are needed. While there are relatively few intervention studies specifically targeted to improve spelling in the broader population, several research syntheses of spelling interventions for learning-disabled and dyslexic populations have shown that explicit instruction improves spelling (Fulk & Stormont-Spurgin, 1995; Gordon, Vaughn, & Schumm, 1993; Wanzek, Vaughn, Wexler, Swanson, Edmonds, & Kim, 2006). However, aside from focusing only on learning-disabled or dyslexic students, these studies primarily addressed aspects of the instructional environment that would support spelling interventions (e.g., the availability of classroom assistants), not substantive characteristics of the literacy intervention itself, such as whether instruction included morphology. Furthermore, none of these studies utilize statistical methods to estimate or compare effect sizes (ESs).

Weiser and Mathes (2012) recently reviewed the effects of 11 encoding interventions on children in kindergarten to Grade 3. Although they did not employ meta-analytic methods to estimate an overall ES, they reported a range of moderate to large ESs from each study. Overall,

they reported that interventions involving encoding, defined by activities such as manipulating phonemes or spelling words, were superior to traditional instruction in improving spelling outcomes for both typical and struggling students.

Although there are relatively few studies that specifically address spelling interventions, there is an abundance of studies of multiple other literacy interventions that include spelling as an outcome. The large majority of these intervention studies, many of which target phonological awareness, have been designed to improve reading skills and only minimally address spelling. To date, no quantitative meta-analysis examining the impact of literacy interventions on spelling ability for all populations has been conducted. The present study builds on the work of previous reviews, but also expands the scope and range of studies included in the analysis.

Research goals. Given the state of the field, there are three major goals of this study: 1) to provide a comprehensive synthesis of current literature on the impact of literacy interventions on spelling ability, 2) to investigate the differential effects of interventions based on participant (e.g., age), intervention (e.g., orthographic word study), and study characteristics (e.g., use of random assignment); 3) to shed light on the literature gap in spelling development and its relationship to other literacy outcomes.

CHAPTER 3: Method

Overview of Meta-Analysis

Meta-analysis is a systematic review of research that involves a quantitative analysis of the relationship between two variables across a selection of studies. Since the first meta-analysis on psychotherapy outcome studies (Smith & Glass, 1977), this method has become increasingly popular in recent decades in many different fields, such as medicine, psychology, and education (Borenstein, 2009). This increase is arguably for good reason, as meta-analyses help to build knowledge in a field by identifying ideas for productive research, by clarifying the current state of knowledge, by increasing the precision of what is known, and by broadening the scope of the current paradigm (Chan & Arvey, 2012). In considering this last characteristic, the use of meta-analysis in this study is especially warranted because of the need to draw more attention in the field to literacy skills beyond phonological knowledge, such as morphological and orthographic knowledge.

Meta-analysis differs in important ways from narrative review, which is typically employed in reports of primary studies. Both methods involve synthesizing a body of research and deriving a summary of results, but the similarities end there. First, unlike narrative review, meta-analysis involves searching for, selecting, and coding studies based on clearly reported objective criteria. Transparency in the methods used to collect and code studies precludes the synthesis from being overly biased, and also allows for the possibility of replication. Details regarding the literature search, study inclusion and exclusion criteria are described in detail in the next section.

Second, meta-analysis allows for a quantitative analysis of ESs, a standardized value for the relationship between two variables (in this case, the magnitude of the intervention effect on

spelling). In computing ESs, larger and thereby more precise studies are weighted more than smaller studies. Thus, the use of multiple study samples affords a more precise estimation of the true effect, as well as higher statistical power to detect small to moderate ESs. Because of its focus on ES estimation rather than arbitrary significance tests, meta-analysis can be especially useful in clarifying seemingly conflicting results in the literature, which often stem from overly crude vote-counting procedures based on whether a certain level of significance was reached (Hunter & Schmidt, 2006).

A third and particularly useful feature of meta-analysis is that ESs heterogeneity among studies can be modeled and analyzed based on relevant moderator variables. Quantitative moderator analysis affords a detailed investigation of issues that would not be possible in either a primary study or a qualitative research synthesis. The extent to which a categorical variable moderates ESs is captured by the *Q* statistic, which is defined as the total amount of variance between studies. The I^2 statistic indicates the proportion of this variance that is attributable to true differences among studies, as opposed to error.

Study Retrieval and Selection

The literature search in this case departs from the typical search in that it serves as data collection and is intended to be as comprehensive as possible. While there is no way to ensure the exhaustiveness of the search, multiple research channels were accessed in an effort to decrease the likelihood of introducing systematic bias to the study retrieval process. Three sources were used to identify viable studies: electronic bibliographic databases, prior research syntheses, and references from retrieved studies. Electronic databases such as ERIC, PsycINFO, Social Sciences Citation Index, and Dissertation Abstracts were searched by abstracts. The search was limited to peer-reviewed articles published in English.

We conducted an initial search of electronic databases with the search terms *spell** and *(train** or *intervention)* to assess whether meta-analysis would be viable. This yielded over 800 results, from which we retained an initial sample of 40 studies. Next, we conducted a more comprehensive search of the literature. Keywords corresponding to the outcome (e.g., *spell*, encode*), intervention (e.g., *intervention, training, phonological*), and study design (e.g., *comparison, group, control, random assignment)* were crossed and entered with wild card operators to locate eligible studies. This search yielded over 1600 results, all of which were screened by order of relevance. Third, prior meta-analysis and reviews were searched for eligible studies. Finally, a snowball method of locating studies was used through the reference section of each retrieved study.

Given the relative dearth of research on spelling, this meta-analysis was designed to cast a wide net in order to comprehensively represent the available literature. Studies were selected in two screening stages. In the first stage, abstracts were screened in a relatively inclusive manner in order to avoid missing potentially eligible studies. At this stage, if the study appeared to report a literacy-related intervention as well as literacy outcomes, the full text was retrieved.

In the second stage, full texts were searched in order to determine whether spelling was reported as an outcome measure, and whether the reported statistics and study design allowed computation of ES corresponding to group contrasts. Spelling was operationally defined as generative (i.e., students spelled dictated words rather than chose correctly spelled words from a list) and writing-based (i.e., students spelled words by writing or typing them on a computer, not simply orally reciting the words). Further, any spelling measures that included primarily practiced words were excluded to preserve generalizability of results. Based on these eligibility

criteria, over 350 full text studies were retrieved and searched for potential inclusion in the meta-analysis.

Out of these retrieved studies, over half were excluded because of invalid design (i.e., correlational or case study design) or because they did not include a valid spelling outcome. Any study that reported insufficient statistics, or for which it was not possible to access the pertinent statistics, was excluded from the analysis. Of the remaining 190 studies, studies were excluded for the following reasons: no viable comparison group (45); a major design flaw, such as reported low fidelity of treatment (12), the intervention was not discernible (12), not a literacy intervention (6), duplicate samples (4), not in a school setting (2), and non-generalizable sample, such as children with cerebral palsy (1).

All research designs were included except for case studies, which are not viable for statistical analysis. Single-group designs were also excluded, as ESs yielded from pre-post contrasts are not comparable to those yielded from group contrasts (Lipsey & Wilson, 1993). While some researchers opt to exclude less methodologically rigorous studies from the meta-analysis based on a priori criteria (Ehri, Torgerson, Brooks, & Hall, 2006), others argue for the benefits of using post-hoc criteria to empirically assess any threats to validity (Smith & Glass, 1977; Cooper & Cooper, 2010), or a combination of the two (Valentine, 2009). Cooper and Cooper further argue that the issue of whether study methodology significantly affects the ES is best treated as an empirical question. With the exception of the a priori exclusion of case study methodology, the post-hoc strategy of assessing the impact of study quality on ESs was used.

The screening process yielded 111 total studies that met all eligibility criteria. However, seventeen of these studies were excluded from most analyses because they only reported spelling

measures at follow-up time points. The remaining 91 studies included post-test spelling measures and comprised the full sample of studies used in the majority of the analyses.

Studies with multiple conditions. All interventions involved explicit, systematic instruction emphasizing at least one of the following skills, ordered from smallest to largest lexical unit: letter-sound correspondences (phonemic), orthographic and morphological units, or a mix of the two. Explicit instruction entailed that content was clearly taught, as the instructor pointed out the relevant content throughout the lesson. Systematic instruction entailed that the content was taught on a regular basis, rather than incidentally throughout the intervention.

Although most studies included one clear intervention group and one control group, nine studies included multiple groups. Comparisons from studies including more than one intervention group were selected based on the following criteria: 1) Treatments had to involve systematic and explicit instruction in phonological, orthographic, and/or morphological knowledge. Any treatment that met this criterion was potentially included in the analysis. 2) Included interventions from a single study had to be distinguishable from each other according to the intervention descriptions in the coding manual. If the study included slightly different iterations of the same interventions, the intervention that was most similar to the others in the same category was chosen; otherwise, one was randomly chosen. The same applied if two interventions involved the same instructional content (i.e., and were both coded as the same intervention category) but differed in intensity (e.g., duration, number of letter-sound correspondences taught, number of spelling rules taught, etc.) or another detail (e.g., phoneme segmentation with colored blocks versus with tiles).

Coding Procedure

All relevant data were carefully extracted from eligible studies and entered into a *Filemaker Pro v.12* database, which was also used to manage study retrieval and bibliographic data. Eligible studies were coded for reported statistics, sample size, participant characteristics, intervention type, type of spelling measure, as well as extrinsic study characteristics such as journal tier. A detailed coding manual was developed in order to systematically and reliably code the relevant variables from each study; this manual was edited as needed throughout the coding process. To establish inter-coder reliability, high-inference codes for 20% of the study sample (n = 22) were independently coded by a trained doctoral student who had previous experience conducting meta-analyses. Reliability based on the Cohen's kappa statistic ranged from .80 to .90 on all moderator variables. Any coding disagreements were resolved through discussion and by consulting the original study.

Computation of Effect Sizes

The ES statistic used in this meta-analysis is Cohen's *d*, which is the standardized mean difference between groups, $ES_{SM} = X_T - X_C / s$, which is the difference between the post-test means of the treatment and control groups divided by the pooled post-test standard deviation. A positive ES in the present study represents the extent to which the intervention group exceeded the comparison group, in units of the post-test outcome measure. For 77 of the 92 included studies, means and standard deviations were used to compute ESs. Eleven studies reported computed ESs, and 6 studies reported *t* statistics for independent samples. All computation of ESs and analyses were conducted using *Comprehensive Meta-Analysis Version 2* (Bornstein, Hedges, Higgins, & Rothstein, 2005), a statistical program specifically designed to conduct meta-analysis.

In choosing a statistical model for the analysis, we considered different available meta-analytic models, which differ primarily in their assumptions about sources of error. The random-effects model is preferable in most cases (Bornstein, 2009) and is appropriate for this study. In this model, the overall error attributed to estimating the summary ES is constituted by two sources: the sampling error in an individual study, as well as the sampling error of the group of studies, which is considered to be sampled from a distribution of studies centered on a true effect mean. The random-effects model accounts for the assumption that the true effect of the intervention may vary across studies, given that studies are likely to be different from each other.

Moderators

Twenty variables were coded from all studies and used as categorical moderators in the analyses. These moderators were selected based on theory and prior meta-analyses relating to literacy (Bus, 1999; Ehri et al., 2001). Four types of moderators associated with the dependent variable, intervention characteristics, participant characteristics, methodological variables, and extrinsic variables were coded. (See Appendix A for the full Coding Manual.) Studies which failed to report sufficient information on certain moderator variables were excluded from the corresponding moderator analysis.

Dependent variable. While the construct of interest in this meta-analysis is spelling ability, this construct was operationalized and measured differently among studies. Moderators associated with the dependent variable included the time of measurement, the type of spelling outcome measure used, and the scoring method used in the outcome measure. The purpose of including these moderators was to assess whether different methods of measuring the construct of spelling impacted the magnitude and distribution of ESs.

Time of measurement. ESs from three time points of measurement were coded. ESs measured within one month after the end of an intervention were coded as post-test. This category also included multi-year interventions which reported results at the end of the first year. Short-term follow-ups included ESs reported within one year after the end of an intervention. Long-term follow-ups included ESs reported more than one year after the end of intervention.

Type of measure. Spelling outcome measures represented in the study sample included standardized and researcher-constructed measures. For most of the measures, students were asked to write dictated words, which were usually stated in isolation and in a sentence. The type of word used in the outcome measure was also coded (i.e., real words versus pseudowords), but almost all words were real words. ESs from both types of measures were included to assess whether one type of measure yielded greater ESs than the other.

Scoring method. Spelling outcome measures varied in whether they were scored by the correct conventional spellings of words or whether they were scored by developmental criteria, which afforded partial credit for words based on accurate phonetic representations of letters or groups of letters. For instance, as a spelling for the word *pencil*, *pensil* would be allotted more points than *penkil*.

Intervention. One of the major goals of this meta-analysis was to explore the differential effects of various interventions on spelling. As such, the type of intervention, as well as instructional parameters such as the length of the intervention, the instructor who taught the intervention, and the size of the instructional unit were the primary substantive moderators of interest.

Intervention type. Six types of interventions were coded based on instructional content. Both phonemic and phonological awareness are defined by the child's ability to be aware of and

manipulate sound units in words. Phonological awareness is the broader term and encompasses phonemic awareness, which is defined by awareness at the level of phonemes, the smallest unit of spoken language. While children as young as preschool-aged may be taught phonological skills such as clapping syllables and rhyming, phonemic awareness typically develops later and is considered more developmentally advanced. Common phonemic awareness activities involve manipulating phonemes by deletion, addition, substitution, and segmenting of sounds. Table 2 provides examples of each intervention type that was coded.

Table 2

Examples of Intervention Types

Name	Cases	Example
Early Phonological Awareness	4	Whole language activities including reading and writing, plus oral metalinguistic games (singing, dancing, looking at pictures), each representing an aspect of phonological awareness above level of phoneme, awareness of sound, worked up to syllables, identifying initial and middle phonemes in words. (Brennan & Ireson, 1997)
Phonemic Awareness	18	Focus of intervention was on establishing sound-symbol connections, in addition to letter-sound and letter-name training. Activities included phoneme segmentation using letters and tiles, sound categorization using rhyme or alliteration, and sound-letter correspondence using CVC words. (Ball & Blachman, 1991).

(*Table 2 continues*)

(*Table 2 continued*)

Name	Cases	Example
Comprehensive Phonics	23	Corrective Reading: Direct Instruction remedial reading program. Level A focuses on reviewing letter-sound correspondences through rhyming, blending and segmenting activities. Level B includes word-attack skills, practicing pronouncing words, identifying sounds of letters or letter combinations, and reading isolated words composed of sounds and sound combinations. (Hempenstall, 2008).
Comprehensive Text Reading	4	Explicit and systematic instruction in word-level reading skills as well as comprehension, vocabulary, decoding, encoding, and fluency. Minimal emphasis was given to instruction in phonological awareness. Majority of lesson time was spent in explicit word-level instruction and practice, with application of skills in connected text reading with teacher support. (Denton, Wexler, Vaughn, & Bryan, 2008).
Orthographic Word Study	9	Structural Analysis. Instruction in phonological and orthographic skills, alphabetic principle training, phonological decoding, and oral reading of connected text for meaning. Emphasized word analysis at the syllable/morpheme level, with some attention to affixes and etymology. Divided words into syllables, sounded out the spelling-sound correspondences within syllables, then blended syllables to make a whole word. (Abbot & Berninger, 1999).

(*Table 2 continues*)

(*Table 2 continued*)

Name	Cases	Example
Orthographic Word Study with Spelling	29	Knowledge of the spelling system was enhanced by teaching common sound-letter combinations, spelling patterns or rules involving long and short vowels, and frequently occurring phonograms or rimes. Students learned how to spell frequently occurring words by correcting misspelling Built words using onsets and rimes. (Graham, Harris, & Chorzempa, 2000).
Other	4	Designed to boost reading and spelling through phonological awareness/word attack skills. Students did dictation using their own recorded voice linked to text. Rationale is that learners should benefit from hearing themselves vocalize text with immediate feedback and constant opportunity for self-correction. Kids worked at their own pace, at different levels of difficulty built into the program. (Brooks, Miles, Torgerson, & Torgerson, 2006)

Phonemic awareness refers to interventions that expose children to letters/word and explicitly connect sounds to letters. *Early phonological awareness* interventions consisted of similar activities as above, but instruction focused primarily on phonemes as sounds and letters were not explicitly taught. For instance, children may be clapping out syllables, or use blocks to represent phonemes in spoken words. *Phonics* interventions are typically multicomponent programs designed to teach children to read, based on explicit instruction in phonemic and phonological awareness (as above). However, along with the latter, phonics programs included other reading-related instruction such as sight words, vocabulary, story reading, etc.

Orthographic word study involved systematic methods of decoding and spelling using larger sub-lexical units such as onsets, rimes, syllables, and morphemes. This is distinct from pure whole word strategies, in which students simply learn words by exposure or repetition without explicit application of sub-lexical analysis. Students may also learn common spelling units such as the /ai/ long vowel sound in *drain* and *stain*. *Orthographic word study with spelling* often included similar activities relating to word analysis as well as phoneme segmentation, but the goal of instruction was specifically to improve spelling. The distinguishing feature of spelling interventions was that students spent a significant amount of time writing correct spellings of words, as opposed to using spelling strategies to decode.

Comprehensive text reading interventions were typically implemented with children in Grade 3 and above and emphasized reading fluency and comprehension as broad goals. Although this sometimes involved some review of phonics, the instruction moved beyond word decoding and included reading text. An *Other* category accounted for miscellaneous interventions that did not fit into the six coding categories, such as an interactive computer intervention that involved spelling error correction through vocal feedback.

Length of intervention. This moderator refers to the range of time during which intervention took place. All interventions except those were field-based and took place on a regular basis, varying from 2 times per week to daily, with sessions ranging from 10 minutes to 45 minutes. Interventions that were coded as less than 2 week typically involved training of a specific skill in a short period of time, ranging from one day to two weeks. Other interventions were coded as less than three months, three to six months, inclusive, one academic year, and multi-year. Because the exact time of pre- and post-testing varies due to logistical factors in

field research, the actual length of a year-long intervention may range from 6.5 months to 8 months.

Instructor. This moderator indicates whether the person who provided instruction for the majority of the intervention. Interventions in which children primarily used a computer program were coded as "Computer." Inclusion of this moderator is warranted by recent research which has pointed to discrepancies in the effects of literacy interventions delivered by regular classroom teachers versus researchers. Interventions are generally more effective when they are taught by children's regular classroom teachers.

Instructional group size. The instructional group size refers to the primary mode in which students received the intervention by the designated instructor. Students received instruction through one-on-one tutoring, small-group, and whole-class units. Small groups ranged from two to ten children, and whole-class groups ranged from ten to an unspecified number of children in a classroom. In three studies, children were instructed in different instructional units in different sessions equal amounts of time (e.g., tutoring one day per week and small groups one day of the week); in these studies, the smaller unit was chosen. Studies in which children were reported to receive instruction in all group sizes (including tutoring, small group, and whole class) were coded as *mixed.* If students were taught the intervention primarily as a class and then worked in small groups or individually with the teacher at some points throughout the lesson, this was coded as *whole-class.*

Participant characteristics.

Country. The country in which the study took place was coded in order to account for differences in formal education systems by country. For instance, children in the UK and New Zealand may start formal schooling up to one year earlier than American children. Studies were

coded as *US*, *earlier*, or *later* based on the age of entry into formal schooling in comparison to the US.

Grade level. The age of participants at the beginning of the intervention was coded as both the mean chronological age and grade level. The mean age was estimated using grade levels in cases where the mean age was not reported. For moderator analyses, grade levels Grade levels were coded by the following categories: pre-kindergarten, kindergarten, 1st grade, 2nd and 3rd grade, 4th and 5th grade, and 6th grade and above.

Literacy profile. Four categories of participants' literacy profiles were coded. Participants who had no documented deficits or who were described as typically achieving were coded as *typical*. This category also included samples for which no screening protocols were mentioned. Participants were coded as *at-risk* for future reading difficulties if their reading or spelling abilities were below average compared to their general or verbal IQ. This status is typically given to younger children in kindergarten, 1st, and 2nd grade. Samples that were predominantly low-income were also coded as *at-risk*. Third, participants in third grade and beyond who had been experiencing difficulty reading and spelling in spite of their relatively typical cognitive/intellectual levels were coded as *learning disabled*. These participants had either been diagnosed with specific disabilities, or they had been screened as poor readers despite having average cognitive scores. Finally, participants who were receiving their core reading curriculum in English and were learning English as a second language were coded as *ELL*. If the majority of the sample was both ELL and at-risk, the sample was coded as *ELL*. Finally, samples were coded as *mixed* only in cases in which two or more disparate groups were included in the same sample, such an otherwise typical sample with significant proportions of ELL and at-

risk children. This occurred when, for instance, a study failed to report separate sub-sample data.

Study methodology. Moderators associated with the methodology reported in studies included study design, the type of control group used, the assessment of treatment fidelity, and the assessment of initial group differences. The purpose of these moderators was to determine the general impact of study quality on ESs. Because no one variable can reliably predict study quality, multiple variables were included to best capture this construct.

Design. All included studies were group comparison designs. Studies that randomly assigned participants to treatment conditions were coded as *experimental*, whereas studies that used intact groups were coded as *quasi-experimental*. Large-scale, cluster-randomized trials in which multiple classrooms from multiple schools were randomly assigned, were also coded as *experimental*.

Type of control group. The type of control group used was coded to determine whether ESs varied depending on the type of treatment the participants received. *No contact* groups continued with their existing practice, which was typically based on traditional or basal literacy curricula. *Literacy* groups received an alternative treatment other than the intervention. These conditions were typically diluted, less systematic versions of the intervention and were intended to control for either instructional intensity or a specific aspect of the intervention. *General language* groups participated in a general language activity, such as talking about stories or semantic categorization of words, with little to no attention to literacy-specific activities. These conditions were typically intended to control for exposure to books and words used in the intervention condition. *Non-language* groups received instruction involving non-language

activities, such as mathematics or general classroom skills. These conditions were intended to control for general exposure to an intervention.

Initial group equivalence. Another indicator of methodological rigor is whether an attempt was made to assess initial group equivalence at the beginning of the intervention. This assessment is necessary in order to attribute intervention effects to the intervention itself, rather than to pre-existing differences between groups. This moderator was coded as *no significant differences* if groups were not statistically different from each other at baseline, as *present and statistically adjusted* if significant group differences were found and statistically adjusted for in the subsequent analysis, and as *present and not statistically adjusted* if significant group differences were found but were not addressed in the analysis. Thirty-five studies failed to report this information and were therefore excluded from this moderator analysis.

Treatment fidelity. Studies varied in the methods that were used to assess treatment fidelity, if any. Studies were coded as *assessed* or *not assessed*, based on whether the reported interventions were directly observed by research staff.

Extrinsic variables. Extrinsic variables were coded to determine the extent to which variables other than those relating to substantive or methodological qualities of the studies themselves influenced ESs **(**Lipsey, 2009). These variables include the journal impact factor, as well as a variable indicating whether means used in ES computation were adjusted for covariates.

Publication year. Studies were coded by publication year as *earlier* than 2000 or *later* than 2000 to detect whether ESs varied by the time they were published. A difference in ES by publication year could suggest several historical effects, including the kinds of interventions implemented, changes in the literacy curriculum, and broader cultural influences such as changes

in technology. These historical effects may also include differences in research reporting practices.

Journal impact factor. A commonly cited drawback of meta-analysis is the file drawer effect, or the potential bias that can occur in the aggregation of ESs due to the fact that published studies are more likely to report significant or greater effects. This is a problem associated with all research and is not unique to meta-analysis. As mentioned previously, all efforts were made to complete a systematic, comprehensive search of the literature. To explore this issue empirically, journal impact factors were coded to explore whether higher tier journals were more likely to report greater ESs. Journals were coded as *first-tier* (> 2.0), *second-tier* (< 2.0), and unpublished, which included doctoral dissertations. Book chapters, as well as journals for which no impact factors were available, were coded as *second-tier*.

Adjusted means. In 24 studies, authors reported post-test means that were adjusted in an analysis of covariance by at least one covariate, such as the pre-test score or other relevant variables. In these cases, adjusted means were used in place of observed means to calculate the mean difference between groups, as it is expected that adjusted means are reasonable and sometimes more accurate estimates than observed means (Bornstein, 2005; Lipsey & Wilson, 1993). Unadjusted standard deviations were still used as the pooled standard deviation to avoid any reduction in variance associated with covariates. ESs were coded as *adjusted* or *unadjusted* to determine whether ESs computed from adjusted means differed from ESs computed from unadjusted means.

CHAPTER 4: Results

The overarching goal of this meta-analysis was to address the impact of systematic, code-based literacy instruction on spelling achievement. In presenting the results, we examine the effects of various interventions on spelling, and to what extent these effects were moderated by relevant variables. Although moderator analyses helped to elucidate certain findings such as distributions of studies across variables of interest, these analyses were somewhat constrained due to a confounding in the sample of studies between two important types of moderators: intervention and participant characteristics. In some ways the confounding of these two types of moderators reflects the simple fact that the nature of interventions inherently differs by the age of the participant based on what is considered developmentally appropriate. For example, kindergarten children are more likely than third-grade children to receive phonics interventions. But this overlap also stems from the limits of the existing literature, as much of the research on interventions with younger children is dominated by phonics-related interventions and includes little else. Similarly, there is a relative dearth of studies on literacy interventions that address spelling for older students. To address this issue, we conducted stratified analyses by three age groups, in addition to the full meta-analysis.

To best convey the findings, results of the meta-analysis are presented in four sections. Presented first are the results of the full-sample meta-analysis, including both overall effects and moderator analyses. Presented next are analyses from three subsets of studies corresponding to the following developmentally defined age groups: pre-readers (preschool and kindergarten), developing readers (first, second, and third grade), and proficient readers (fourth grade and above). These smaller analyses were conducted to provide a more meaningful lens through which to interpret the data.

The majority of the analyses were conducted on studies that measured spelling as an outcome at post-test, as this was the most common dependent variable reported in studies. Seventeen studies that included only follow-up spelling measures were excluded from most analyses (See Appendix C for a list of references), except for the moderator analysis of *time of measurement.*

Overall Effects

Ninety-one studies yielding 153 ESs measured at post-test were analyzed using a random-effects model (See Table 3 for a summary). The total number of participants represented by the sample of studies was 9,046, and the sample size for each study ranged from 12 to 912, with a median of 61. ESs ranged from 0 to 2.45. The mean ES at the study level was moderate, d = 0.58, 95% CI [.49, .68], indicating that systematic code-based literacy interventions are more effective at improving spelling outcomes than non-code-based or less systematic code-based interventions. These ESs also varied significantly from each other, Q = 336.76 (90), p = .001, and 73.28% of this heterogeneity was attributable to differences among studies, rather than to chance, as indicated by the I^2 statistic. This variation was not surprising, as the study sample comprised several different types of interventions and participants. At the ES level, the mean ES was moderate, d = 0.61, 95% CI [.53, .68], p = .001. There was a significant amount of heterogeneity in computed ESs, Q = 552.46 (152), p = .001. (See Appendix B for a full descriptive sample of studies, each with its corresponding ES and selected moderators.)

Sensitivity analyses. Before proceeding to moderator analyses, we conducted several sensitivity analyses to determine the extent to which the magnitude of the overall ES was susceptible to statistical vagaries in the data, such as outliers and potential bias from any missing studies. First, we examined the distribution of ESs for outliers. Visual inspection of the 153

Table 3

Summary of Effect Sizes for Overall Effects at the Level of Studies and Comparisons (Numbers of Studies and Comparisons are in Parentheses).

	Level of Analysis	Cohen's *d*	95% CI	Z	*p*-value (Z)	*N*	*Q*	*df* (*Q*)	*p*-value (*Q*)	I-squared
Studies (91)	Fixed	.49	[.45, .53]	22.15	.001	9046	336.76	90	.001	73.28
	Random	.58	[.49, .68]	12.53	.001					
Comparisons (153)	Fixed	.52	[49, .56]	27.70	.001	12471	552.46	152	.001	72.49
	Random	.61	[.53, .68]	15.67	.001					

computed ESs revealed two studies with three ESs that were clear outliers (Armand, Le Francois, et al., 2004; Castle, Riach, & Nicholson, 1997 Study 1). These outliers were more than two standard deviations greater than the mean ES (d = 3.14 and 5.27, respectively). Upon closer examination of these two studies, we did not detect any evidence that these studies were substantively different from the rest of the studies in the sample. Therefore, we opted to retain the data from these studies rather than to remove them entirely from the study sample. To avoid skewing the ES distribution, ESs from these studies were "windsorized" (Lipsey and Wilson, 1993) by imputing a value that was two standard deviations from the mean ES of all the studies (d = 2.45). A value of two standard deviations was chosen, as there was a clear break in the distribution is this range. Three ESs total from these studies were windsorized to a value of 2.45.

Next, we conducted an assessment of publication bias using Duval & Tweedie's (2000) "trim-and-fill" method, which estimates the mean ES after imputing values for potentially missing studies, usually those to the left of the mean (i.e., those with small or null results). Using this method, the estimated value of the mean ES after imputing 29 missing studies based on the random effects model was d = .33, 95% CI [.25, .46]. Based on this estimation, the magnitude of the overall ES would be reduced from moderate to small. It is important to acknowledge that "this is a method used to adjust for the effects of publication bias, not a method for determining whether it exists" (p. 508, Hunter & Schmidt, 2000). Essentially, the trim-and-fill method yields a more conservative estimate of the true ES by accounting for the possibility of Type I error.

We also conducted Orwin's failsafe-K analysis as another way of assessing publication bias. At the level of studies with a two-tailed alpha set to .05, this analysis suggests that 12,083

studies would be needed to reduce the effects to non-significance. This large number of studies suggests that the overall ES is relatively robust to the impact of missing studies. Although the failsafe-K value provides some information, it is only one statistic, and therefore should be interpreted judiciously.

Moderator Analysis

To explore whether intervention effects varied based on variables of interest, we conducted an exploratory analysis of twenty moderators. These moderators were categorized as variables associated with characteristics of the intervention, participant, outcome, methodology, and extrinsic factors of each study (Cooper, 2009). We emphasize that moderator analyses used in meta-analysis are essentially correlational in nature, as the moderators in question were not randomly assigned to studies. Therefore, the following results should be interpreted as suggestive of possible associations between variables, rather than indicative of causal relationships. With the exception of *time of measurement*, *type of measure*, *type of control group*, *type of measure*, *scoring method*, and *adjusted means*, which were analyzed at the level of comparisons, all moderator analyses were conducted using study as the unit of analysis.

It is important to note that including multiple comparisons from individual studies in the analysis violates the assumption of statistical independence and may lead to overestimation of the precision of the ES estimate. However, given the research goals of this study, multiple comparisons were included in order to maximize the data coded from studies. Thus, we opted to include the comparisons and interpret them carefully rather than exclude them altogether (Lipsey & Wilson, 2001).

Effects moderated by time of measurement. In order to assess the robustness of overall effects over time, we conducted an analysis at the level of comparisons using time of

measurement as a moderator (See Table 4). This analysis showed that ESs computed at post-test (d = .60) were significantly greater than those computed at follow-up time points (*d* = 0.46), *Q* = 6.12 (1), *p* =.013. Thus, there is evidence of a slight decay in intervention effects between the time of immediate post-test and up to two years afterward, though the effects remain moderate. These results suggest that the benefits of code-based interventions on spelling outcomes for children may attenuate if the instructional content of these interventions is not continuously reinforced in children's regular literacy curriculum.

Table 4

Effect Sizes Moderated by Time of Measurement at the level of Comparisons.

Time of Measurement	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Post-test	.60	[.54, .67]	17.57	.001	153	12471		
Follow-up	.46	[.36, .55]	9.50	.001	68	7600		
Between-classes effect							6.12 (1)	.013

Effects moderated by characteristics of interventions. One of the primary goals of this meta-analysis was to explore the heterogeneity of effects, if any, among different kinds of interventions (See Table 5). Contrary to our expectations, there was no significant heterogeneity in ESs, *Q* = 9.73 (6), *p* = .001 based on intervention type. One reason for this could be that effects of code-based interventions were robust across all interventions, such that any intervention was superior to control conditions. Another possible interpretation of these results is that there is not enough statistical power to detect differences among interventions, as generally statistical power in moderator analyses is low (Lipsey & Wilson, 2001).

Table 5

Effect Sizes Moderated by Type of Intervention at the level of Studies.

Intervention	*d*	95% CI	*Z*	*p-value (Z)*	*K*	*N*	*Q (df)*	*p-value (Q)*
Early Phonological Awareness	.66	[.20, 1.12]	2.83	.005	4	216		
Phonemic Awareness	.79	[.57, 1.00]	7.25	.001	18	1296		
Phonics	.47	[.30, .64]	5.40	.001	23	3035		
Text Reading	.43	[.01, .84]	2.02	.043	4	479		
Word Study	.52	[.23, .82]	3.51	.001	9	895		
Word Study with Spelling	.67	[.50, .84]	7.67	.001	29	2645		
Other	.25	[-.16, .66]	1.19	.233	4	481		
Between-classes effect							9.73 (6)	.137

This issue may be exacerbated by the issue of confounds among intervention characteristics and other moderators, as discussed above. It is possible that the interventions do, differ in important ways, but that these differences are not statistically detectable in this sample of studies.

Despite the lack of statistically significant variation among studies, mean ESs for all interventions except *comprehensive text reading* and *other* were statistically greater than zero. Participants in the intervention group exceeded participants in the control group on spelling outcomes after receiving interventions in early phonological awareness ($d = 0.66$, $p = .005$), phonemic awareness ($d = 0.79$, $p = .001$), comprehensive phonics ($d = 0.47$, $p = .001$),

comprehensive text reading (*d* = 0.43, *p* = .043), *orthographic word study* (*d* = 0.52, *p* = .001), and *orthographic word study with spelling* (*d* = 0.67, *p* = .001). Interventions coded as *other* did not have a significant impact students' spelling achievement in comparison to control conditions, *p* = .233, though this ES estimate was only based on four studies. This moderator analysis was based on 18 studies involving phonemic awareness, 23 studies involving phonics, 9 studies involving orthographic word analysis¸ 29 studies involving orthographic word analysis and spelling, 4 studies involving comprehensive text reading, and 4 studies involving miscellaneous interventions.

Effects of encoding instruction. In fact, upon further inspection of the ESs among the different interventions, it was evident that the ESs for both *phonemic awareness* and *spelling* interventions were slightly greater than the ESs of the other interventions (See Table 6). One explanation for this pattern is that both of these intervention types involved generating spellings by segmenting phonemes and synthesizing these phonemes to build words. Therefore, we collapsed these two intervention types into one category, and compared these "encoding" interventions with all other interventions. This analysis showed spelling interventions

Table 6

Effect Sizes Moderated by Encoding Interventions

Adjusted Means	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Encoding	.72	[.58, .85]	10.54	.001	47	3941		
Other	.47	[.34, .59]	7.30	.001	44	5106		
Between-classes effect							7.12 (1)	.008

(d = 0.72, 95% CI [.58, .85]), or those interventions that involved tasks such as phoneme segmentation and writing out spellings for words, were significantly different from other interventions (d = 0.47, 95% CI [.34, .59]), Q = 7.12, p = .008. This finding suggests that the act of generating spellings of words, as well as the act of attending to the segmentation of phonemes in words, may play important roles in the effectiveness of interventions with respect to spelling outcomes.

Moderator analyses at the study level of both type of instructor and instructional group size were conducted to inform the extent to which instructional features of the interventions impacted spelling outcomes (See Tables 7 and 8). Interventions were taught by students' regular classroom teacher in 29 studies, by hired teachers or other school personnel in 57 studies, and by a computer program in 5 studies. There was no significant variation in ESs by instructor, Q = 1.92 (2), p = .383, indicating that whether instruction was primarily implemented by the regular classroom teacher (d = 0.51), researchers or other personnel (d = 0.64), or a computer program (d = 0.52) did not moderate the impact of the intervention. This finding suggests that interventions taught by instructors other than students' regular classroom teacher may still be effective.

In regards to the instructional group size used in the intervention, 18 studies involved one-to-one tutoring, 42 studies involved small groups, 28 studies involved the classroom, and 3 studies involved mixed groups. There was no significant heterogeneity among group size, Q = 0.93 (3), p = .818. It may be that interventions were effective in improving spelling outcomes across a variety of instructional group sizes, and that any differential effects are too small to detect. However, given the evidence for the benefits of tutoring and small-group literacy instruction (Vaughn, Hughes, Moody, & Elbaum, 2000), especially for at-risk students, it is possible that this finding against the moderating effects of group size on instruction may be an

artifact of the study sample. More research is needed to identify the effects of instructional group size on literacy outcomes, with respect to the population of students taught.

Table 7

Effect Sizes Moderated by Instructor at Level of Studies.

Instructor	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Regular Teacher	.51	[.36, .66]	6.63	.001	29	4641		
Other	.64	[.52, .76]	10.57	.001	57	3859		
Computer	.52	[.14, .89]	2.70	.007	5	546		
Between-classes effect							1.92 (2)	.383

Table 8

Effect Sizes Moderated by Instructional Group Size at the Level of Studies.

Instructional Group Size	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Tutoring	.53	[.33, .74]	5.07	.001	18	1484		
Small-group	.57	[.42, .71]	7.77	.001	42	2879		
Whole class	.63	[.47, .78]	7.76	.001	28	4520		
Varied	.76	[.19, 1.33]	2.60	.009	3	164		
Between-classes effect							0.93 (3)	.818

Although interventions of all lengths of time significantly impacted spelling outcomes, heterogeneity of ESs across this moderator was not significant, $Q = 0.24$ (4), $p = .993$ (See Table 9). The majority of studies lasted from less than three months to a year; there were only four studies of interventions that lasted less than two weeks, and only five studies that spanned multiple years. The mean ES of interventions that lasted less than two weeks ($d = 0.54$) was comparable to the mean ES of those that lasted less than three months ($d = 0.61$), three to six months ($d = 0.55$), one year ($d = 0.60$) and more than one year ($d = 0.59$). These finding suggests that even relatively short code-based interventions can impact spelling outcomes at post-test. Whether longer interventions have a proportionate impact on outcomes is not directly discernible by the data, though it is reasonable to suspect that, with a greater amount of instruction, children would make proportionate gains. More research in this area is needed.

Table 9

Effect Sizes Moderated by Length of Intervention at the Level of Studies

Length of Intervention	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Less than 2 weeks	.54	[.06, 1.01]	2.22	.026	4	222		
Less than 3 months	.61	[.42, .79]	6.49	.001	26	1625		
3 to 6 months	.55	[.37, .73]	6.04	.001	24	2220		
Year	.60	[.44, .75]	7.69	.001	32	4389		
Multi-year	.59	[.22, .97]	3.09	.002	5	591		
Between-classes effect							0.24 (4)	.993

Effects moderated by participant characteristics. The grade levels represented in the study sample were 11 preschool, 29 kindergarten, 21 first grade, 20 second and third grade, 7 fourth and fifth grade, and 9 sixth grade and up (See Table 10). Three studies reported samples with a large age range, such as 2nd to 8th grade; we excluded these studies from the analyses, as we could not meaningfully code them into a grade category. Intervention effects were significantly greater than zero for kindergarten (d = .66), grade one (d = 0.59), grades two and three (d = 0.43), grades four and five (d = 0.69), and grades six and up (d = 0.77). Intervention effects for preschool (d = 0.39) and mixed (d = 0.21), both of which represented only 2 and 3 studies, respectively, were not significantly greater than zero. Because a large portion of the literature on literacy research is comprised of studies on phonics-related interventions for younger children, children in grade two and below are overrepresented in the study sample in comparison to older students. This uneven distribution of studies may account for the lack of significant heterogeneity of ESs among studies, Q = 7.47 (6), p = .280, as grade level would otherwise be expected to moderate the impact of interventions.

For possibly related reasons, there was no significant variation in ESs across participant literacy profile, Q = 7.78 (4), p = .100 (See Table 11). Among the study samples represented, 30 were typically achieving, 35 were at-risk, 15 were reading or learning disabled, 7 were ELL, and 4 were mixed. Intervention effects were statistically greater than zero for all profile types, including typical (d = 0.56), at-risk (d = 0.57), reading or learning disabled (d = 0.48), ELL (d = 0.97), and mixed (d = 0.42). It is possible that intervention effects were robust across participants with different levels of reading ability. Despite the statistical findings of this moderator analysis, it is also possible that a child's literacy profile has at least some effect on

Table 10

Effect Sizes Moderated by Age at the Level of Studies.

Age	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Preschool	.39	[-.21, .99]	1.28	.199	2	182		
Kindergarten	.66	[.49, .82]	7.87	.001	29	3015		
Grade 1	.59	[.40, .78]	6.11	.001	21	2379		
Grades 2 and 3	.43	[.23, .63]	4.18	.001	20	1779		
Grades 4 and 5	.69	[.33, 1.06]	3.75	.001	7	385		
Grades 6 and up	.77	[.48, 1.05]	5.22	.001	9	1031		
Mixed	.21	[-.30, .72]	0.81	.420	3	276		
Between-classes effect							7.47 (6)	.280

Table 11

Effect Sizes Moderated by Literacy Profile at the Level of Studies.

Literacy Profile	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Typical	.56	[.41, .72]	7.22	.001	30	3725		
At-risk	.57	[.43, .72]	7.93	.001	35	3364		
RD/LD	.48	[.24, .72]	3.90	.001	15	739		
ELL	.97	[.67, 1.28]	6.24	.001	7	651		
Mixed	.42	[.01, .82]	2.03	.042	4	567		
Between-classes effect							7.78 (4)	.100

their responsiveness to interventions, but that these moderating effects were not detected in this study sample. Because of confounds between intervention and participant characteristics, it is not possible to confirm this claim based on findings from this data.

We also conducted analyses of the effects of SES (See Table 12). Intervention effects did not vary based on SES, $Q = 2.08$ (2), $p = .556$, as ESs for low SES ($d = 0.65$) were comparable to the mean ES for middle ($d = 0.47$) and mixed SES ($d = 0.63$). Because the *at-risk* literacy profile was partially defined by low SES, there was some redundancy between these two moderators, and this may account for the lack of significant heterogeneity across this moderator. It is also important to acknowledge that there was some amount of overlap between SES and literacy profile, as by definition any low SES participant was coded as *at-risk*, as researchers have argued that SES is one of the strongest predictors of reading scores (Lee & Burkam, 2002).

Table 12

Effect Sizes Moderated by SES at the Level of Studies

SES	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Low	.65	[.50, .80]	8.57	.001	32	3851		
Middle	.47	[.27, .67]	4.52	.001	21	1567		
Mixed	.63	[.36, .90]	4.59	.001	11	1037		
Between-classes effect							2.08 (2)	.556

In regards to the country where the intervention took place, there was no significant difference between intervention effects for students in the US (d = 0.58) versus for students in other countries (d = 0.59), Q = 0.03 (1), p = .864 (See Table 13). In other words, nothing in the findings from this data suggests that children in the US are any more or less responsive to interventions than children from other countries.

Table 13

Effect Sizes Moderated by Country at the Level of Studies

Country	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
US	.58	[.46, .70]	9.68	.001	55	6149		
Other	.59	[.45, .74]	7.90	.001	36	2897		
Between-classes effect							0.03 (1)	.864

Effects moderated by characteristics of dependent variable. At the level of comparisons, there was no significant difference between intervention effects based on researcher-constructed (d = 0.63) versus standardized measures (d = 0.58), Q = 0.38 (1), p = .535. Although previous research suggests that researcher-constructed measures are more likely to be biased toward study conditions and therefore yield higher ESs, findings from this dataset did not confirm these claims.

Neither did ESs significantly vary based on the scoring method used, Q = 1.56 (1), p = .212, as spelling ESs computed from conventional scoring methods (d = .58) did not differ from those computed from developmental scoring methods (d = .68) (See Table 15). This analysis of scoring method was based on researcher-constructed measures only, as most (54 out of the 63) of

the standardized measures were scored conventionally. If anything, we would expect developmental scoring methods to yield higher ESs because they are designed to be more sensitive to variation. It is possible that developmental scoring methods do, in fact, yield higher ESs, but that this difference was not detected in this study sample.

Table 14

Effect Sizes Moderated by Type of Measure at the Level of Comparisons

Type of Measure	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Standardized	.58	[.46, .70]	9.34	.001	62	4492		
Researcher-constructed	.63	[.53, .73]	12.47	.001	91	7979		
Between-classes effect							0.38 (1)	.535

Table 15

Effect Sizes Moderated by Scoring Method at the Level of Comparisons

Scoring Method	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Conventional	.58	[.48, .67]	12.28	.001	108	8278		
Developmental	.68	[.54, .82]	9.70	.001	45	4193		
Between-classes effect							1.56 (1)	.212

Effects moderated by characteristics of study quality. Differential effects on ESs based on variables associated with study methodology would suggest that results from more methodologically rigorous studies should be considered more valid than results from less methodologically rigorous studies. In this meta-analysis, however, there was no significant heterogeneity in ESs based on any indicator of study quality, including study design, Q = 3.48 (1), p = .062, type of control group, Q = 5.73 (3), p = .125, assessment of treatment fidelity, Q = 0.14 (2), p = .933, and assessment of initial group differences, Q = 1.52 (3), p = .678 (See Tables 16, 17, 18, and 19, respectively). It is possible that this lack of heterogeneity stems from variation in reporting practices, which would entail that these moderators do not adequately capture study quality. For instance, 13 studies failed to report assessment of initial group differences, and 28 studies failed to report assessment of treatment fidelity.

With respect to study design, our findings were contrary to research indicating that quasi-experimental studies tend to yield higher ESs than experimental studies (Bornstein, Hedges, et al., 2009). Instead, experimental studies (d = 0.65, 95% CI [.54, .77]) yielded slightly higher ESs than quasi-experimental studies (d = 0.48, 95% CI [.35, .62]), though this difference was not statistically significant. Based on the study sample, one reason for this may be that the experimental studies included a proportionately higher number of studies that involved interventions which were shorter in length and more focused; these interventions may have more likely to produce training effects at post-test. In fact, there are multiple reasons why studies with an experimental design would yield different ESs from those with a quasi-experimental design, and additional research is needed to speak to this issue.

Table 16

Effect Sizes Moderated by Design at the Level of Studies

Design	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Experimental	.65	[.54, .77]	11.02	.001	57	4311		
Quasi-experimental	.48	[.35, .62]	6.90	.001	34	4735		
Between-classes effect							3.48 (1)	.062

At the level of comparisons, there was no significant heterogeneity with respect to the type of control group used as a comparison to the intervention group. Given the importance of controlling for Hawthorne effects in intervention research, this moderator analysis was conducted to determine whether ESs yielded from no-contact control groups were greater than those yielded from contact control groups. (See Table 17.) Our findings showed that ESs yielded from no contact groups (d = 0.61, 95% CI [.50, .71]) were not significantly different from those yielded from non-language (d = 0.41, 95% CI [.15, .66]), general language (d = 0.83, 95% CI [.59, 1.07]), and alternative literacy (d = 0.59, 95% CI [.45, .73]) groups. There was some overlap between the type of control group and study design, as most experimental studies included a contact control group.

Effects moderated by extrinsic characteristics. Analysis of moderators extrinsic to substantive features of the study themselves, such as *journal impact factor* and *adjusted means*, was conducted to assess whether research artifacts influenced estimates of ESs. Journal impact factor was considered an extrinsic variable because the journal in which the study was published should not theoretically impact the computed ES for that study. Nevertheless, because studies reported in higher-tiered journals are more likely to report significant findings and greater ESs

Table 17

Effect Sizes Moderated by Type of Control Group at the Level of Comparisons.

Type of Control Group	*d*	95% CI	*Z*	*p-value (Z)*	*K*	*N*	*Q (df)*	*p-value (Q)*
No contact	.61	[.50, .71]	11.50	.001	76	8509		
Non-language	.41	[.15, .66]	3.05	.002	13	651		
General language	.83	[.59, 1.07]	6.76	.001	15	988		
Literacy	.59	[.45, .73]	8.22	.001	49	2323		
Between-classes effect							5.73 (3)	.125

Table 18

Effect Sizes Moderated by Assessment of Treatment Fidelity at the Level of Studies

Assessment of Treatment Fidelity	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Assessed	.57	[.44, .70]	8.58	.001	45	4371		
Not assessed	.61	[.41, .82]	5.94	.001	18	2778		
Not reported	.59	[.42, .77]	6.61	.001	28	1898		
Between-classes effect							0.14 (2)	.933

than those published in lower-tiered journals, it is important to account for how this may impact ES estimates. Based on this study sample, there was no evidence of publication bias, $Q = 0.89$ (2), $p = .640$, as studies from first-tier journals ($d = 0.61$, $p = .001$) reported comparable ESs to those from second-tier journals ($d = 0.55$, $p = .001$), as well as to unpublished studies ($d = 0.67$, $p = .001$) (See Table 20).

Table 19

Effect Sizes Moderated by Initial Group Differences at the Level of Studies

Initial Group Differences	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
None	.56	[.44, .69]	8.76	.001	48	4085		
Statistically adjusted	.64	[.47, .80]	7.45	.001	26	2547		
Not statistically adjusted	.74	[.32, 1.15]	3.50	.001	4	498		
Not reported	.49	[.25, .74]	3.98	.001	13	1917		
Between-classes effect							1.52 (3)	.678

Table 20

Effect Sizes Moderated by Journal Impact Factor at the Level of Studies

Impact Factor	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
First-tier (< 2.0)	.61	[.46, .76]	7.70	.001	31	2855		
Second-tier (> 2.0)	.55	[.42, .67]	8.40	.001	49	4910		
Unpublished	.67	[.41, .94]	5.05	.001	11	1281		
Between-classes effect							0.89 (2)	.640

Publication year also did not significantly impact ESs, $Q = 0.30$ (1), $p = .582$. Mean ESs from studies published between 1981 and 2000 ($d = 0.62$, $p = .001$) were comparable to those published between 2001 and 2012 ($d = 0.57$, $p = .001$) (See Table 21). Thus, based on this study sample, there is no statistical evidence that significant historical factors influenced the impact of

interventions on spelling outcomes. Policy decisions and gradual changes in culture may certainly affect educational outcomes, but our findings do not validate this claim.

Table 21

Effect Sizes Moderated by Publication Year at the Level of Studies

Publication Year	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Early (1981 - 2000)	.62	[.46, .78]	7.75	.001	35	3082		
Later (2001 – 2012)	.57	[.45, .68]	9.72	.001	56	5964		
Between-classes effect							0.30 (1)	.582

Because of differences among studies in overall design and statistical methodology, covariate-adjusted means were used in the computation of 35 ESs (See Table 22). We compared ESs computed from adjusted and unadjusted means to determine whether there was any systematic difference between the two methods. If anything, we would expect adjusted means to yield slightly higher and more accurate ESs, as adjustment for covariates tends to result in higher statistical power. However, the analysis showed that ESs computed from adjusted (d = 0.63) did not differ statistically from ESs computed from unadjusted means (d = 0.60), Q = 0.08 (1), p = .773. This finding suggests there is no reason to assume that one type of ES is more valid or accurate than the other.

Summary of moderator analyses. We conducted an exploratory analysis of 17 moderators representing characteristics of the intervention, participant, outcome measure, study methodology, as well as extrinsic variables. With the exception of time of measurement, there was no statistically significant heterogeneity across any of the moderators. As previously

discussed, this is more likely due to the uneven distribution of studies and low statistical power rather than to a true lack of variation. To address this issue further, we present results in the following three sections by age group, and focus on intervention and participant characteristics.

Table 22

Effect Sizes Moderated by Adjusted Means at the Level of Comparisons

Adjusted Means	*d*	95% CI	*Z*	*p-value (Z)*	*k*	*N*	*Q (df)*	*p-value (Q)*
Adjusted	.63	[.47, .79]	7.68	.001	35	2885		
Unadjusted	.60	[.52, .69]	13.60	.001	118	9586		
Between-classes effect							0.08 (1)	.773

Effects of Interventions for Pre-Readers

Results of intervention effects in this section are presented for a subset of 27 studies on pre-school- and kindergarten-aged children, who may be considered pre-readers according to theoretical models of spelling and reading. With the exception of two studies that sampled pre-school children, nearly all of the studies sampled kindergarten children. To facilitate interpretation for the subsequent analyses, we collapsed the early phonological awareness intervention code with phonemic awareness, to yield an overall code for phonological awareness. (This broader phonological awareness code applied to analyses for developing readers, as well.)

Three studies that included a mixed sample of age groups (Hempenstall, 2008; Lovett, Lacerenza, et al., 2000; Lovett, Warren-Chaplin, et al., 1990), as well as four studies that included a mixed sample of literacy profiles (Fuchs et al., 2001; Schedule, Justice, et al., 2008;

Simmons, Coyne, et al., 2011; Slattery-Gursky, 2003) were excluded from the subsequent analyses.

The overall effect of all interventions on spelling achievement for a total sample of 2,631 pre-readers was moderate, $d = 0.67$, 95% CI [.49, .77], $p = .001$. This finding suggests that code-based early literacy interventions have a moderate effect on pre-readers' spelling achievement. That children as young as five and six years-old may improve their spelling skills is notable, at least because of the notion that spelling cannot be taught, much less with this age group. But if children can improve in reading tasks from phonemic awareness instruction, as some research suggests (Ref), then there is no reason to believe that they cannot also improve their early spelling ability. There was also significant heterogeneity among studies in this subsample, $Q = 84.14$, $p = .001$. This heterogeneity is not surprising, particularly because children in this age group are developing rapidly in many different areas, and there may be considerable variability in each child's rate of development.

Table 23 shows the effects of interventions implemented for this age group, divided into *typical* and *non-typical* samples. Nine studies involved typically achieving pre-readers, and 22 studies involved non-typical samples, nearly all of which were at-risk. The larger number of studies on non-typical participants reflects the fact that many interventions in this age group are targeted toward at-risk populations as a way to prevent future reading deficits. In fact, only five out of the nine studies on typical children were conducted in the US. Although results show no statistically significant difference in ESs between typical and non-typical children, $Q = 1.31$, $p = .252$, it is possible that any difference may be obscured by the imbalanced distribution of studies, as only 9 studies represented typical samples.

Table 23

Effects of Interventions for Pre-Readers at the Level of Studies

Profile	Intervention	*k*	Cohen's *d*	95% CI	*Z*	*p-value (Z)*
Typical	Phonological Awareness	4	.65	[.31, .99]	3.71	.001
	Comprehensive Phonics	1	.88	[.30, 1.45]	2.99	.003
	Invented Spelling	3	.34	[.04, .64]	2.22	.026
	Other	1	.30	[-.55, 1.16]	0.69	.488
At-risk	Phonological Awareness	8	.98	[.75, 1.22]	8.21	.001
	Comprehensive Phonics	8	.47	[.29, .65]	5.04	.001
	Invented Spelling	2	1.00	[.53, 1.48]	4.17	.001

Interventions for both groups of pre-readers consisted of phonological awareness, comprehensive phonics, and spelling (which, for this age group, was characterized as invented spelling). One intervention which was defined as *Other* (Shepherd, 2011) involved a systematic "word wall" intervention designed to teach children to identify simple phonetic and orthographic patterns in words. The overall effect on spelling for both typical (d = 0.54, 95% CI [.29, .80], p = .001) and non-typical (d = 0.73, 95% CI [.55, .92], p = .001) pre-readers was moderate. The most notable finding is that non-typical children seem to benefit from supplementary phonological awareness instruction (d =.98), over and beyond comprehensive phonics programs (d = .47), as indicated by a significant post-hoc comparison, Q = 11.36, p = .001. It is possible that, in comparison with typical children, non-typical children may be more responsive to

additional explicit instruction in phonological awareness. As this finding is only based on eight studies per category, further research is necessary to reach any substantial conclusions on whether at-risk children learn to spell best when their regular phonics curriculum is supplemented with more intensive, supplementary instruction in phonological awareness.

Five of the interventions for this age group involved invented spelling methods designed for younger children. Invented spelling interventions had a small effect for typical children (d = .34), and large effect for non-typical children (d = 1.00), though these ESs were based on 3 and 2 studies, respectively. It is important to highlight the fact that all of these invented spelling interventions were designed to strengthen children's reading and spelling skills by awareness of phonemic and orthographic features of words (Santoro, Coyne, & Simmons, 2006; Senechal, 2012). These methods contrast with some less code-based invented spelling methods that have been associated with whole-language or meaning-oriented literacy curricula. Practitioners of this method may simply encourage children to independently "invent" their own spellings without explicitly emphasizing the phonemic and orthographic properties of words.

Although more research in this area is needed, it seems reasonable to suspect that the beneficial effects of these interventions reported from these studies were at least partially due to the instructional component emphasizing the integration of phonemic and orthographic knowledge. The theoretical reasoning behind both of these studies is that reading and spelling knowledge should be integrated in early literacy curriculum, and that this integration might be particularly helpful for those at-risk readers with specific deficits such as phonemic awareness.

In summary, findings suggest that at-risk children may benefit from supplementary instruction in phonological awareness, as children who received this intervention fared better in than those who received a comprehensive phonics intervention. In addition, invented spelling

interventions that emphasized phonemic and orthographic connections between words seemed effective for at-risk children, but less so for typical children.

Effects of Interventions for Early Readers

Results presented in this section are presented for a subset of 41 studies on children in first, second, and third grades, who may be considered developing readers. For ease of interpretation, as well as to facilitate analyses, all subsequent analyses (including those involving proficient readers), *orthographic word study* and *orthographic word study with spelling* were collapsed into one category called *orthographic word study.* These categories were collapsed for the following analyses because of the small number of studies in the *orthographic word study* category, as the majority of the interventions involved spelling.

It is not surprising that nearly half of the studies in the sample involved children in this age group, as much of literacy instruction is targeted to children in this area; if there is any interest in spelling, then, it is generally in regards to this age group. The overall effect of interventions on spelling achievement for a total sample of 4,157 participants was moderate, d = 0.51 (.38, .54), p = .001. There was also significant heterogeneity in this sub-sample, Q = 14.34 (40), p = .001. Based on these findings, it seems that existing code-based interventions have a similar impact on pre-readers as they do on developing readers. Table 24 displays intervention effects for this age group, for both typical and non-typical samples. Similar to the pattern in the subsample of pre-readers, there were more studies involving non-typical samples.

Interventions which involved explicit instruction in detecting and manipulating larger orthographic and morphological units had a statistically significant moderate effect on typical children, d = 0.55, 95% CI [.33, .77], p = .001. Given the extensive empirical evidence citing the benefits for phonological awareness interventions for children in this age group (Ref), t is likely

that the effect of these interventions was at least comparable to interventions involving phonological awareness (d = 0.40). However, to the small number of cases in the latter the ES for phonological awareness interventions did not reach significance, p = .082. Phonological awareness interventions did have a moderately large effect on non-typical children, d = 0.87, p = .001. Although the difference is not statistically significant, there seems to be a similar pattern of results as for at-risk pre-readers in regards to the additional benefits of focused phonological awareness intervention over and beyond comprehensive phonics.

Table 24

Effects of Interventions for Developing Readers at the Level of Studies

Profile	Intervention	*k*	Cohen's *d*	95% CI	*Z*	*p-value (Z)*
Typical	Phonological Awareness	2	.40	[-.05, .85]	1.74	.082
	Comprehensive Phonics	3	.45	[.14, .75]	2.86	.004
	Orthographic Word Study	11	.55	[.33, .77]	4.91	.001
	Other	1	.29	[-.21, .80]	1.15	.251
Non-Typical	Phonological Awareness	5	.87	[.41, 1.33]	3.72	.001
	Comprehensive Phonics	9	.46	[.15, .78]	2.89	.004
	Orthographic Word Study	7	.33	[-.06, .71]	1.66	.097
	Text Reading	2	.70	[.03, 1.38]	2.06	.040
	Other	1	.38	[-.55, 1.29]	0.82	.413

Interventions which involved explicit instruction in detecting and manipulating larger orthographic and morphological units had a statistically significant moderate effect on typical children, $d = 0.55$, 95% CI [.33, .77], $p = .001$. However, these types of interventions did not seem to be as effective for non-typical children, $d = 0.33$, though this ES did not reach significance, $p = .097$, as this category was only based on 7 studies. Still, it is conceivable that the benefits of interventions involving larger lexical units for non-typical children in this age group may be limited by these children's deficiencies in one or more areas, particularly phonemic awareness.

Interventions involving text reading had a moderate to large effect on non-typical children, $d = 0.70$, 95% CI [.03, 1.38], $p = .040$. Closer inspection of these two studies revealed that both interventions involved explicit word-level instruction and emphasized connections among multiple lexical units. For instance, one study conducted in Norway describes the Epi-Meta-Mastery-Approach, a comprehensive remedial reading program that employs an embedded approach to reading, and whose practitioners "[work] in mediating ways with letters, syllables, words and sentences that aim to integrate linguistic levels above the phoneme/grapheme and word level…" (Frost & Sorenson, 2007). Similarly, the intervention implemented by Blachman, Schatschneider, Fletcher, Francis, Clonan, Shawitz, & Shaywitz, 2004 incorporated a similar emphasis on "phonologic and orthographic connections in words." It is important to highlight the fact that both of these interventions were highly structured and systematic, such that a sufficient amount of time was spent on explicit word-level instruction. Although this ES was based on only two studies, the evidence from these two studies suggests that

In summary, findings for this age group point toward a possible benefit of interventions focusing on the integration of phonemic awareness with larger orthographic and morphological

units. Phonological awareness interventions had a large effect on non-typical children, which was similar to the pattern seem in pre-readers.

Effects of Interventions for Proficient Readers

Results of intervention effects for this section are presented for a subsample of 17 studies on proficient readers. This sample of participants ranged in grade level from 4th grade to 11th grade. (Two of these studies included a partial sample of third graders.) With the exception of one speech-recognition intervention categorized as other (d = 0.05) and two comprehensive text-reading interventions (d = 0.17), the rest of the interventions involved orthographic and morphological units. The two comprehensive reading interventions were distinct from those mentioned with developing readers, as these interventions had a broader focus and did not explicitly emphasize word-level connections.

The overall effect of interventions on spelling achievement for a total sample of 1,416 proficient readers was moderate, d = 0.76, 95% CI [.45, 1.07]. Although based on the point estimate of the ES it may appear that code-based interventions may have a slightly greater impact on spelling ability of older students, this difference was not statistically significant. However, given the relative dearth of studies that address spelling within this age group, more research is needed to substantiate this claim. There was significant heterogeneity among studies within the subset, Q = 95.99, p = .001. Table 25 displays the effects of interventions for typical and non-typical participants.

Perhaps the clearest finding is the relative dearth of intervention studies addressing spelling in this age group. Based on the sample of studies we retrieved, only four studies addressed spelling for typically achieving older children, one of which was conducted in England. Even while spelling is not commonly researched in this age group, we did not expect

Table 25

Effects of Interventions for Proficient Readers at the Level of Studies

Profile	Intervention	*k*	Cohen's *d*	95% CI	*Z*	*p-value (Z)*
Typical	Orthographic Word Study	3	1.15	[.90, 1.40]	9.05	.001
	Other	1	.05	[-.30, .39]	0.27	.790
Non-Typical	Orthographic Word Study	10	.84	[.49, 1.20]	4.65	.001
	Text Reading	2	.16	[-.71, 1.03]	0.36	.717

the number of studies in this age group to be so low. It is clear that more research is needed in this area. Despite the lack of literature in this area, there are studies that address interventions involving larger lexical units. In fact, the effect of these interventions was large for both typical (d = 1.15, 95% CI [.90, 1.40], p = .001) and non-typical (d = .84, 95% CI [.49, 1.20], p = .001). Again, as has been the pattern with the pre-readers and developing readers, more of the studies are conducted with non-typical samples. Given that many of these participants had either been diagnosed with reading disability, or were below grade level in one or more reading predictors, the fact that these interventions benefited spelling achievement is suggests that explicit instruction in larger lexical units can be beneficial for these students. That is, although reading disability is often characterized by deficiencies in phonemic knowledge, there are other types of knowledge involved in reading disability.

In summary, these findings suggest that interventions involving orthographic word study had a large effect on both typical and non-typical participants. However, the lack of studies indicates a need for further research, particularly for typical students.

CHAPTER 5: Discussion

Overall Findings

The primary goal of the meta-analysis was to investigate the extent to which code-based literacy interventions impact spelling achievement in school-aged children. The overall analysis suggests that explicit and systematic code-based literacy interventions had a greater impact (d = 0.58) on generative spelling outcomes than control conditions, which varied from typical classroom practice, limited code-based instruction, general language activities, and non-language activities. This finding directly opposes the claim that the acquisition of spelling knowledge is a passive process that occurs in all literacy contexts, as well as lay assumptions that knowledge about spelling cannot or does not need to be taught. Rather, the findings from this meta-analysis suggest that students who had received interventions that incorporated explicit instruction in phonological, orthographic, morphological knowledge fared better than their control-group peers on spelling outcomes. The implication of these findings is that, spelling, like reading, depends on code-based linguistic knowledge, and that most interventions that target some aspect of code-based knowledge are superior to the status quo.

A further analysis revealed that the effects of encoding interventions (i.e., *phonemic awareness* and *orthographic word study with spelling* collapsed together) were superior to all other interventions. The counterargument that the effects of spelling interventions were mainly attributed to mere exposure to words fails to hold, as in many of control group students were exposed to or spelled the same words (but were not taught specific word analysis strategies). Students who received spelling interventions students spent a considerable amount of time generating spellings of words, either by segmenting phonemes, or by applying various types of sub-lexical knowledge. While this result may initially seem self-explanatory, this finding sheds

more light on the growing focus in the literature on the role of larger lexical units in learning to read and spell (Bear, Invernizi, Templeton, & Johnston, 2004; Berninger, 1998; Goswami & Bryant, 1990; Nunes, 2003). In regards to impacting spelling instruction, incorporating these larger lexical units in decoding instruction may not be sufficient. Rather, students may need to directly apply this knowledge in the process of spelling out words in order to improve their spelling ability, not simply in decoding them (Oulette, 2010).

A major constraint in the moderator analyses of this meta-analysis was the confound between intervention and age, namely that interventions focusing on the phoneme level were nearly all implemented with children younger than grade two, while interventions focusing on larger units such as onset-rimes and morphemes were implemented with older students. This confound may explain why age was not a significant moderator of ESs in the analyses. Further, the literature shows that there is a dearth of studies on spelling interventions for older children. Out of 29 studies of spelling interventions, only 8 of these studies reported interventions implemented with children in fourth grade and above. This limitation is a reflection of the state of the literature and the state of instructional practice, both of which have placed far more emphasis on phoneme level activities, particularly for younger students.

With respect to moderators such as group size, we did not find that it significantly moderated ES. Despite the fact that tutoring and small group instruction has been found to b significantly more beneficial than classroom-instruction, some researchers have argued that it is difficult to actually localize the effect of group apart from the population, as most interventions involving tutoring or small groups are done with at-risk or disabled populations. Contrary to our expectations, intervention effects did not vary significantly by participant literacy profile. Given the increasing number of students categorized as English-language learners, or diagnosed with

learning disabilities, language disorders, and other profiles, the question of how spelling should be taught must account for these different populations. For instance, recent empirical evidence shows that SES (and therefore at-risk status) accounts for more unique variation in reading scores than any other factor (Lee & Burkam, 2002).

Implications Pre-Readers

Overall, the effects of code-based interventions on spelling outcomes for pre-readers was moderate. This finding was similar to the effects reported by Ehri's (2001) meta-analysis (d = 0.67), though this meta-analysis included first graders and was limited to the effects of systematic phonics instruction. However, the findings from the present study also suggest that focused, supplemental phonological and phonemic awareness instruction had a large effect on the spelling achievement for at-risk children. Although results are inconclusive because of the limitations of the study sample, this large effect of focused, supplemental phonological seems to be superior to the effect of comprehensive phonics interventions. This result would suggest that at-risk children may need more intense or a greater amount of instruction in phonological awareness than is typically given in comprehensive phonics programs. More research is needed to determine the length and intensity of intervention that is needed in order for students to receive maximum benefits.

Because of the recent popularity of whole language programs, which often incorporate invented spelling activities with little explicit code-based instruction, it is important to address the effectiveness of this type of instruction for pre-readers. There is some evidence from our findings that invented spelling interventions that included explicit code-based instruction had a moderate impact on pre-readers' spelling achievement, though more studies are needed to validate this claim. It seems that invented spelling can be an effective method to improve

spelling outcomes for pre-readers, as long as the instruction focuses on letter-sound correspondences and incorporates guided feedback on orthographic patterns. However, more research is needed in this area, as there were only five studies of invented spelling in this study sample.

It is possible that encouraging children to create their own spellings of words without explicitly guiding them conventional spellings may hinder their spelling development. Studies with college students and adults have shown that exposure to misspellings may negatively impact spelling performance shortly after exposure (Brown, 1988; Jacoby & Hollingshead, 1990). If adults, who have many more years of print exposure and reading experience than pre-readers, are vulnerable to misspellings of words, then it is reasonable to assume that younger children would be at least equally as vulnerable. Invented spelling instruction for this age group should incorporate at least some amount of guided and explicit instruction, especially because children of this age are just beginning to become aware of sound-letter correspondences.

Implications for Developing Readers

Based on the distribution of studies involving this age group, there seems to be a greater emphasis on orthographic instruction. This is not surprising, given that children at this age are developing more sophisticated ways to decode and spell words beyond simple sound-letter correspondences. Similar to the pattern seen in studies among pre-readers, focused phonological awareness interventions also had a large effect on non-typically achieving children, including those who were at-risk, reading disabled, or learning disabled. As other studies have shown, this finding suggests that struggling spellers may benefit most from spelling training at the phonological level, whereas typical spellers may gain most from more advanced instruction that highlights other aspects of word knowledge (Worthy & Invernizi, 1990).

Although further research is needed to reach more substantial conclusions, there is some evidence that interventions emphasizing the connection between phonological and orthographic features of words may impact spelling achievement for children in this age group (Stage & Wagner, 1992). Orthographic word study interventions had a small to moderate effect for non-typical and typical children, respectively. At this stage of literacy development, children may benefit from an integration of reading and spelling instruction, in order to maximize their ability to make connections among words. For instance, practice in spelling words using word building exercises (McCandliss, Beck, Sandak, & Perfetti, 2003) may allow children to rehearse their knowledge of sound-letter correspondences, while continued experiences with print may expand their mental repertoire of orthographic patterns. In addition, practice in spelling words may also benefit children's vocabulary development (Ehri & Rosenthal, 2010).

Implications for Proficient Readers

The need for research on spelling is perhaps most evident in addressing this age group. Based on our retrieved sample of studies, only 17 studies addressed spelling in children past grade four, 12 of which sampled students with reading disabilities. Of these 12 studies, only five were conducted in the US. Surprisingly, only four studies on typical students were found. Of these four studies, one was an unpublished dissertation, and two were conducted in the 1980s. Even accounting for retrieval bias (i.e., the possibility of missing eligible studies), and other studies that may have addressed spelling but were ineligible in this meta-analysis (such as correlational studies), the distribution of studies in our sample suggests that spelling ability in this age group only tends to be addressed concomitantly with delays or disabilities in reading ability. This lack of research may certainly stem, at least partially, from the assumption that spelling is only a concern for younger children who are developing literacy skills.

But spelling is arguably no less important for older students. Much of the curriculum for this age group involves writing, and good spelling is undoubtedly a prerequisite for good writing. In fact, research has shown that poor spelling can be a significant hindrance to writing, as the act of pausing over spellings of words can interrupt the thought process involved in conveying ideas (Berninger, 2000; Graham, 2001). Longitudinal studies also show that spelling ability is linked with later writing ability. Thus, interventions that involve morphological and etymological instruction may enhance students' command of the English language and give them new linguistic insight. Although more research in this area is needed, it seems reasonable to posit that orthographic and morphological knowledge may be a beneficial supplement to these students' language arts curriculum.

There is also budding evidence that, for this age group, interventions that benefit spelling also impact other important skills, such as reading comprehension and vocabulary development (Moats, 2005). In an unpublished dissertation study on English language learners aged fifteen to seventeen, students who received morphological instruction that incorporated knowledge of etymological word roots and English affixes demonstrated greater achievement in spelling, vocabulary, and reading comprehension than those who received traditional classroom spelling instruction (Diaz, 2010). Another relatively large study done with both English language learners and those who spoke English as their first language in sixth grade showed that, in both samples, morphological awareness was significantly correlated with reading and spelling skills, even after controlling for phonological awareness an oral language skills (Siegel, 2008). As a supplement to existing language arts curricula, then, morphological instruction may not only improve students' spelling, but may also enhance other skills such as reading comprehension and writing.

It is important to emphasize that spelling instruction is typically not incorporated into the language arts curriculum in US classrooms, and sometimes eschewed completely (Graham, 1999, 2000; Fresch, 2003). As such, whether or not they need it, students are generally not receiving explicit instruction in spelling words past Grade four. However, many of these interventions that specifically targeted spelling were implemented with older students such as fifth and sixth graders proved to be effective. Thus, although spelling instruction is not commonly associated with older students, there is growing empirical evidence that explicit spelling instruction may, in fact, positively impact students' literacy development.

General Discussion and Future Directions

According to the most recent assessments conducted by the National Association of Educational Progress, 67% of fourth graders and 76% of eighth graders performed at or above basic proficiency in reading achievement (National Center for Education Statistics, 2011). These statistics are less encouraging than they may initially seem, given that the NAEP defines basic proficiency as "partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade." In other words, two-thirds of fourth and eighth graders may not actually be proficient in grade level work.

It seems reasonable to posit that a diligent consideration of spelling issues would help to inform the situation, as without remediation difficulties in reading and spelling are likely to persist and have long-term consequences for literacy development (Maughan, et al., 2009). Efforts at reading remediation may also be bolstered upon consideration of the increasing evidence that spelling benefits reading (Cataldo & Ellis, 1988; Ehri & Wilce, 1987). This is particularly true for younger children, for whom phoneme segmentation and encoding-related activities have a significant impact on reading skills (Uhry & Shepherd, 1993). For developing

readers in later primary grades, spelling ability may be an indicator of an undetected weakness in reading or phonological skills.

Spelling ability also plays an important role in the development of writing ability, an increasingly crucial skill in today's society. Although the notion of spelling has generally been divorced from categories of knowledge associated with general language abilities, it is not hard to imagine that someone who has difficulty spelling words would also have difficulty writing. In fact, research has shown that poor spelling hinders the ability to write by interfering with the transcription process (Berninger, 1999; Graham & Harris, 2000; Singer & Bashir, 2004). Longitudinal studies have shown that spelling ability in grade school is moderately correlated with later writing ability (Abbott, 2010; Juel, 1988). In fact, Abbott's (2010) longitudinal study of children in grades one to seven showed that the correlation between spelling and writing was comparable to the correlation between reading and writing. Thus, research suggests that when spelling ability is ignored in children, their future writing ability may be at risk.

Although further research on spelling is needed for all age groups, there is a pronounced need for research in spelling development for older students in 4th grade and beyond. Relative to younger children, much less is known about how spelling ability develops in older children and adults. As would be expected, there is some evidence that phonological knowledge loses predictive power for reading and spelling as children progress from early to later primary grades (Mehta et al., 2005). In this vein, it is possible that the process of reading and spelling in older children may be qualitatively different from the process of reading and spelling acquisition in younger children.

For instance, Fink's (1998) interviews of successful adults with dyslexia showed that, despite their dyslexia, many of these adults managed to attain high levels of educational and

occupational achievement by being avid readers in an area in which they were highly interested. Results from another study comparing adult literacy students and reading-level-matched children suggest that adults' may employ different strategies to read and spell words than children (Greenberg, Ehri, & Perin, 2002). In this study, although both groups of participants had trouble spelling similar words, the adults tended to rely more on orthographic strategies in spelling words, whereas the younger students relied more on phonological strategies. These results suggest that the exposure to print and experiences with reading may be an important variable; over time, these experiences may impact the strategies that proficient readers use to spell unknown words. In other words, reading may produce a feedback loop of benefits, or "Matthew effects," such that the more one reads, the more one's lexical and linguistic repertoire is bolstered, and vice versa (Stanovich, 1986). Thus, a thorough investigation of how spelling development interacts with print exposure and language experiences is needed.

Measuring spelling. In order to know more about which instructional methods are most effective in improving spelling ability, more research is needed on how to measure spelling development for multiple age groups. One method that may be particularly suited for this is quantitative and qualitative error analysis, which is more sensitive than all-or-none scoring methods to detect deficiencies at multiple levels, including phonological, orthographic, and morphological. Spelling development inventories that are designed to capture spelling patterns within words, rather than to simply score words based on conventional spellings should be used to better portray the full range of learners' abilities (Ganske, 1999; Tangel & Blachman, 1995).

However, error analysis is a complex task and requires considerable linguistic knowledge on the part of the researcher. Perhaps this is one reason that spelling has been left relatively unaddressed in the literature, as some researchers may not have the knowledge or resources to

undertake this complexity (Moats, 1994). In these cases, interdisciplinary collaboration among educational researchers and linguistic specialists may be especially helpful. Developmental and cognitive psychologists could also serve as helpful consultants on how spelling ability develops over time and interacts with other cognitive abilities.

Spelling in other languages. Because the results from this study were intended to generalize to English-speaking children in the US, otherwise eligible studies of children who spoke languages which were drastically different from English (e.g., Finnish, Chinese) were excluded. However, further research could address the issue of how spelling development differs across languages. There is burgeoning evidence that speakers of other alphabetic languages may learn to spell in different ways, which may be contingent on language-specific features. One of the most salient differences among different languages is the extent to which the sound-spelling and spelling-sound correspondences are consistent, or where the language falls on a spectrum of shallow to deep orthography.

Extant research suggests that there are both similarities and differences among predictors of reading and spelling across languages varying in orthographic depth. Longitudinal studies comparing predictors of non-word decoding, reading fluency, and spelling in Finnish, Greek, and English showed that predictive models for non-word decoding in Greek were similar to that in Finnish, and predictive models for spelling in Greek was similar to that in English (Georgiou, Parrila, & Papadopoulos, 2008; Georgiou, Torppa, Manolitsis, Lyytinen, & Parrila, 2012). As the authors explain, Greek and Finish both are fairly consistent in letter-to-sound correspondences, whereas English and Greek are relatively inconsistent in sound-to-letter correspondences. Thus, although languages of different orthographic depth may share common

predictors of reading and spelling, the primacy of each of these predictors may differ depending on the features of the language.

Further research is also needed in assessing the degree to which spelling development in alphabetic languages may differ from that in logographic languages, such as Chinese. Results from one study of Chinese-English bilingual kindergarten children support this claim (Leong & Rickard Liow, 2011). Half of the children in this study spoke Mandarin and half spoke English as their first language. For English L1 students, phonemic awareness significantly predicted spelling achievement at the end of six months, after accounting for age, nonverbal IQ, vocabulary, and prior spelling achievement. However, for Mandarin L1 speakers, syllable awareness, along with letter-sound awareness, was a more potent predictor of spelling achievement. These results suggest provide evidence that, even in a logographic language, letter-sound awareness still predicts spelling achievement, but that Mandarin speakers may be more attuned to syllables than English speakers.

In some ways, the overemphasis on studies of English speakers in the extant literature may be a hindrance to the advancement of research (Share, 2008). Further research in languages other than English, including those languages both deeper and more shallow orthographies, would help to elucidate existing findings, generate new hypotheses. Converging results across orthographically different languages would strengthen existing theories, whereas differential results by language would afford the opportunity to reexamine existing theories and to generate new ones.

Given the rapidly increasing number of English language learners in the United States, further research in spelling development across languages may be especially timely and help to inform educational practice. For example, Spanish is much more phonetically consistent than

English, and Spanish words are more easily parsed in syllables. As such, Fashola et al. (1996) emphasize the importance of teachers' being knowledgeable about the phonological and orthographic structure of Spanish, so that they can accurately monitor students' progress in Spanish-speaking English language learners.

The need for teacher training. Any progress in research and theoretical debate will be of little use if this progress is not substantiated in practice. As such, any endeavor to improve spelling outcomes in school-aged children will require the appropriate training of teachers and other educational professionals. Teacher training is undoubtedly an integral part of educational progress, as research clearly indicates that a primary determinant of students' academic outcomes is teachers' depth of knowledge in their subject area (Monk, 1994; Podgursky, Monroe, & Watson, 2001). To effectively teach spelling, teachers must have a sophisticated understanding of multiple aspects of English orthography, including patterns of sound-letter correspondences, morphological units, and etymology (McCutcheon & Berninger, 1999).

But, as recent national surveys have shown (Fresch, 2003, 2007; Graham, 2000), the majority of teachers do not have this knowledge. Many teachers report a lack of knowledge about the process of spelling development and about how to best develop spelling lessons (Moats, 2009; Moats & Lyon, 1996; Templeton & Morris, 1999). It is clear that teacher training in spelling instruction and linguistic knowledge is an area of much needed improvement. Given increasing demands on teachers' time and lack of school resources, changes in curricula to incorporate spelling-directed activities may only be accomplished with sufficient administrative support.

Although incorporating spelling instruction into the existing curriculum may be challenging, Moats (2005) describes ways in which spelling instruction can be incorporated into

the curriculum with just 15-20 minutes of daily activity. Further, these activities can be tailored to be developmentally appropriate for children of different ages. Younger students may benefit most from studying orthographic patterns in common Anglo-Saxon words (similar to "sight words"). For older students, who are typically learning about science and history, instruction in word etymology may be particularly useful in bolstering their vocabulary development and reading comprehension. The take-home message seems to be that all children would benefit from instruction that increases their sensitivity to different aspects of words.

Limitations of meta-analysis. While meta-analysis can be an indispensable methodology to advance the state of knowledge in this and other research areas, its use must be informed by several limitations. The first is that it is virtually impossible for a meta-analysis to be exhaustive. Not all studies can be located, and some otherwise viable studies will not be usable based on insufficient reporting of data. As mentioned previously, this limitation was addressed by performing a systematic and careful search of studies, as well as by performing sensitivity analyses designed to assess the likelihood of publication bias in the sample of studies. The issue of publication bias is not easy to assess, and some researchers argue that methods of locating unpublished studies, such as searching for dissertations and grey literature, may introduce additional bias (Ferguson & Brannick, 2012). In the present study, dissertations were included as a way to comprehensively represent the literature on spelling interventions; this was particularly relevant for finding studies on morphology interventions, which is only a recently burgeoning topic in the published literature base.

Second, the substantive knowledge gained from this meta-analysis is limited by the range of existing research. While phonological and phonemic awareness has been studied thoroughly, studies on children's awareness and application of larger lexical units are lacking. For this

reason, we could not reach more substantive conclusions based on the statistical constraints Because many more studies have been done on interventions focusing on phoneme units than on larger units, a moderator analysis at this grain size is not possible at this time. In addition, given the correlation between intervention type and participant age, an analysis of the differential effects of intervention by age is not feasible, as almost all phoneme-level interventions were implemented with younger students below grade four, and interventions with larger units were implemented with older children. Both correlational and intervention studies are needed to parse apart these confounding variables and to further investigate the differential effects of interventions by age.

A third limitation of meta-analysis is that the validity of inferences drawn from studies is contingent on the methodological rigor with which the studies themselves were conducted. The assessment of study quality is far from a straightforward task, as many researchers widely disagree on the issue (Valentine, 2009), and there is no sole indicator of a study's methodological strength. Rather, study quality can be measured with an array of different variables. For this reason, we conducted analyses on multiple moderators, including study design, type of control group, assessment of treatment fidelity, and assessment of initial group differences. Although none of these moderators significantly contributed to variance in the ESs, it is possible that this is the result of "noise" in the data resulting from correlated moderators. For example, research shows that quasi-experimental studies tend to report larger ESs than experimental studies (Bornstein & Hedges, 2009), but our results failed to confirm this. Thus, it is important to distinguish between substantive and methodological variation in ESs, particularly when analyzing moderator effects. Variance in reporting practices also constrains the types of data that can be extracted from studies.

Despite these limitations, meta-analysis is a useful methodology that allows researchers to investigate relationships in the literature in a way that would not be possible using only primary study. To the extent that the quantitative results of meta-analysis are subjected to careful interpretation, meta-analysis can be a powerful research tool.

Conclusion

In sum, this study provides a comprehensive synthesis of the state of the literature on the impact of literacy interventions on spelling achievement. This meta-analysis is especially warranted given the lack of research on spelling ability, and the lack of attention to spelling instruction in educational contexts. Based on the sample of studies analyzed in the meta-analysis, the overall effect of code-based interventions on spelling was moderate and was robust across age groups, participant literacy profiles, and types of interventions. Although more research is needed to arrive at more specific conclusions, it seems reasonable to conclude that any intervention that involves explicit code-based instruction is superior to the status quo in benefiting spelling achievement.

These findings are also contrary to the broad misconception that the process of spelling is divorced from general linguistic knowledge such as reading ability and linguistic aptitude. But to state that reading and spelling each draws upon two mutually exclusive sets of cognitive abilities would be akin to stating that arithmetic knowledge is distinct from algebraic knowledge. Just as reading interventions are often comprised of explicit instruction in phonological and other types of code-based knowledge, interventions designed to improve spelling outcomes should arguably include similar types of instruction. Although the present study may serve as a stepping stone for advancement in the areas of spelling development and spelling intervention research, it is clear that further research on the topic of spelling is needed to inform educational practice.

Indeed, sufficient consideration of the topic of spelling is long overdue, especially given the increasing push for evidence-based practice in education, among other fields.

Rather than being an isolated skill, spelling is clearly a necessary component of a body of literacy and communication skills crucial for successful integration into society and the workplace. As such, disregard for spelling instruction can only be deemed a disservice to students of all abilities. To put it succinctly, spelling matters.

Appendix A: Studies included in the Meta-Analysis

Author (s)	Country	Age	Profile	Intervention	Total n	Cohen's d	SE
Abbott, 2001	US	3rd	Typical	Spelling	16	0.98	0.53
Amendum, Vernon-Feagans & Ginsberg, 2011	US	K, 1st	At-risk	Phonics	167	0.40	0.16
Armand, Lefrancois et al., 2004	CA	1st	ELL	Phonemic awareness	162	2.45	0.21
Arnbak & Elbro, 1998	DK	4th, 5th	RD/LD	Word Analysis	60	0.34	0.26
Ball & Blachman, 1991	US	K	Typical	Phonemic awareness	59	0.86	0.27
Berninger et al. JEP, 2002	US	3rd	Typical	Spelling	48	0.74	0.30
Berninger et al., 1998	US	2nd	At-risk	Spelling	32	0.13	0.35
Bingham, Hall-Kenyon, & Culatta, 2010	US	K	At-risk	Phonics	63	0.14	0.26
Birgisdottir, Nunes, et al., 2006	UK	5th	At-risk	Spelling	153	0.26	0.16
Blachman et al. JEP, 2004	US	2nd, 3rd	At-risk	Text reading	69	0.89	0.25
Blachman, Ball et al., 1994	US	K	At-risk	Phonemic awareness	149	0.93	0.17
Bond, Ross, et al. (1st grade subsample), 1995	US	1st	Typical	Phonics	219	0.31	0.24
Bond, Ross, et al. (2nd grade subsample), 1995	US	2nd	Typical	Phonics	236	0.18	0.13
Bowyer-Crane, Snowling, et al., 2008	UK	Pre-K	At-risk	Phonics	143	0.50	0.17
Brennan & Ireson, 1997	UK	K	Typical	Early phonological awareness	25	0.75	0.42
Brooks, Miles et al., 2006	UK	6th	Typical	Other	130	0.05	0.18
Brown & Felton, 1990	US	1st	At-risk	Phonics	47	0.59	0.30
Butyniec-Thomas, Woloshyn,1997	CA	3rd	Typical	Spelling	25	0.56	0.41

Author (s)	Country	Age	Profile	Intervention	Total n	Cohen's d	SE
Castiglioni-Spalten & Ehri, 2003	US	K	At-risk	Early phonological awareness	28	0.96	0.40
Castle, Riach & Nicholson (1 of 2), 1994	NZ	K	At-risk	Phonemic awareness	30	1.81	0.44
Castle, Riach & Nicholson (2 of 2), 1994	NZ	K	At-risk	Phonemic awareness	32	1.38	0.39
Center & Freeman, 1997	AU	K	At-risk	Phonics	156	0.25	0.16
Center, Freeman, & Robertson, 2001	AU	K, 1st	Typical	Phonics	313	0.77	0.12
Christensen & Bowey, 2005	AU	2nd	Typical	Word Analysis	78	0.67	0.23
Darch & Siimpson, 1995	US	5th	RD/LD	Spelling	28	1.23	0.41
Darch, Eaves, et al., 2006	US	2nd, 3rd, 4th	RD/LD	Spelling	42	0.45	0.31
Darch, Kim, et al., 2000	US	4th	RD/LD	Spelling	30	1.76	0.43
Denton, Nimon, et al., 2010	US	1st	At-risk	Phonics	422	0.63	0.10
Denton, Wexler, et al., 2008	US	6th, 7th, 8th	RD/LD	Text reading	38	0.08	0.33
Devonshire & Fluck (2 of 2), 2010	UK	2nd, 3rd	Typical	Spelling	68	0.24	0.24
Diaz, 2010	US	10th, 11th	ELL	Word Analysis	140	1.78	0.20
Donnell, 2005	US	3rd	At-risk	Word Analysis	418	0.08	0.10
Duff, Hayiou-Thomas, & Hulme, 2012	UK	K, 1st	RD/LD	Phonemic awareness	59	0.55	0.27
Dunsmuir, Thomas, et al., 2008	UK	3rd	RD/LD	Spelling	33	0.10	0.35
Foorman, Fletcher et al., 1998	US	1st, 2nd	At-risk	Phonics	119	0.39	0.19
Frost & Sorensen, 2007	NO	3rd	At-risk	Text reading	72	0.52	0.24
Fuchs et al., 2001	US	K	Mixed	Phonemic awareness	268	0.58	0.12
Gettinger, et al., 1982	US	3rd, 4th, 5th	RD/LD	Spelling	39	0.92	0.35
Good, 2011	US	3rd	RD/LD	Spelling	16	0.52	0.51

Author (s)	Country	Age	Profile	Intervention	Total n	Cohen's d	SE
Graham, Harris & Chorzempa, 2002	US	2nd	At-risk	Spelling	54	0.79	0.28
Gustafson, Ferreira, & Ronnberg, 2007	SE	4th, 5th	RD/LD	Word Analysis	59	0.16	0.28
Hatcher, Hulme, & Ellis, 1994	UK	2nd	At-risk	Phonemic awareness	63	0.32	0.25
Hecht & Close, 2002	US	K	At-risk	Phonemic awareness	76	1.20	0.25
Hempenstall, 2008	AU	3rd to 8th	RD/LD	Phonics	206	0.25	0.15
Hilte & Reitsma, 2011	NL	2nd	Typical	Spelling	259	0.58	0.16
Ise & Korne, 2010	DE	5th and 6th	RD/LD	Spelling	37	0.36	0.33
Iversen & Tunmer, 1993	NZ	1st	At-risk	Phonics	64	0.25	0.25
Joseph, 2000	US	1st	Typical	Phonemic awareness	48	0.00	0.29
Kent, 2006	US	1st	Typical	Other	224	0.29	0.13
Kernaghan & Woloshyn, 1995	CA	1st	Typical	Spelling	24	1.01	0.44
Kirk & Gillon, 2009	NZ	3rd to 6th	At-risk	Word Analysis	16	1.14	0.54
Kwan (ELL subsample), 2005	US	K	ELL	Phonics	76	0.51	0.24
Kwan (Typical subsample), 2005	CA	K	Typical	Phonics	92	0.88	0.22
Lie, 1991	NO	1st	Typical	Early phonological awareness	102	0.67	0.20
Lovett, Lacerenza et al., 2000	US	3rd to 9th	RD/LD	Word Analysis	34	0.14	0.34
Lovett, Warren-Chaplin et al., 1990	CA	2nd to 8th	RD/LD	Word Analysis	36	0.22	0.33
Lum & Morton, 1984	CA	2nd	Typical	Spelling	18	0.54	0.50
Martinussen & Kirby, 1998	US	K	At-risk	Phonemic awareness	27	0.49	0.39
Nag-Arulmani, Reddy, & Buckley, 2003	IN	3rd	ELL	Early phonological awareness	61	0.39	0.26
O'Connor & Padeliadu, 2000	US	1st	At-risk	Phonemic awareness	12	0.09	0.58

Author (s)	Country	Age	Profile	Intervention	Total n	Cohen's d	SE
Ouellette & Senechal, 2008	CA	K	Typical	Spelling	46	0.80	0.31
Piasta, Purpura, & Wagner, 2010	US	Pre-K	Typical	Phonemic awareness	39	0.23	0.32
Pittman, 2007	US	6th	Typical	Spelling	142	1.08	0.18
Post & Carreker, 2002	US	2nd, 3rd, and 4th	Typical	Spelling	128	0.18	0.18
Rieben, Ntamakiliro, et al., 2005	CH	K	Typical	Spelling	73	0.46	0.24
Robinson & Hesse, 1981	US	7th	Typical	Spelling	172	1.23	0.21
Robinson, 2010	US	K	Typical	Phonemic awareness	89	0.69	0.22
Roth & Guinee, 2011	US	1st	At-risk	Other	101	0.38	0.20
Rubin & Eberhardt, 1996	US	K	Typical	Spelling	912	0.14	0.07
Saint-Laurent & Giasson, 2001	CA	K	At-risk	Phonemic awareness	61	0.47	0.26
Santa & Hoien, 1999	US	1st	At-risk	Phonics	49	0.61	0.29
Santoro, Coyne, & Simmons, 2006	US	K	At-risk	Spelling	77	1.16	0.25
Savage, Carless, & Stuart, 2003	UK	1st	At-risk	Word Analysis	54	0.32	0.27
Schuele, Justice, et al. (Typical sub-sample), 2008	US	K	Mixed	Phonemic awareness	77	0.30	0.23
Senechal, Oulette, et al., 2012	CA	K	At-risk	Spelling	37	0.76	0.34
Shepherd, 2011	US	K	Typical	Other	26	0.30	0.39
Simmons, Coyne, et al., 2011	US	K	Mixed	Phonics	206	0.34	0.14
Slattery-Gursky, 2003	US	K	Mixed	Phonemic awareness	16	0.38	0.50
Stephens & Hudson (Remedial subsample), 1984	AU	7th	RD/LD	Spelling	22	0.94	0.48
Stephens & Hudson (Typical subsample), 1984	US	7th	Typical	Spelling	50	1.19	0.32

Author (s)	Country	Age	Profile	Intervention	Total n	Cohen's d	SE
Sussman, 1998	US	1st	Typical	Spelling	42	0.64	0.32
Torgesen, Wagner, et al., 1999	US	K	At-risk	Phonics	70	0.66	0.25
Uhry,Joanna K.;Shepherd,Margaret Jo, 1993	US	1st	Typical	Spelling	22	0.75	0.44
Vadasy & Sanders (LM subsample), 2010	US	K	ELL	Phonics	84	0.37	0.22
Vadasy & Sanders (non-LM subsample), 2010	US	K	At-risk	Phonics	64	0.94	0.26
Vadasy, Sanders & Peyton, 2006	US	K	At-risk	Phonics	67	0.58	0.25
Vadasy, Sanders, & Tudor, 2007	US	2nd and 3rd	At-risk	Phonics	43	0.24	0.31
Vandervelden & Siegel, 1997	CA	K	At-risk	Phonemic awareness	29	1.14	0.40
Vaughn, Cirino et al. (1 of 2), 2006	US	1st	ELL	Phonics	90	0.33	0.21
Vaughn, Cirino et al., 2010	US	6th	At-risk	Text reading	300	0.22	0.12
Vaughn, Mathes et al., 2006	US	1st	ELL	Phonics	39	0.78	0.33

*Note. C*ountry abbreviations, AU = Australia, CA = Canada, CH = Switzerland, DE = Germany, DK = Denmark, IL = Israel, NL = IN = India, NL = Netherlands, NO = Norway, NZ = New Zealand, SE = Sweden, UK = United Kingdom.

Appendix B: Coding Manual

Study-Level Codes

Study Characteristics

1) Country
    1. Entered value

    ** (1) Country in which study takes place.

2) Design
    1. Experimental
    2. Quasi-experimental
    3. Retrospective study

    ** (1) *Participants* were randomly assigned to conditions. A large-scale randomized control trial in which multiple classrooms from multiple schools are randomly assigned to conditions would also be experimental. (2) Participants were assigned to conditions non-randomly, or no random assignment was mentioned. Or, groups are intact and therefore are considered non-equivalent. This includes any study that involves at least one intact classroom or group (even if they are randomly assigned). (3) In a retrospective study, the treated sample is compared to a non-treated sample, based on existing retrospective data after the intervention has already been carried out.

3) Group Assignment
    1. Participants
    2. Classes
    3. Schools
    4. None

    ** Indicates unit of random assignment to conditions. (1) Participants. (2) Classes. (3) Schools. (4) Groups are intact, no random assignment of any unit is mentioned, or conditions are pre-existing (such as a classroom already implementing an intervention).

Study Quality

4) Existence of initial group differences
    1. None found
    2. Present and statistically adjusted
    3. Present and not statistically adjusted

    ** (1) Equivalence of groups on baseline variables relevant to the outcome was statistically analyzed, and no significant differences were found. (2) Significant group differences were found on at least one pretest variable, and this was statistically

adjusted for in the analysis. (3) Significant group differences at baseline were reported, but this was not accounted for in the analysis.

5) Effort was made at assessing treatment fidelity
    1. Yes
    2. No
    3. Not reported

    ** (1) Researchers reported assessment of treatment fidelity by directly observing the intervention. Observation includes both classroom visits (either by researchers or trained personnel) and/or review of video/audiotaped lessons. (2) Researchers employed some measures of assessing treatment fidelity (such as submitting instructor logs, receiving ongoing consultation, etc.), but did not directly observe the intervention. (3) The article does not report assessment of treatment fidelity.

<u>Participant</u>

6) Language
    1. Entered value

    ** (1) The primary language of the sample. For English language or bilingual learners, this is the language of their core reading instruction.

7) SES
    1. Low
    2. Middle
    3. Mixed
    4. Not reported

    ** (1) Over half of the sample is either explicitly described as "low" in socioeconomic status, or by other relevant characteristics (e.g., living in housing projects, eligibility for free lunch, etc.). (2) The study reports SES by using the word "middle" (e.g., lower-middle class, upper-middle class). (3) The sample is heterogeneous in SES (e.g., a sample comprised of 70% middle-class and 30% low-income students). Or, if the study does not report SES details but takes place in an urban or suburban area, this would be "Mixed."

8) Literacy Profile
    1. Typical
    2. At-risk
    3. RD/LD
    4. ELL
    5. Mixed

    ** (1) Participants have no documented deficits or are described as typically achieving. No screening protocols are mentioned. (2) Participants are defined as at-

risk for future reading difficulties if their reading/spelling abilities are below average compared with their general or verbal IQ. This status is typically given to younger children in kindergarten, 1st, or 2nd grade. If the sample is predominantly low-income, they are at-risk. (3) Participants' literacy difficulties persist in spite of their relatively typical cognitive/intellectual levels. These participants are usually older students in 3rd grade and beyond who may have specific language or reading impairments. They may also have been screened as poor readers who have average cognitive scores. (4) ELL indicates that participants are second language learners. If the majority of the sample is both ELL and at-risk (as many are), choose ELL, the more salient feature. (5) Choose this option only when 2 or more disparate groups are included in the sample, such as Typical and At-risk. The only time this applies is when a study fails to report separate sub-sample data, and therefore the sample included in the comparison includes both sub-samples.

Intervention

9) Intervention Type
    1. Early Phonological Awareness
    2. Phonemic Awareness
    3. Comprehensive Phonics
    4. Comprehensive Text Reading
    5. Word analysis (Decoding)
    6. Word analysis (Spelling)
    7. Other

** All interventions provide systematic and explicit instruction in lexical (word-level) and/or sub-lexical (within-word-level) units.

(1) Both phonemic and phonological awareness are defined by the child's ability to be aware of and manipulate sound units in words. Phonological awareness is the broader term and encompasses phonemic awareness, which is defined by awareness at the level of phonemes, the smallest unit of spoken language. While children as young as preschool-aged may be taught phonological skills such as clapping syllables and rhyming, phonemic awareness is typically develops later. Common phonemic awareness activities involve manipulating phonemes by deletion, addition, substitution, and segmenting of sounds. "Print" refers to interventions that expose children to letters/word and explicitly connect sounds to letters.

Example: Sound Foundations (Byrne & Fielding-Barnsley, 1993). Program focuses on phoneme invariance by teaching kids that different words can begin or end with the same sound (phoneme identity) and phoneme segmentation. 5 consonants (s, m, l, t, p) and one vowel /ae/ in initial position only.

(2) An early phonological awareness intervention may consist of similar activities as above, but instruction is primarily oral and letter-sound correspondences are not explicitly taught. For instance, in a strictly phonological awareness intervention,

children may be clapping out syllables, or use blocks to represent phonemes in spoken words.

Example: In the say-it-and-move it activity, children move tiles below rectangles representing phonemes. In some interventions, these tiles may have letters on them. In an oral intervention, the tiles may be different colors, in which case the children are attending to phonemes (sounds) without reference to their corresponding letters.

(3) Comprehensive phonics interventions are typically multicomponent programs designed to teach children to read, based on explicit instruction in phonemic and phonological awareness (as above). However, along with the latter, phonics programs will usually include other reading-related instruction such as sight words, vocabulary, story reading, etc.

Example: Corrective Reading: Decoding (Hempenstall, 2008). Synthetic phonics-emphasis Direct Instruction remedial reading program. 2 major features: decoding (phonics) and the Direct Instruction approach to teaching phonics content. Lessons are scripted and use choral responses prompted by teacher. Level A program focuses attention on word structure through reviewing LSC, regular rhyming, blending and segmenting activities. Then irregular word are introduced. Level B includes word-attack skills, practicing pronouncing words, identifying sounds of letters or letter combinations, and reading isolated words composed of sounds and sound combinations. Also did story-reading, individual reading, and workbook activities.

(4) A comprehensive text reading intervention is distinguished from phonics by its focus on reading fluency and comprehension as broad goals. These interventions are typically carried out with children in Grade 3 and above. Although this may involve some review of phonics, the instruction moves beyond word decoding and includes reading text.

Example: Responsive Reading (Denton, Nimon, et al. 2010). Direct, explicit instruction in phonics skills and text-reading strategies, as well as modeling and instructional scaffolding as students apply these skills and strategies while reading and writing connected text. Activities include word word phonemic awareness, LSC, sight words, phonemic decoding, and spelling), print concepts/fluency and assessment, supported reading, and supported writing.

(5) Orthographic word study involves systematic methods of decoding and spelling using larger sub-lexical units such as onsets, rimes, and syllables (e.g., in the word cartoon, the syllable is /car/, and in this syllable, the onset is /c/, and the rime is /ar/). This is distinct from pure whole word strategies, in which students simply learn words by exposure or repetition without explicit application of sub-lexical analysis. Students may also learn common spelling units such as the /ai/ long vowel sound in *drain* and *stain*.

Example: Structural Analysis (Abbot & Berninger, 1999). Both conditions included instruction in phonological and orthographic skills, alphabetic principle training, phonological decoding, and oral reading of connected text for meaning

(comprehension monitoring, not explicit instruction). However, this condition emphasized analysis at the syllable/morpheme level, with some attention to affixes. For decoding, kids were taught to divide word into syllables, sound out the spelling-sound correspondences within syllables, then blend syllables to make a whole word. They were taught some etymological strategies for decoding, reading, and spelling long words. Encouraged to use letter-sound correspondences only after attempting the morpheme and syllable strategies.

(6) Orthographic word study with spelling may include any of the above literacy activities, but the goal of instruction is specifically to improve spelling. The content of some word analysis interventions may overlap with spelling interventions. However, the distinguishing feature of spelling interventions is that students spend a significant amount of time writing correct spellings of words, as opposed to using spelling strategies to decode.

Example: Spelling (Graham, Harris, & Chorzempa, 2000). Students' lexical knowledge was strengthened by teaching them how to spell correctly words that frequently occur in the writing of primary-grade children. Their knowledge of the spelling system was enhanced by teaching them common sound-letter combinations (e.g., consonants, blends, etc.), spelling patterns or rules involving long and short vowels (e.g., the doubling rule for adding a suffix), and frequently occurring phonograms or rimes (e.g., ig, at, etc.). With the word-sorting, word-building, and peer-practice activities used to teach these skills, both accuracy and fluency were stressed. Verifying the correctness of their spelling was emphasized as students learned how to spell frequently occurring words (i.e., misspellings made during study practice were corrected) and when building words using onsets and rimes (i.e., they examined each built word to determine whether it was a real or made-up word).

(7) Other intervention that does not fit any of the above descriptions.

10) Type of comparison group

1. No contact
2. Contact control, literacy
3. Contact control, language
4. Contact control, non-language

** (1) No treatment was given, and researchers made no contact with this group. Options 2-5 indicate that researchers either implemented a comparison condition or influenced instruction in some way, even if it is just by instructing teachers to increase the intensity of their teaching during the intervention period. (2) Often a "diluted" or less systematic/explicit version of the intervention. Some studies may describe the condition as "limited" or "implicit" version of the intervention. (3) Students were given a general language arts activity that does not specifically pertain to literacy. An example of this would be students who were told to talk about their experiences with spelling. (4) This category refers to non-literacy and non-language activities, such as mathematics or general academic strategies.

11) Follow-Up Type
    1. Post-test
    2. Short-term follow-up
    3. Long-term follow-up.

    ** (1) Immediate post-test within 1 month of the end of intervention. This category also included multi-year interventions which reported results at the end of the first year. (2) Delayed follow-up within one year of the end of intervention. (3) Delayed follow-up more than one year after the end of intervention.

Outcome

12) Outcome Measure Type
    1. Researcher-constructed
    2. Standardized

13) Outcome Word Type
    1. Real words
    2. Real words- regular only
    3. Real words- irregular only
    4. Pseudo-words only
    5. Mixed- Both real and pseudowords

    ** (1) The measure uses real words which may include both regular and irregular words. Typically most grade-representative words used on spelling measures (both standardized and researcher-constructed) include irregular words, as these are high-frequency. (2) Only phonetically predictable real words were used (e.g., cat, fig). (3) Only phonetically irregular real words were used (e.g., sight words like "are" and "the" and/or more words with inconsistent grapheme-phoneme correspondence, like "rough" and "right"). (4) Only pseudowords were used (e.g., nuf, mab, stup).

14) Scoring criteria
    1. Correct spelling
    2. Developmental criteria

    ** (1) Each correctly spelled word is given 1 point. (2) Partial credit is given based on phonetically or orthographically rated spellings (e.g., "pensil").

Intervention Instructional Details

15) Duration
    1. Less than 2 weeks
    2. Less than 3 months
    3. 3 to 6 months
    4. Year
    5. Multi-year

** The range of time during which intervention took place. All interventions except those coded as "lab" are field-based and took place on a regular basis, varying from 2 times per week to daily, with sessions ranging from 10 minutes to 45 minutes. (1) Less than 2 weeks. These studies are typically conducted in a non-field-based setting and involve training of a specific skill in a short period of time, typically anywhere from one day to two weeks. (2) Less than 3 months. (3) Three to 6 months, inclusive. (4) One academic year. Because the exact time of pre- and post-testing varies due to logistical factors in field research, the actual duration of a year-long intervention may range from 6.5 months to 8 months. (5) More than one academic year.

16) Instructor
    1. Regular instructor
    2. Other
    3. Computer

** (1) Any instructor who is in regular contact with the children, such as the classroom teacher, teacher aides, or remedial teachers. Choose this only if it is clear that all children were taught by their regular teacher. (2) Any other instructors who do not regularly teach the children, including school personnel, certified teachers hired by the researchers, and/or researchers. In some studies that include multiple classrooms within a school, a group of school staff which includes classroom teachers may be selected to teach the intervention. (3) Children received the intervention primarily via interaction with a computer, though they may have been supervised by an adult.

17) Instructional Unit
    1. Tutoring
    2. Small group
    3. Whole-class
    4. Varied

** All choices refer to the primary mode in which children received *instruction*. For example, if the children were taught the intervention primarily as a class and then worked in small groups or individually with the teacher at some points throughout the lesson, this would be considered "whole-class." If children were instructed in different instructional units in different sessions equal amounts of time (e.g., tutoring one day per week and small groups one day of the week), choose the smallest unit. (1) Children received the intervention primarily on a one-on-one basis. (2) Children received the intervention primarily in small groups from between 2 to 10 children. (3) Children received the intervention primarily through whole-class instruction, or in a group larger than 10. (4) Choose this option only if children received instruction in all group sizes (including tutoring, small group, and whole class.

Appendix C: References for Studies Included in Meta-Analysis

Abbott, M. (2001). Effects of traditional versus extended word-study spelling instruction on students' orthographic knowledge. *Reading Online, 5*(3).

Amendum, S. J., Vernon-Feagans, L., & Ginsberg, M. C. (2011). The Effectiveness of a Technologically Facilitated Classroom-Based Early Reading Intervention: The Targeted Reading Intervention. *Elementary School Journal*, 112(1), 107-131.

Armand, F., Lefrançois, P., Baron, A., Gomez, M., & Nuckle, S. (2004). Improving reading and writing learning in underprivileged pluri-ethnic settings. *British Journal of Educational Psychology, 74*(3), 437-459.

Arnbak, E., & Elbro, C. (2000). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research, 44*(3), 229-251.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26*(1), 49-66.

Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., . . . Graham, S. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94*(2), 291-304.

Berninger, V., Abbott, R., Rogan, L., Reed, E., Abbott, S., Brooks, A., & Others, A. (1998). Teaching spelling to children with specific learning disabilities: The mind's ear and eye beat the computer or pencil. *Learning Disability Quarterly, 21*(2), 106-22.

Bingham, G. E., Hall-Kenyon, K., & Culatta, B. (2010). Systematic and engaging early literacy: Examining the effects of paraeducator implemented early literacy instruction. *Communication Disorders Quarterly, 32*(1), 38-49.

Birgisdottir, F., Nunes, T., Pretzlik, U., Burman, D., Gardner, S., & Bell, D. (2006). An intervention program for teaching children about morphemes in the classroom. In T. Nunes, P. Bryant, U. Pretzlik, & J. Hurry (Eds.), *Improving literacy by teaching morphemes (*pp. 104-120). London: Routledge.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., Shaywitz, B. A., & Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology, 96*(3), 444-461.

Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing: An Interdisciplinary Journal, 11*(3), 239-73.

Bond, C. E., Ross, S. M., Smith, L. J., & Nunnery, J. A. (1995). The effects of the sing, spell, read and write program on reading achievement of beginning readers. *Reading Research and Instruction, 35*(2), 122-141.

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., & ... Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal Of Child Psychology And Psychiatry,* 49(4), 422-432.

Brennan, F., & Ireson, J. (1997). Training phonological awareness: A study to evaluate the effects of a program of metalinguistic games in kindergarten. *Reading and Writing: An Interdisciplinary Journal, 9*(4), 241-63.

Brooks, G., Miles, J. N. V., Torgerson, C. J., & Torgerson, D. J. (2006). Is an intervention using computer software effective in literacy learning? A randomised controlled trial. *Educational Studies, 32*(2), 133-143.

Brown, I. S., & Felton, R. H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. *Reading and Writing, 2*(3), 223-241.

Butyniec-Thomas, J., & Woloshyn, V. E. (1997). The effects of explicit-strategy and whole-language instruction on students' spelling ability. *Journal of Experimental Education, 65*(4), 293-302.

Byrne, B., & F. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology, 85*, 104-111.

Byrne, B., & F. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology, 87*, 488-503.

Castiglioni-Spalten, M., & Ehri, L. C. (2003). Phonemic awareness instruction: Contribution of articulatory segmentation to novice beginners' reading and spelling. *Scientific Studies of Reading, 7*(1), 25-52.

Castle, J. M., Riach, J., & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language program. *Journal of Educational Psychology, 86*(3), 350-359.

Cates, W. M., & Goodling, S. C. (1997). The relative effectiveness of learning options in multimedia computer-based fifth-grade spelling instruction. *Educational Technology Research and Development, 45*(2), 27-46.

Center, Y., & Freeman, L. (1997). A trial evaluation of SWELL (schoolwide early language and literacy): A whole class early literacy program for at-risk and disadvantaged children. *International Journal of Disability, Development and Education, 44*(1), 21-39.

Center, Y., Freeman, L., & Robertson, G. (1998). An evaluation of schoolwide early language and literacy (SWELL) in six disadvantaged schools. *International Journal of Disability, Development and Education, 45*(2), 143-172

Center, Y., Freeman, L., & Robertson, G. (2001). The relative effect of a code-oriented and a meaning-oriented early literacy program on regular and low progress australian students in year 1 classrooms which implement reading recovery. *International Journal of Disability, Development and Education, 48*(2), 207-232.

Christensen, C. A., & Bowey, J. A. (2005). The efficacy of orthographic rime, grapheme-phoneme correspondence, and implicit phonics approaches to teaching decoding skills. *Scientific Studies of Reading, 9*(4), 327-349.

Cirino, P. T., Vaughn, S., Linan-Thompson, S., Cardenas-Hagan, E., Fletcher, J. M., & Francis, D. J. (2009). One-year follow-up outcomes of spanish and english interventions for english language learners at risk for reading problems. *American Educational Research Journal, 46*(3), 744-781.

Clarke, L. K. (1988). Invented versus traditional spelling in first graders' writings: Effects on learning to spell and read. *Research in the Teaching of English, 22*(3), 281-309.

Cunningham, A. J., & Carroll, J. M. (2011). The development of early literacy in steiner- and standard-educated children. *British Journal of Educational Psychology, 81*(3), 475-490.

Darch, C., Eaves, R. C., Crowe, D. A., Simmons, K., & Conniff, A. (2006). Teaching spelling to students with learning disabilities: A comparison of rule-based strategies versus traditional instruction. *Journal of Direct Instruction, 6*(1), 1-16.

Darch, C., Kim, S., Johnson, S., & James, H. (2000). The strategic spelling skills of students with learning disabilities: The results of two studies. *Journal of Instructional Psychology, 27*(1), 15-26.

Darch, C., & Simpson, R. G. (1990). Effectiveness of visual imagery versus rule-based strategies in teaching spelling to learning disabled students. *Research in Rural Education, 7*(1), 61-70.

Denton, C. A., Kethley, C., Nimon, K., Kurz, T. B., Mathes, P. G., Minyi, S., & Swanson, E. A. (2010). Effectiveness of a Supplemental Early Reading Intervention Scaled Up in Multiple Schools. Exceptional Children, 76(4), 394-416.

Denton, C. A., Wexler, J., Vaughn, S., & Bryan, D. (2008). Intervention provided to linguistically diverse middle school students with severe reading difficulties. *Learning Disabilities Research & Practice, 23*(2), 79-89.

Devonshire, V., & Fluck, M. (2010). Spelling development: Fine-tuning strategy-use and capitalising on the connections between words. *Learning and Instruction, 20*(5), 361-371.

Diaz, I. (2010). *The effect of morphological instruction in improving the spelling, vocabulary, and reading comprehension of high school english language learners (ELLs).* TUI University). *ProQuest Dissertations and Theses,* 106. (250906858)

Donnell, W. J. (2005). *The effects of multisensory vowel instruction during word study for third grade students.* University of Missouri - Kansas City). *ProQuest Dissertations and Theses,* 238-238 p. (305458227).

Duff, F. J., Hayiou-Thomas, M., & Hulme, C. (2012). Evaluating the effectiveness of a phonologically based reading intervention for struggling readers with varying language profiles. *Reading and Writing, 25*(3), 621-640.

Dunsmuir, S., Thomas, C., May, R., Monroe, J., Roiter, T., & Wellman, S. (2008). Developing an intervention for pupils with writing difficulties: Conceptualisation and analysis. *Educational and Child Psychology, 25*(3), 150-164.

Durgunoğlu, A. Y., & Öney, B. (2002). Phonological awareness in literacy acquisition: It's not only for children. *Scientific Studies of Reading, 6*(3), 245-266.

Ecalle, J., Magnan, A., & Calmus, C. (2009). Lasting effects on literacy skills with a computer-assisted learning using syllabic units in low-progress readers. *Computers & Education, 52*(3), 554-561.

Everatt, J., Al-Sharhan, A., Al-Azmi, Y., Al-Menaye, N., & Elbeheri, G. (2011). Behavioural/attentional problems and literacy learning difficulties in children from non-English language/cultural backgrounds. *Support for Learning, 26*(3), 127-133.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37-55.

Foorman, B. R., Chen, D., Carlson, C., Moats, L., Francis, K. D., & Fletcher, J. M. (2003). The necessity of the alphabetic principle to phonemic awareness instruction. *Reading and Writing, 16*(4), 289-324.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology, 31*(1), 1-29.

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.

Frost, J., & Sorensen, P. M. (2007). The effects of a comprehensive reading intervention programme for grade 3 children. *Journal of Research in Reading, 30*(3), 270-286.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., & ... O'Connor, R. E. (2001). Is Reading Important in Reading-Readiness Programs? A Randomized Field Trial with Teachers as Program Implementers. *Journal Of Educational Psychology, 93*(2), 251-67.

Gettinger, M. (1993). Effects of error correction on third graders' spelling. *The Journal of Educational Research, 87*(1), 39-45.

Gettinger, M., & Others, A. (1982). Designing spelling instruction for learning-disabled children: An emphasis on unit size, distributed practice, and training for transfer. *Journal of Special Education, 16*(4), 439-48.

Gillon, G. T. (2005). Facilitating phoneme awareness development in 3- and 4-year-old children with speech impairment. *Language, Speech, and Hearing Services in Schools, 36*(4), 308-324.

Good, J. E. (2011). *The effects of morphological awareness training on reading, spelling, and vocabulary skills.* (University of Arkansas for Medical Sciences). *ProQuest Dissertations and Theses* 176. (897106256).

Graham, S., Harris, K. R., & Chorzempa, B. F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*(4), 669-686.

Gustafson, S., Ferreira, J., & Rönnberg, J. (2007). Phonological or orthographic training for children with phonological or orthographic decoding deficits. *Dyslexia: An International Journal Of Research And Practice, 13*(3), 211-228.

Hatcher, P. J., & And, O. (1994). Ameliorating Early Reading Failure by Integrating the Teaching of Reading and Phonological Skills: The Phonological Linkage Hypothesis. *Child Development, 65*(1), 41-57.

Hecht, S. A., & Close, L. (2002). Emergent literacy skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. *Journal of Experimental Child Psychology, 82*(2), 93-115.

Hempenstall, K. (2008). Corrective Reading: An Evidence-Based Remedial Reading Intervention. *Australasian Journal Of Special Education, 32*(1), 23-54.

Hilte, M., & Reitsma, P. (2011). Activating the meaning of a word facilitates the integration of orthography: Evidence from spelling exercises in beginning spellers. *Journal of Research in Reading, 34*(3), 333-345.

Hurry, J., & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading, 30*(3), 227-248.

Ise, E., & Schulte-Korne, G. (2010). Spelling deficits in dyslexia: Evaluation of an orthographic spelling training. *Annals of Dyslexia, 60*(1), 18-39.

Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery Program. *Journal Of Educational Psychology, 85*(1), 112-126.

Joseph, L. M. (2000). Using word boxes as a large group phonics approach in a first grade classroom. *Reading Horizons, 41*(2), 117-27.

Kent, M. A. (2006). *The effect of systematic word wall instruction on the literacy achievement of first grade students.* (University of Houston). *ProQuest Dissertations and Theses*, 271-271. (305323441).

Kernaghan, K., & Woloshyn, V. E. (1995). Providing grade one students with multiple spelling strategies: Comparisons between strategy instruction, strategy instruction with metacognitive information, and traditional language arts. *Applied Cognitive Psychology, 9*(2), 157-166.

Kirk, C., & Gillon, G. T. (2007). Longitudinal effects of phonological awareness intervention on morphological awareness in children with speech impairment. *Language, Speech & Hearing Services in Schools, 38*(4), 342-352.

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech & Hearing Services in Schools, 40*(3), 341-351.

Kwan, A. B. (2005). *Impact of systematic phonics instruction on young children learning English as a second language.* University of Toronto (Canada)). *ProQuest Dissertations and Theses*, 190-190 p. (276420075).

Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children. *Reading Research Quarterly, 26*, 234-250.

Lovett, M. W., Lacerenza, L., Borden, S. L., Frijters, J. C., Steinbach, K. A., & De Palma, M. (2000). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. *Journal of Educational Psychology*, *92*(2), 263-283.

Lovett, M. W., Warren-Chaplin, P., Ransby, M. J., & Borden, S. L. (1990). Training the word recognition skills of reading disabled children: Treatment and transfer effects. *Journal of Educational Psychology, 82*(4), 769-780.

Lum, T., & Morton, L. L. (1984). Direct instruction in spelling increases gain in spelling and reading skills. *Special Education in Canada, 58*(2), 41-45.

Lundberg, I., & F. (1988). Effects of an extensive program for stimulating phonologicalawareness in preschool children. *Reading Research Quarterly, 23*, 263-284.

Lyster, S. H. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing, 15*(3-4), 261-294.

Mantzicopoulos, P., Morrison, D., Stone, E., & Setrakian, W. (1992). Use of the SEARCH/TEACH tutoring approach with middle-class students at risk for reading failure. *The Elementary School Journal, 92*(5), 573-586.

Martinussen, R. L., Kirby, J. R., & Das, J. P. (1998). Instruction in successive and phonological processing to improve the reading acquisition skills of at-risk kindergarten children. *Developmental Disabilities Bulletin, 26*(2), 19-39.

Nag-Arulmani, S., Reddy, V., & Buckley, S. (2003). Targeting phonological representations can help in the early stages of reading in a non-dominant language. *Journal of Research In Reading, 26*(1), 49-68.

Nancollis, A., Lawrie, B., & Dodd, B. (2005). Phonological awareness intervention and the acquisition of literacy skills in children from deprived social backgrounds. *Language, Speech, and Hearing Services in Schools, 36*(4), 325-335.

O'Connor, M., Arnott, W., McIntosh, B., & Dodd, B. (2009). Phonological awareness and language intervention in preschoolers from low socio-economic backgrounds: A longitudinal investigation. *British Journal of Developmental Psychology, 27*(4), 767-782.

Ouellette, G., & Sénéchal, M. (2008). Pathways to literacy: A study of invented spelling and its role in learning to read. *Child Development, 79*(4), 899-913.

Piasta, S. B., Justice, L. M., McGinty, A. S., & Kaderavek, J. N. (2012). Increasing young children's contact with print during shared reading: Longitudinal effects on literacy achievement. *Child Development, 83*(3), 810-820.

Piasta, S. B., Purpura, D. J., & Wagner, R. K. (2010). Fostering alphabet knowledge development: A comparison of two instructional approaches. *Reading and Writing: An Interdisciplinary Journal, 23*(6), 607-626.

Pittman, R. T. (2007). *Improving spelling ability among speakers of African American vernacular English: An intervention based on phonological, morphological, and*

orthographic principles. (Texas A&M University). *ProQuest Dissertations and Theses*, 169. (304727528).

Post, Y. V., & Carreker, S. (2002). Orthographic similarity and phonological transparency in spelling. *Reading and Writing, 15*(3-4), 317-340.

Rieben, L., Ntamakiliro, L., Gonthier, B., & Fayol, M. (2005). Effects of various early writing practices on reading and spelling. *Scientific Studies of Reading, 9*(2), 145-166.

Robinson, S. (2010). *The effects of embedded phonological awareness training on the reading and spelling skills of kindergarten students.* The University of North Dakota). *ProQuest Dissertations and Theses,*117. (814760105).

Robinson, J. W., & Hesse, K. D. (1981). A morphemically based spelling program's effect on spelling skills and spelling performance of seventh grade students. *Journal of Educational Research, 75*(1), 56-62.

Roth, K., & Guinee, K. (2011). Ten minutes a day: The impact of interactive writing instruction on first graders' independent writing. *Journal of Early Childhood Literacy, 11*(3), 331-361.

Rubin, H., & Eberhardt, N. C. (1996). Facilitating invented spelling through language analysis instruction: An integrated model. *Reading and Writing: An Interdisciplinary Journal, 8*(1), 27-43.

Saint-Laurent, L., & Giasson, J. (2001). Effects of a multicomponent literacy program and of supplemental phonological sessions on at-risk kindergartners. *Educational Research and Evaluation, 7*(1), 1-33.

Santoro, L. E., Coyne, M. D., & Simmons, D. C. (2006). The reading-spelling connection: Developing and evaluating a beginning spelling intervention for children at risk of reading disability. *Learning Disabilities Research & Practice, 21*(2), 122-133.

Savage, R., Carless, S., & Stuart, M. (2003). The effects of rime- and phoneme-based teaching delivered by learning support assistants. *Journal of Research in Reading, 26*(3), 211-233.

Schneider, W., & K. (1997). Short-and long-term effects of training phonological awareness in kindergarten: Evidence from two german studies. *Journal of Experimental Child Psychology, 66*, 311-340.

Schuele, C. M., Justice, L. M., Cabell, S. Q., Knighton, K., Kingery, B., & Lee, M. W. (2008). Field-based evaluation of two-tiered instruction for enhancing kindergarten phonological awareness. *Early Education and Development, 19*(5), 726-752.

Senechal, M., Ouellette, G., Pagan, S., & Lever, R. (2012). The role of invented spelling on learning to read in low-phoneme awareness kindergartners: A randomized-control-trial study. *Reading and Writing: An Interdisciplinary Journal, 25*(4), 917-934.

Shepherd, J. M. (2011). *Finger-point reading instruction using storybooks: The effects on kindergarten children's early literacy skills.* Fordham University). *ProQuest Dissertations and Theses,* 222. (884225879).

Simmons, D. C., Coyne, M. D., Hagan-Burke, S., Kwok, O., Simmons, L., Johnson, C., . . . Crevecoeur, Y. C. (2011). Effects of supplemental reading interventions in authentic contexts: A comparison of kindergarteners' response. *Exceptional Children, 77*(2), 207-228.

Slattery Gursky, R. L. (2003). *Effects of spontaneous phonemic awareness knowledge, contextualized phonemic awareness instruction, and decontextualized phonemic awareness training on phonemic awareness, reading, and spelling development in kindergarten students.* Northern Illinois University). *ProQuest Dissertations and Theses,* 252-252 p. (305315599).

Snider, V. E. (1990). Direct instruction reading with average first-graders. *Reading Improvement, 27*(2), 143-48.

Stephens, M., & Hudson, A. (1984). A comparison of the effects of direct instruction and remedial english classes on the spelling skills of secondary students. *Educational Psychology, 4*(4), 261-267.

Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology, 69*(4), 587-605.

Sussman, G. L. (1998). *The effects of phonologically constructed spelling on first graders' literacy development.* Fordham University). *ProQuest Dissertations and Theses,*141 p.(304426614).

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*(4), 579-593.

Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/Spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly, 28*(3), 218-33.

Uhry, J. K., & Shepherd, M. J. (1997). Teaching phonological recoding to young children with phonological processing deficits: The effect on sight-vocabulary acquisition. *Learning Disability Quarterly, 20*(2), 104-125.

Vadasy, P. F., & Sanders, E. A. (2010). Efficacy of supplemental phonics-based instruction for low-skilled kindergarteners in the context of language minority status and classroom phonics instruction. *Journal Of Educational Psychology*, *102*(4), 786-803.

Vadasy, P. F., & Sanders, E. A. (2012). Two-year follow-up of a kindergarten phonics intervention for English learners and native English speakers: Contextualizing treatment impacts by classroom literacy instruction. *Journal of Educational Psychology, 104*(4), 987-1005.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Code-oriented instruction for kindergarten students at risk for reading difficulties: A randomized field trial with paraeducator implementers. *Journal of Educational Psychology, 98*(3), 508-528.

Vadasy, P. F., Sanders, E. A., & Tudor, S. (2007). Effectiveness of paraeducator-supplemented individual instruction: Beyond basic decoding skills. *Journal of Learning Disabilities, 40*(6), 508-525.

Vandervelden, M. C., & Siegel, L. S. (1997). Teaching phonological processing skills in early literacy: A developmental approach. *Learning Disability Quarterly*, *20*(2), 63-81.

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., . . . Francis, D. J. (2006). Effectiveness of a spanish intervention and an english intervention for english-language learners at risk for reading problems. *American Educational Research Journal, 43*(3), 449-487.

Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., & ... 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.

Vaughn, S., Cirino, P. T., Wanzek, J., Wexler, J., Fletcher, J. M., Denton, C. D., . . . Francis, D. J. (2010). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. *School Psychology Review*, *39*(1), 3-21.

Appendix D: References for Follow-up Studies

Bruck, M., Teiman, R., Caravolas, M., Genesee, F., & Cassar, M. (1998). Spelling skills of children in whole language and phonics classrooms. *Applied Psycholinguistics, 19*(4), 669-84.

Byrne, B., & F. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology, 83*, 451-455.

Byrne, B., & F. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology, 85*, 104-111.

Byrne, B., & F. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology, 87*, 488-503.

Center, Y., Freeman, L., & Robertson, G. (1998). An evaluation of schoolwide early language and literacy (SWELL) in six disadvantaged schools. *International Journal of Disability, Development and Education, 45*(2), 143-172.

Chall, J. S. (1996). *Learning to read: The great debate* (3rd ed.). Fort Worth, TX: Harcourt Brace College.

Cirino, P. T., Vaughn, S., Linan-Thompson, S., Cardenas-Hagan, E., Fletcher, J. M., & Francis, D. J. (2009). One-year follow-up outcomes of spanish and english interventions for english language learners at risk for reading problems. *American Educational Research Journal, 46*(3), 744-781.

Ecalle, J., Magnan, A., & Calmus, C. (2009). Lasting effects on literacy skills with a computer-assisted learning using syllabic units in low-progress readers. *Computers & Education, 52*(3), 554-561.

Gillon, G. T. (2005). Facilitating phoneme awareness development in 3- and 4-year-old children with speech impairment. *Language, Speech, and Hearing Services in Schools, 36*(4), 308-324.

Hurry, J., & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading, 30*(3), 227-248.

Johnston, R. S., & Watson, J. E. (2004). Accelerating the development of reading, spelling and phonemic awareness skills in initial readers. *Reading and Writing: An Interdisciplinary Journal, 17*(4), 327-357.

Kirk, C., & Gillon, G. T. (2007). Longitudinal effects of phonological awareness intervention on morphological awareness in children with speech impairment. *Language, Speech & Hearing Services in Schools, 38*(4), 342-352.

Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, *23*(3), 263-284.

Mantzicopoulos, P., Morrison, D., Stone, E., & Setrakian, W. (1992). Use of the SEARCH/TEACH tutoring approach with middle-class students at risk for reading failure. *The Elementary School Journal, 92*(5), 573-586.

Nancollis, A., Lawrie, B., & Dodd, B. (2005). Phonological awareness intervention and the acquisition of literacy skills in children from deprived social backgrounds. *Language, Speech, and Hearing Services in Schools*, *36*(4), 325-335.

O'Connor, M., Arnott, W., McIntosh, B., & Dodd, B. (2009). Phonological awareness and language intervention in preschoolers from low socio-economic backgrounds: A longitudinal investigation. *British Journal of Developmental Psychology*, *27*(4), 767-782.

Schneider, W., & K. (1997). Short-and long-term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology, 66*, 311-340.

Snider, V. E. (1990). Direct instruction reading with average first-graders. *Reading Improvement, 27*(2), 143-48.

Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, *69*(4), 587-605.

References

Apel, K., Masterson, J. J., & Niessen, N. L. (2004). Spelling assessment frameworks. In E. R. Silliman, B. J., Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders*, 644-660.

Arra, C. T., & Aaron, P. G. (2001). Effects of psycholinguistic instruction on spelling performance. *Psychology in the Schools, 38*(4), 357.

Baron, D. E. (1982). *Grammar and Good Taste: Reforming the American Language*. Yale University Press, 302 Temple St., New Haven, CT 06511 (25.00 cloth--ISBN-0-300-02799-0; $9.95 paperback--ISBN-0-300-03080-0). National Council of Teachers of English, 1111 Kenyon Rd., Urbana, IL.

Barry, C. (1994). Spelling routes (or roots or rutes). In Brown, G.D. & Ellis, N. C. (Eds.), *Handbook of spelling: Theory, process and intervention*, (pp. 27-49). John Wiley & Sons Ltd.

Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2000). Words their way: Word study for phonics, vocabulary, and spelling instruction (2nd ed.). Columbus, OH: Merrill.

Berninger, V., & Fayol, M. (2008). Why spelling is important and how to teach it effectively. *Encyclopedia of language and literacy development*.

Bissex, G. (1980). GNYS AT WRK: A child learns to read and write. *Cambridge, MA: Harvard*.

Borenstein, M. (2009). *Introduction to meta-analysis*. Chichester, U.K.: John Wiley & Sons.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive meta-analysis version 2 (computer software). Englewood, N.J.: Biostat.

Bosman, A. M., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A., Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages*, (pp. 173-194). Routledge.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80* (2), 144-179.

Bradley, L., & B. (1983). Categorizing sounds and learning to read—A causal connection. *Nature, 301*, 419-421.

Brown, A. S. (1988). Encountering misspellings and spelling performance: Why wrong isn't right. *Journal of Educational Psychology*, 80(4), 488-94.

Brown, G. D., & Loosemore, R. P. (1994). Computational approaches to normal and impaired spelling. In E. R. Silliman, B. J., Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders*, 319-335.

Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403-

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, *45*(4), 751-774.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, *45*(4), 464-487.

Cataldo, S., & Ellis, N. (1988). Interactions in the development of spelling, reading and phonological skills. *Journal of Research in Reading*, 11(2), 86-109.

Chan, M. E., & Arvey, R. D. (2012). Meta-analysis and the development of knowledge. *Perspectives on Psychological Science, 7*(1), 79-92.

Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*, (pp. 6-23). Oxford: Blackwell.

Coltheart, M. (2006). Dual route and connectionist models of reading: An Overview. *London Review Of Education*, *4*(1), 5-17.

Cooke, A. (1997). Learning to spell difficult words: why look, cover, write and check is not enough. *Dyslexia (10769242)*, *3*(4), 240-243.

Cooper, H. M., & Cooper, H. M. (2010). *Research synthesis and meta-analysis: A step-by-step approach*. Los Angeles: Sage.

Cummings, D. W. (1988). *American English spelling: An informal description*. JHU Press.

Dewey, G. (1971). *English Spelling: Roadblock to Reading.*

Diaz, I. (2010). *The effect of morphological instruction in improving the spelling, vocabulary, and reading comprehension of high school english language learners (ELLs).* TUI University). *ProQuest Dissertations and Theses,* 106. (250906858)

Duval, S. & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455-463.

Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In

Perfetti, C. A., Rieben, L., & Fayol, M. (Eds., *Learning to spell: Research, theory, and practice across languages*. (pp. 237-269). Mahwah, N.J: Lawrence Erlbaum Associates.

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy*, 3-40.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393-447.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250-287.

Ehri, L. C., & Rosenthal, J. (2010). Spellings of words: A neglected facilitator of vocabulary learning. In *Literacy Development and Enhancement Across Orthographies and Cultures* (pp. 137-152). Springer US.

Ehri, L. C., & Wilce, L. S. (1982). Recognition of spellings printed in lower and mixed case: Evidence for orthographic images. *Journal of Literacy Research*, *14*(3), 219-230.

Ehri, L. C., & Wilce, L. S. (1987). Does learning to spell help beginners learn to read words?. *Reading Research Quarterly*, 47-65.

Fashola, O. S., Drum, P. A., Mayer, R. E., & Kang, S. J. (1996). A cognitive theory of orthographic transitioning: Predictable errors in how Spanish-speaking children spell English words. *American Educational Research Journal*, 33(4), 825-843.

Ferguson, C. J. & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*(1), 120-128.

Fink, R. P. (1998). Literacy development in successful men and women with dyslexia. *Annals of Dyslexia*, *48*(1), 311-346.

Fischer, F., Shankweiler, D., & Liberman, I. Y. (1985). Spelling proficiency and sensitivity to word structure. Journal of Memory and Language, 24(4), 423-441.

Foorman, B. R., Francis, D. J., Novy, D. M., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal Of Educational Psychology*, *83*(4), 456-469.

Fresch, M. J. (2003). A national survey of spelling instruction: Investigating teachers' beliefs and practice. *Journal of Literacy Research*, *35*(3), 819-848.

Fresch, M. J. (2007). Teachers' concerns about spelling instruction: A national survey. *Reading Psychology*, *28*(4), 301-330.

Frith, U. (Ed.). (1980). Cognitive processes in spelling. London: Academic Press.

Frost, R. (1992). Orthography and phonology. *The linguistics of literacy*, 255-74.

Fulk, B. M., & Stormont-Spurgin, M. (1995). Spelling interventions for students with disabilities: A review. *The Journal of Special Education*, *28*(4), 488-513.

Fry, E. (2004). Phonics: A large phoneme-grapheme frequency count revised. *Journal of Literacy Research*, 36(1), 85-98.

Ganske, K. (1999). The developmental spelling analysis: A measure of orthographic knowledge. *Educational Assessment*, *6*(1), 41-70.

Gentry, J. R. (1978). Early spelling strategies. *The Elementary School Journal*, 79(2), 88-92.

Gentry, J. R. (1982). An Analysis of Developmental Spelling in" GNYS AT WRK". *The Reading Teacher*, 36(2), 192-200.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, 100(3), 566.

Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing*, *25*(2), 321-346.

Gettinger, M. (1993). Effects of invented spelling and direct instruction on spelling performance of second-grade boys. *Journal of Applied Behavior Analysis*, *26*(3), 281-291.

Gill, C. H., & Scharer, P. L. (1996). " Why Do They Get It on Friday and Misspell It on Monday?" Teachers Inquiring About Their Students as Spellers. *Language Arts*, 73(2), 89-96.

Gordon, J., Vaughn, S., & Schumm, J. S. (1993). Spelling interventions: A review of literature and implications for instruction for students with learning disabilities. *Learning Disabilities Research & Practice*.

Graham, S., Morphy, P., Harris, K. R., Fink-Chorzempa, B., Saddler, B., Moran, S., & Mason, L. (2008). Teaching spelling in the primary grades: A national survey of instructional practices and adaptations. *American Educational Research Journal*, 45(3), 796-825.

Greenberg, D., Ehri, L. C., & Perin, D. (2002). Do adult literacy students make the same word-reading and spelling errors as children matched for word-reading age? *Scientific Studies*

of Reading, 6(3), 221-243.

Hanna, P.R., Hanna, J.S., Hodges, R. E., & Rudorf, H. (1966). *Phoneme-grapheme correspondences as cues to spelling improvement.* Washington, DC United States Office of Education Cooperative Research.

Hempenstall, K. (2008). Corrective Reading: An Evidence-Based Remedial Reading Intervention. *Australasian Journal Of Special Education, 32*(1), 23-54.

Henderson, E. H., & Beers, J. W. (1980). Developmental and cognitive aspects of learning to spell: A reflection of word knowledge. International Reading Association, 800 Barksdale Rd., PO Box 8139, Newark, DE.

Henderson, E. H. (1981). Learning to read and spell: The child's knowledge of words. Dekalb, IL: Northern Illinois University Press.

Hilte, M., Bos, M., & Reitsma, P. (2005). Effects of spelling pronunciations during spelling practice in Dutch. *Written Language & Literacy*, *8*(2), 61-77.

Homer, B. D. (2009). Literacy and metalinguistic development. In D. R. Olson, N. Torrance, D. R. Olson, N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 487-500). New York, NY US: Cambridge University Press.

Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology*, 20(2), 115-162.

Hunter, J. E., & Schmidt, F. L. (2006). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, Calif. Sage.

International Phonetic Association. (1999). Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet. Cambridge Univ Pr.

Jacoby, L. L., & Hollingshead, A. (1990). Reading student essays may be hazardous to your spelling: Effects of reading incorrectly and correctly spelled words. *Canadian Journal of Psychology*, *44*(3), 345-358.

Johnston, F. (2001). Exploring classroom teachers' spelling practices and beliefs. *Reading Research & Instruction, 40,* 143-156.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437-447.

Kamhi, A. G., & Hinton, L. N. (2000). Explaining individual differences in spelling ability. *Topics in Language Disorders*, *20*(3), 37-49.

Moats, L. C. (2005). How spelling supports reading. *American Educator, 29*(4), 12-22.

Kessler, B., & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language*, 44(4), 592-617.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18(2), 201-212.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., & National Inst. of Child Health and Human Development (NIH), B. D. (NIH), Bethesda, MD. (1989). *The Alphabetic Principle and Learning To Read*.

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

Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, *23*(3), 263-284.

Martinet, C., Valdois, S., & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition*, *91*(2), B11-B22.

Maughan, B. B., Messer, J. J., Collishaw, S. S., Pickles, A. A., Snowling, M. M., Yule, W. W., & Rutter, M. M. (2009). Persistence of literacy problems: Spelling in adolescence and at mid-life. *Journal of Child Psychology & Psychiatry*, *50*(8), 893-901.

McCandliss, B., Beck, I. L., Sandak, R., & Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: Design and preliminary tests of the word building intervention. Scientific Studies of Reading, 7(1), 75-104.

Mcclurg, P. A., & Kasakow, N. (1989). Wordprocessors, spelling checkers, and drill and practice programs: Effective tools for spelling instruction?. *Journal of Educational Computing Research*, *5*(2), 187-198.

McCutchen, D., & Berninger, V. W. (1999). Those who know, teach well: Helping teachers master literacy-related subject-matter knowledge. *Learning Disabilities Research & Practice*, *14*(4), 215-226.

Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. (2005). Literacy as a Unidimensional Multilevel Construct: Validation, Sources of Influence, and Implications in a Longitudinal Study in Grades 1 to 4. *Scientific Studies Of Reading*, *9*(2), 85-116.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological Bulletin*, *138*(2), 322.

Mitchell, P., Kemp, N., & Bryant, P. (2011). Variations Among Adults in Their Use of Morphemic Spelling Rules and Word-Specific Knowledge When Spelling. *Reading Research Quarterly*, *46*(2), 119-133.

Moats, L. C. (1995). *Spelling: Development, disabilities, and instruction*. Timonium, MD: York Press.

Moats, L. (2009). Still wanted: Teachers with knowledge of language. *Journal of Learning Disabilities*, *42*(5), 387-391.

Moats, L., & Lyon, G. (1996). Wanted: Teachers with knowledge of language. *Topics in Language Disorders*, *16*(2), 73-86.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*,*13*(2), 125-145.

Morris, D., Blanton, L., Blanton, W. E., & Perney, J. (1995). Spelling instruction and achievement in six classrooms. *The Elementary School Journal*, 145-162.

Morris, D., & Perney, J. (1984). Developmental spelling as a predictor of first-grade reading achievement. *The Elementary School Journal*, *84*(4), 441-457.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

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 reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Nies, K. A., & Belfiore, P. J. (2006). Enhancing spelling performance in students with learning disabilities. *Journal of Behavioral Education*, *15*(3), 162-169.

Nunes, T., & Bryant, P. (2009). *Children's reading and spelling: Beyond the first steps*. Chichester: Wiley-Blackwell.

Olson, D. R. (1994). *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.

Olson, D. R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition*, 60(1), 83-104.

Ormrod, J. E., & Spivey, N. R. (1990). Overlearning and speeded practice in spelling instruction. *Psychological Reports*, *67*(2), 365-366.

Ouellette, G. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology*, 107(1), 50-58.

Podgursky, M., Monroe, R., & Watson, D. (2002). Teacher mobility, pay, and academic quality. *Economics of Education Review*.

Pollo, T. C., Kessler, B., & Treiman, R. (2005). Vowels, syllables, and letter names: Differences between young children's spelling in English and Portuguese. *Journal of Experimental Child Psychology, 92*(2), 161-181.

Pollo, T. C., Treiman, R., & Kessler, B. (2008). Three perspectives on spelling development. In Naples, A. J. (Ed.), *Single-word reading: Behavioral and biological perspectives*, (pp. 175-189). Psychology Press.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2(2), 31-74.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346-362.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Schlagal, B. (2002). Classroom spelling instruction: History, research, and practice. *Literacy Research and Instruction*, 42(1), 44-57.

Schermerhorn, P. K., & McLaughlin, T. F. (1997). Effects of the add-a-word spelling program on test accuracy, grades, and retention of spelling words with fifth and sixth grade regular education students. *Child & family behavior therapy*, *19*(1), 23-35.

Schlagal, B. (2007). Best practices in spelling and handwriting. In S. Graham, C. MacAruthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (pp. 179-201). New York: Guilford.

Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, Mass: Harvard University Press.

Shallice, T. (1981). Phonological agraphia and the lexical route in writing. *Brain: a journal of neurology*, *104*(3), 413.

Shankweiler, D., Lundquist, E., Dreyer, L. G., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and consequences. *Reading and Writing: An Interdisciplinary Journal, 8,* 267-294.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an" outlier" orthography. *Psychological Bulletin*, *134*(4), 584.

Siegel, L. S. (2008). Morphological awareness skills of English language learners and children with dyslexia. *Topics in Language Disorders*, *28*(1), 15-27.

Singer, B. D., & Bashir, A. S. (2004). Developmental variations in writing composition skills. In E. R. Silliman, B. J., Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders*, 559-82.

Smith, M. L. & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752-760.

Stage, S. A., & Wagner, R. K. (1992). Development of young children's phonological and orthographic knowledge as revealed by their spellings. *Developmental Psychology*, 28(2), 287-296.

Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research*, *59*(1), 87-116.

Stanback, M. L. (1992). Syllable and rime patterns for teaching reading: Analysis of a frequency-based vocabulary of 17,602 words. *Annals of Dyslexia*, 42(1), 196-221.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 360-407.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology*, *38*(2), 175-190.

Swanson, H., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research*, *73*(4), 407-440.

Templeton, S., & Morris, D. (1999). Questions teachers ask about spelling. *Reading Research Quarterly*, 34(1), 102-112.

Thompson, G. B., Fletcher-Flinn, C. M., & Cottrell, D. S. (1999). Learning correspondences between letters and phonemes without explicit instruction. *Applied Psycholinguistics*, 20(1), 21-50.

Torgerson, C. J., & Elbourne, D. (2002). A systematic review and meta-analysis of the effectiveness of information and communication technology (ICT) on the teaching of spelling. *Journal of Research in Reading*, *25*(2), 129-143.

Torgerson, C., Brooks, G., Hall, J. (2006). A systematic review of the research literature on the use of phonics in the teaching of reading and spelling. Research Report No. 711. London: Department of Education and Skills. (University of Sheffield)

Treiman, R. (1993). Beginning to spell. Oxford University Press, 200 Madison Avenue, New York, NY 10016.

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, 98(3), 642.

Treiman, R., Kessler, B., & Bourassa, D. (2001). Children's own names influence their spelling. *Applied Psycholinguistics*, *22*(4), 555-570.

Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly*, *28*(3), 218-233. doi:10.2307/747995

National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2011* (NCES 2012–457). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Vaughn, S., Hughes, M. T., Moody, S. W., & Elbaum, B. (2001). Instructional Grouping for Reading for Students with LD Implications for Practice. *Intervention in School and Clinic*, *36*(3), 131-137.

Valentine, J. C. (2009). Judging the quality of primary research. In H. Cooper, L. V. Hedges, J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101,* 192-212.

Wanzek, J., Vaughn, S., Wexler, J., Swanson, E. A., Edmonds, M., & Kim, A. (2006). A synthesis of spelling and reading interventions and their effects on the spelling outcomes of students with LD. *Journal of Learning Disabilities, 39*(6-), 528-543.

Watson, A. (1935). Experimental studies in the psychology and pedagogy of spelling. *The Teachers College Record*, *37*(3), 233-235.

Worthy, M. J., & Invernizzi, M. (1990). Spelling errors of normal and disabled students on achievement levels one through four: Instructional implications. *Annals of Dyslexia*, 40(1), 138-151.

Wu, H. (1999). Basic skills versus conceptual understanding. *American Educator*, 23(3), 14-19.

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3-29.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for-ough and the spelling for/u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers*, 29(4), 600-618.

Zutell, J., & Rasinski, T. (1989). Reading and spelling connections in third and fifth grade students. *Reading Psychology: An International Quarterly*, 10(2), 137-155.