Reading and Writing Relations Are Not Uniform: They Differ by the Linguistic Grain Size, Developmental Phase, and Measurement

Young-Suk Grace Kim iD
Alissa Wolters
Joong won Lee
University of California, Irvine

We conducted a meta-analysis to investigate reading-writing relations. Beyond the overall relation, we systematically investigated moderation of the relation as a function of linguistic grain size (word reading and spelling versus reading comprehension and written composition), measurement of reading comprehension (e.g., multiple choice, open-ended, cloze), and written composition (e.g., writing quality, writing productivity, writing fluency, writing syntax), and developmental phase of reading and writing (grade levels as a proxy). A total of 395 studies (k = 2,265, N = 120,669) met inclusion criteria. Overall, reading and writing were strongly related (r = .72). However, the relation differed depending on the subskills of reading and writing such that word reading and spelling were strongly related (r =.82) whereas reading comprehension and written composition were moderately related (r =.44). In addition, the word reading-spelling relation was stronger for primary-grade students (r =.82) than for university students/adults (r =.69). The relation of reading comprehension with written composition differed depending on measurement of reading comprehension and written composition—reading comprehension measured by multiple choice and open-ended tasks had a stronger relation with writing quality than reading comprehension measured by oral retell tasks; and reading comprehension had moderate relations with writing quality, writing vocabulary, writing syntax, and writing conventions but had weak relations with writing productivity and writing fluency. Relations tended to be stronger when reliability was higher, and the relation between word reading and spelling was stronger for alphabetic languages (r = .83) than for Chinese (r = .71). These results add important nuances about the nature of relations between reading and writing.

KEYWORDS: reading, writing, literacy, meta-analysis, reading-writing relations

1

Comprehending and producing written texts involve the "interplay of mind and text that brings about new interpretations, reformulations of ideas, and new learnings" (Langer, 1986, pp. 2–3). They both involve processing print and meaning, and, therefore, it is reasonable to hypothesize that reading and writing skills are related and that individuals who read well also write well, and those who write well also read well (Fitzgerald & Shanahan, 2000; Kim, 2020a; Shanahan, 2016). In the present study, we conducted a meta-analysis to summarize reading-writing relations. Specifically, based on a theoretical model of reading-writing relations, the interactive dynamic literacy model (Kim, 2020a, 2022), we systematically examined whether reading-writing relations differ by several important factors: (a) linguistic grain size (units or chunks of language, such as sublexical, lexical, and discourse units), specifically lexical literacy (word reading and spelling) versus discourse literacy (reading comprehension and written composition); (b) measurement/assessment of reading comprehension (cloze task, multiple choice, open-ended, and retell) and written composition (writing quality, writing productivity, writing fluency, writing syntax, writing vocabulary, and writing conventions); and (c) developmental phase of reading and writing development (e.g., emergent reading, early reading, advanced) using grade levels as a proxy.

Theoretical Background

Theoretical models of writing such as the cognitive model of writing (Hayes, 1996) and the direct and indirect effects model of writing (Kim, 2020b) explicitly recognize the interactions between the reading process and writing process. Writing often involves reading source texts and reading one's own text during a revision process, and, therefore, constructing an accurate and rich representation of meaning of source texts or one's own texts (reading comprehension) is a central part of the writing process (Deane et al., 2008; Hayes, 1996; Kim & Graham, 2022).

Several hypotheses have been proposed for the relation between reading and writing skills (Shanahan, 2016). According to one account, reading and writing are related because they are functionally related—reading activities often require writing, and writing activities also accompany reading (Langer & Applebee, 1987). Another line of work explains reading-writing relations with a focus on shared knowledge that reading and writing draw on. According to the shared knowledge hypothesis (Fitzgerald & Shanahan, 2000), the reading and writing relation exists because reading and writing largely share common sources of knowledge. For example, they both draw on metaknowledge about functions and purposes of reading and writing, as well as monitoring one's own meaning-making processes and strategies. Reading and writing also involve domain knowledge about substance and content; linguistic knowledge (vocabulary); and procedural knowledge such as how to access, use, and generate knowledge and smoothly instantiate various processes.

The interactive dynamic literacy model (Kim, 2020a, 2022) articulates shared skills and knowledge between reading and writing, specifying shared skills by linguistic grain size: lexical literacy skills, word reading and spelling, and discourse literacy skills—reading comprehension and written composition. Word reading and spelling are related because they involve essentially the same

processes and skills, phonology, orthography, and morphology (Adams, 1990; Bahr et al., 2012; Kim, 2020a) as well as domain-general cognitions such as working memory and attentional control (Kim, 2020a). The discourse literacy skills reading comprehension and written composition are related because they both involve meaning-making processes and draw on language skills, such as vocabulary, syntactic knowledge, and discourse oral language skills (listening comprehension and oral production; e.g., Ahmed et al., 2014; Berninger & Abbott, 2010; Kim & Graham, 2022); higher order cognitions such as inference, perspective taking, and self-regulation (e.g., setting goals and monitoring one's comprehension and performance; Cain et al., 2004; Graham & Harris, 2003; Kim & Schatschneider, 2017); background knowledge such as topic knowledge and discourse knowledge (knowledge about characteristics of different genres and about procedures and strategies to present content appropriate for the genre; e.g., Kim, 2020b; Olinghouse & Graham, 2009); and domain-general cognitive skills (Berninger & Winn, 2006; Berninger et al., 2010; Kim & Graham, 2022).

Beyond the shared skills between reading and writing, the interactive dynamic literacy model has hypotheses about structural relations. Directly relevant to the present study is the dynamic relations hypothesis, which states that the magnitudes or strengths of reading-writing relations differ by linguistic grain size (i.e., lexical vs. discourse skills), measurement of reading comprehension and written composition (e.g., reading comprehension tasks and dimensions of written composition), and developmental phase (e.g., early vs. later phase of reading and writing development).

Linguistic Grain Size

According to the interactive dynamic literacy model (Kim, 2020a, 2022), lexical literacy skills, word reading and spelling, are hypothesized to be more strongly related than are discourse literacy skills, reading comprehension and written composition (i.e., dynamic relations as a function of linguistic grain sizes). Word reading and spelling have stronger relations because they essentially involve the same processes; although, the sequence of processes is different—word reading involves retrieving letters and graphemes and their associated phonological (speech sound structures such as syllables, rimes, and phonemes) and morphological (morphemes, the smallest unit of meaning) information whereas spelling involves retrieving phonological and morphological information, mapping them with graphemes (letters and groups of letters that represent a phoneme), and then assembling and writing them in the correct sequence (Ehri, 1997; Kim, 2022). As such, word reading and spelling rely on essentially the same set of skills—that is, phonological, orthographic, and morphological awareness (Adams, 1990; Bahr et al., 2012; Kim, 2020a). If word reading and spelling draw on a limited number of the same skills and their processes are highly similar, they are likely to be strongly related such that students who have a strong word reading skill have a high likelihood of having a strong spelling skill, and those with a weak word reading skill have a high likelihood of having a weak spelling skill (Berninger et al., 2008; Graham et al., 2021; Kim, 2020a, 2022).

It should be noted that the hypothesis of a strong relation between word reading and spelling does not entail that they are identical skills (Kim, 2022). In most

languages with alphabetic writing systems, phoneme-grapheme correspondences are less consistent than for grapheme-phoneme consistency (Moll & Landerl, 2009). Word reading is a receptive skill where the primary task is to recognize and retrieve grapheme-phoneme/morpheme correspondences whereas spelling is an expressive skill that requires encoding phoneme and morpheme information to graphemes in an accurate sequence and formation. For example, reading/decoding words with r-controlled vowels, such as *her, bird*, and *surf*, requires recognizing the -er, -ir, and -ur orthographic patterns for /ɚ/, and therefore one can successfully read words using partial cues or incomplete mental representations of orthographic knowledge. In contrast, accurately spelling these words requires precise representation and retrieval of word-specific orthographic knowledge (Ehri, 1997; Shanahan, 2016). Evidence also suggests dissociation between word reading and spelling skills. For example, working with a representative sample of German-speaking elementary school children, and using the same words for word reading and spelling tasks, Moll and Landerl (2009) identified students who had discrepant profiles between word reading and spelling: good readers/poor spellers and poor readers/good spellers. Evidence also revealed subgroups with isolated reading difficulties, isolated spelling difficulties, and combined reading and spelling difficulties for children learning to read English and other European languages (Furnes et al., 2019; Moll & Landerl, 2009; Torppa et al., 2017; Wimmer & Mayringer, 2002).

The relation between reading comprehension and written composition is expected to be more moderate than the word reading-spelling relation because their processes are more divergent, and the processes differently tap into skills and knowledge (Kim, 2020a, 2022). For example, one prominent difference between reading comprehension and written composition is whether or not the text is provided or generated. In reading comprehension, the text is given to readers and, although the reader activates relevant topic knowledge and engages in meaning-building and meaning-generating processes, the extent of meaning-making is confined by the given text. Therefore, readers focus on comprehension and validation of the given text (Langer, 1986; Langer & Flihan, 2000). Writers, on the other hand, generate a text, and, therefore, writers are more concerned with setting goals and subgoals and generating texts and associated aspects such as mechanics, syntax, and lexical choices (Langer & Flihan, 2000). These differences in reading comprehension and written composition require readers and writers to differentially tap self-directed processes and to employ nonautomatic strategies and corrective/repair actions in reading comprehension versus written composition. Consequently, although both reading comprehension and written composition involve meaning-building and meaning-generating processes, their relation is hypothesized to be weaker than that for word reading and spelling (Kim, 2020a, 2022).

Measurement of Reading Comprehension and Written Composition

The interactive dynamic literacy model (Kim, 2020a, 2022) hypothesizes that reading-writing relations vary as a function of the measurement of constructs, reading comprehension and written composition in particular (i.e., dynamic relations as a function of measurement). Reading comprehension and written

composition are complex constructs. As such, they are measured in multiple ways. Reading comprehension is widely measured by using multiple-choice, open-ended, cloze, and retell tasks, and studies have shown a varying degree of relations among them (e.g., Cao & Kim, 2021; Francis et al., 2005; Howell & Nolet, 2000; Keenan et al., 2008; Kim, 2020c; Nation & Snowling, 1997). In multiple-choice and short open-ended tasks, individuals read passages and are asked to answer questions designed to tap into the readers' representation of explicitly stated information (literal comprehension questions) and implicit information (inferential comprehension questions) as well as the readers' evaluation of information (evaluation question; Mazany et al., 2015). In cloze tasks, every nth word (typically 5th or 7th word) in the passage is omitted, and individuals are required to fill in the blanks with the words that were deleted. In retell tasks, individuals are asked to retell verbally or to write what they read after passages. Research suggests that reading comprehension tasks vary in the extent to which they tap into comprehension processes. Multiple choice tasks, for example, can tap into deep or inferential comprehension when constructed well (Mazany et al., 2015), whereas cloze and retell tasks are limited in tapping into inferential or higher order integration processes (Cao & Kim, 2021; Francis et al., 2005; Howell & Nolet, 2000; Keenan et al., 2008; Nation & Snowling, 1997).

Like reading comprehension, written composition is measured in multiple ways. Written composition is one of the most widely used performance-based assessments and is evaluated on multiple dimensions, such as quality and conventions (Kim et al., 2014, 2015). Writing quality is the presumed construct in theoretical models of writing and is typically evaluated for the extent to which ideas are coherent and clearly represented, organized using rich and appropriate vocabulary and sentences, and presented with appropriate writing conventions (e.g., Coker et al., 2018; Graham et al., 2002; Hooper et al., 2002). Beyond writing quality, other dimensions are also widely evaluated, including productivity, fluency, vocabulary use, syntax, and conventions. Writing productivity refers to the amount of texts in written composition and is typically measured by the number of words or sentences (e.g., Abbott & Berninger, 1993; Berman & Verhoevan, 2002; Kim et al., 2011; Mackie & Dockrell, 2004; Olinghouse & Graham, 2009). Writing fluency refers to the amount of text written within a specified time (e.g., 1 minute or 3 minutes; Ahmed et al., 2014; Babayigit & Stainthorp, 2011; Kim et al., 2015). Writing productivity and fluency are not regarded as the ultimate outcomes of writing, but they are widely used as indicators of writing quality for developing writers because of their moderate to fairly strong relation to writing quality (e.g., Abbott & Berninger, 1993; Kim et al., 2011, 2014; Mackie & Dockrell, 2004; Wagner et al., 2011) and their utility for screening and progress monitoring purposes (Graham et al., 2011; McMaster & Espin, 2007). Additionally, studies have investigated language features in written composition, such as vocabulary use (e.g., the extent of sophisticated vocabulary and vocabulary diversity; Malpique et al., 2020; Olinghouse & Leaird, 2009; Shanahan, 1980) and sentences and syntax in writing (e.g., syntactic complexity, text cohesion, and diversity of sentence structures; Malpique et al., 2020; Shanahan, 1980; Smith, 2011). Lastly, the extent to which writing conventions such as capitalization, punctuation, and spelling are accurately

employed is widely evaluated (Kim et al., 2014; Malpique et al., 2020; Smith, 2011).

Studies have shown that the various dimensions of written composition noted previously are related but dissociable (Coker et al., 2018; Kim et al., 2014, 2015; Puranik et al., 2008; Wagner et al., 2011). Furthermore, they differentially tap into skills and knowledge (e.g., language, transcription skills). For instance, whereas writing quality draws on a comprehensive set of skills such as transcription skills, language and cognitive skills, and background knowledge (e.g., Abbott & Berninger, 1993; Berninger & Abbott, 2010; Coker, 2006; Graham & Santangelo, 2014; Kim, 2020b; Kim & Graham, 2022; Santangelo & Graham, 2016), writing productivity largely relies on transcription skills (Kim & Graham, 2022; Kim et al., 2014, 2015). Based on these findings, the interactive dynamic literacy model (Kim, 2020a, 2022) hypothesizes that the relation between reading comprehension and written composition varies depending on the measurement of reading comprehension and written composition: reading comprehension is differentially related to written composition as a function of dimensions of written composition, and written composition is differentially related to reading comprehension depending on the type of reading comprehension tasks. For example, reading comprehension is hypothesized to have a stronger relation with writing quality than with writing productivity or writing fluency (Kim, 2020a, 2022). Reading comprehension captures one's representation of a coherent mental model of the text (Kintsch, 1988). Of the various dimensions of written composition, writing quality is the one that captures overall quality of rich, accurate, clear, and coherent representation of ideas. Hence, writing quality is expected to be more strongly related to reading comprehension than writing productivity or writing fluency, and a recent study with grade 2 children supported this hypothesis (Kim & Graham, 2022). Similarly, if multiple choice and short open-ended tasks more readily tap into deep comprehension than cloze or retell tasks, reading comprehension measured by multiple choice and short open-ended tasks are likely to have a stronger relation with writing quality than reading comprehension measured by cloze or retell tasks.

Developmental Phase of Reading and Writing

The interactive dynamic relations model also posits dynamic/differential relations as a function of developmental phase (Kim, 2022). According to this hypothesis, the magnitudes of reading-writing relations vary depending on the developmental phase. Specifically, word reading and spelling are hypothesized to have a stronger relation in the beginning phase than a later phase of development. In the early phase, the majority of individuals are beginning to develop word reading and spelling skills relying on the phonological, orthographic, and morphological processes whereas at a later phase of development, many individuals would have developed sufficient skills and reach asymptotes in word reading and spelling. This entails less variation between individuals at an advanced phase of word reading and spelling development and, therefore, a weaker relation than in the beginning phase of word reading and spelling development. In contrast, reaching asymptotes is not expected in reading comprehension and written composition because these skills continue to develop into adulthood (Chall, 1983; Kellogg, 2008).

Present Study

The goal of this study was to systematically capture the relation between reading and writing, examining whether the magnitude of the relation varies as a function of linguistic grain size, measurement of reading comprehension and written composition, and developmental phase of reading and writing proxied by grade levels. The following were specific research questions.

1. What is the overall magnitude of the relation between reading and writing?
2. How are different subskills of reading (word reading, text reading fluency, reading comprehension) and writing (spelling and written composition) related to one another? Does the relation differ as a function of linguistic grain size (word reading and spelling versus reading comprehension and written composition)?
3. Does the relation between reading comprehension and written composition vary as a function of the types of reading comprehension tasks (e.g., multiple choice, open-ended responses, cloze, and retell) and dimensions of written comprehension (writing quality, writing productivity, writing fluency, writing syntax, writing vocabulary, and writing conventions)?
4. Does the relation differ as a function of the developmental phase of reading and writing proxied by grade levels?

We hypothesized that the overall relation between reading and writing will be at least moderate in magnitude. We also expected that the relation would be stronger for word reading and spelling than for reading comprehension and written composition as stated previously. Differential relations were anticipated between reading comprehension and written composition depending on the types of reading comprehension and dimensions of written composition. Reading comprehension as measured by multiple-choice and open-ended tasks was expected to have a stronger relation with writing quality than does reading comprehension measured by cloze or retell tasks. Overall reading comprehension (across measures) was posited to have a stronger relation with writing quality than with other dimensions, such as writing productivity and writing conventions. Word reading and spelling were expected to have a stronger relation in the earlier phase (e.g., elementary grades) than the later phase of development (e.g., university).

Method

Literature Search

The literature search was conducted on the following databases through ProQuest: Educational Resources Information Center (ERIC), APA PsycInfo, Sociological Abstracts, ProQuest Dissertations & Theses A&I, and ProQuest Dissertations and Theses Global for studies printed between 1980 and 2021. A complete list of search terms is found in the appendix. The year 1980 was chosen because systematic empirical investigations on reading-writing relations began during this time period (e.g., Shanahan & Lomax, 1986). Inclusion criteria were

as follows: (1) word reading, text reading fluency (excluding reading prosody), or reading comprehension was assessed; (2) spelling or written composition was assessed; (3) a Pearson's *r* correlation or raw data of reading and writing measures were reported (excluding composite scores across skills such as word reading and text reading fluency together); (4) intervention studies were included if pretest or control condition data were reported, which were used in coding and analysis; and (5) studies were reported in English—studies of languages and writing systems other than English were included as long as they were reported in English. Studies were not excluded based on publication status and therefore gray literature was included. Reading prosody was not included in the present meta-analysis for two reasons. First, to our knowledge, there is no theoretical account that specifies the relation between reading prosody and writing. Second, although reading prosody is part of a widely accepted definition of text reading fluency (i.e., reading connected texts with accuracy, speed, and expression; National Institute of Child Health and Human Development, 2000), the vast majority of previous work on text reading fluency focused on text reading efficiency (e.g., the number of words read correctly within a specified time), and reading prosody is rarely examined in relation to writing. Spelling was defined as the ability to encode words according to the orthographic system of a language, and therefore, word reading and spelling had to involve reading/decoding and spelling/encoding of words, respectively, in isolation or out of context, and spelling had to be a production task such that recognition of correct spelling (e.g., orthographic choice task) was not included as a spelling task. Text reading fluency and reading comprehension had to involve reading of connected texts: text reading fluency was primarily measured as the number of words read correctly within a specified time, and reading comprehension tasks required individuals to read connected texts and answer comprehension questions, retell the read texts, or fill in clozes. Written composition had to involve producing connected texts (i.e., contrived tasks such as a grammar task and sentence ordering were not included as written composition).

Titles, abstracts, and keywords were uploaded onto the meta-analytic online software, Rayyan, and screened. First and second screening was conducted by the second author and third author who were PhD students in education and had prior experience in meta-analysis on the topic of literacy. Interrater reliability on the first screening process (abstracts) was calculated based on 200 randomly selected studies, yielding 95% exact agreement. In the second screening, full texts of the included studies were examined. Using 150 randomly selected studies, exact agreement was 94%. Disagreements in the first and second screening were resolved through discussion. A total of 344 studies met inclusion criteria in database search. In addition, the four journals with the most included studies (i.e., eight or more studies) through database search, *Reading and Writing* (*n* = 34), *Scientific Studies of Reading* (*n* = 12), *Journal of Educational Psychology* (*n* = 8), and *Journal of Learning Disabilities* (*n* = 8), were digitally hand searched and screened. This resulted in an additional 51 studies that met the inclusion criteria.

As shown in the PRISMA flow diagram (Figure 1), a total of 395 studies were included in the final sample, which included 25 different languages (Arabic, Cantonese, Circassian, Croatian, Danish, Dutch, English, Finnish, French, German, Greek, Korean, Hebrew, Indonesian, Italian, Japanese, Malaysian,
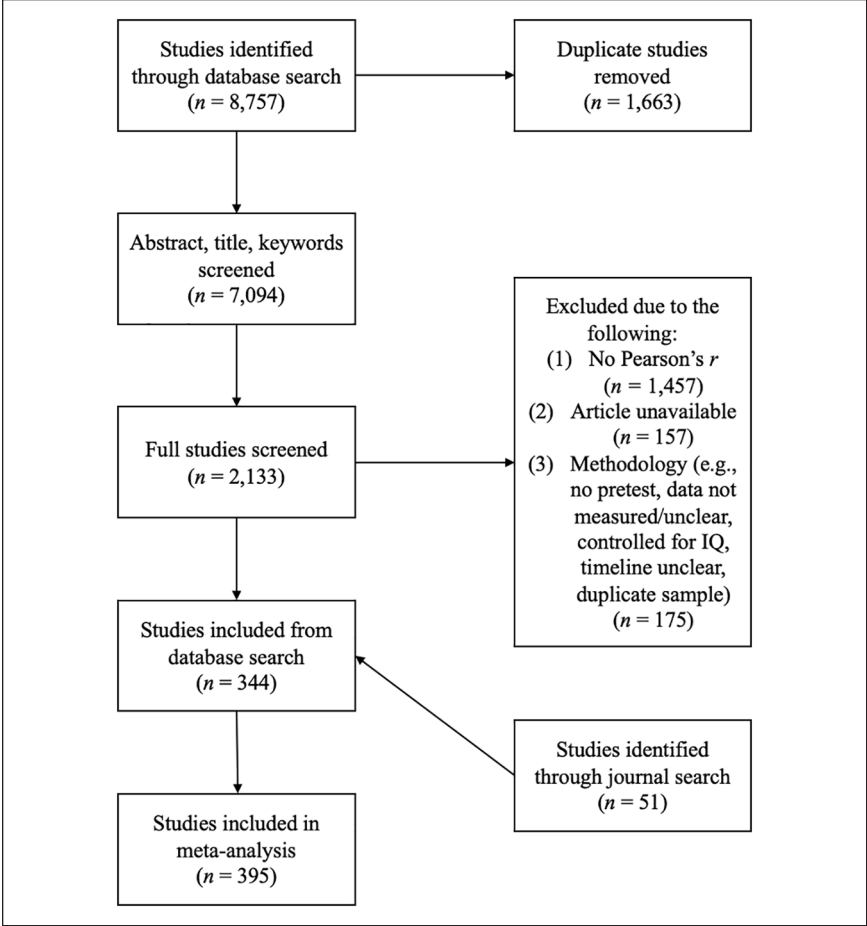
FIGURE 1. *PRISMA chart of the screening process for the current systematic review and meta-analysis*

Mandarin, Norwegian, Persian, Portuguese, Russian, Spanish, Swedish, and Turkish) and a range of developmental stages from preschool to adulthood. Some of the participants had disabilities (e.g., attention deficit disorder, autism spectrum disorder, dyslexia, speech and/or language impairment). However, the vast majority of the studies did not report correlations by subsamples or subgroups of students by these characteristics, and therefore, moderation analysis was not conducted by disability status or type (but sensitivity analysis was conducted; see below). The majority of unavailable documents (see Figure 1) were dissertations, and author contact information was not available; therefore, author contact was not conducted.

Coding Procedures

Included studies were coded for sample size, effect size (Pearson's *r*), reading and writing measures, aspect of written compositions evaluated (quality, productivity, fluency, vocabulary, syntax, and conventions), participant biological sex (percentage female), disability (percentage reported who were diagnosed or receiving support), and grade levels (as a proxy for the developmental phase of reading and writing skills). Reading comprehension tasks were coded for multiple choice, open-ended, cloze, or oral or written retell. Codes for the dimensions of written compositions were as follows: "Quality" included scales that evaluated features such as quality of ideas, structure, cohesion, thematic development, and overall quality (i.e., holistic score); "Productivity" included total number of words and sentences; "Fluency" included words written within a specified time; "Vocabulary" included lexical measures such as type token ratio, lexical density, and percent academic words; "Syntax" included grammatically correct word sequences and sentences or syntax scales that examined syntactic features; "Conventions" included capitalization, punctuation, and spelling.

Thirty studies were randomly selected to establish reliability in coding. Each study was coded for the following: group identification (to differentiate between multiple groups in one study), publication status, effect size, sample size, reading skill (e.g., word reading, reading comprehension), writing skill (e.g., spelling written composition), name of reading measure, name of writing measure, age, grade level, disability status (%), disability type, language, second language status (%), female (%), ethnicity or race, and socioeconomic status. Exact agreement reliabilities for the various codes ranged from 98% to 99%.

A total of 2,265 effect sizes from 120,669 participants (646 unique samples) were included in the final coding document. Pearson's *r* values were converted to Fisher's *z*, and variance was calculated from sample size (Borenstein et al., 2011) by using the following equations.

$z = 0.5$ x $\ln(1+r/1-r)$ $Vz = 1/(n-3)$

Data Analysis

Data were imported into R for analysis. The meta-regression package, *robumeta* (RVE; Hedges et al., 2010), was used to address the research questions. *Robumeta* calculates effect sizes using Fisher's *z* and variance using a random effects model (Kreft & de Leeuw, 1998; Viechtbauer, 2005), taking into account nested structure of effect sizes within each unique sample and weights appropriately for small or large samples (Tipton, 2015). It also tests whether there is a statistically significant difference between effect sizes.

Research questions were addressed using the following analytic procedures. For the first research question, an overall effect size was calculated based on all effect sizes ($k = 2,265$). For the second research question, relations between subskills of reading and writing (e.g., word reading and written composition) were estimated (see Table 1). In addition, I^2 was calculated to determine the heterogeneity of the samples. Both the overall relation and the grain size relations suggested that 81% or more of the variation across the studies was due to differences between studies (Higgins & Thompson, 2002; Higgins et al., 2003). To test

TABLE 1

The relations between reading and writing

	b	SE	p	df	CI.LB	CI.UB	k	Levels	I^2
Overall relation	0.72	0.01	<.001	633	0.70	0.75	2,265	646	94.25
Word reading & spelling	0.82	0.01	<.001	508	0.80	0.85	1,195	518	94.07
Word reading & written composition	0.42	0.03	<.001	50.40	0.37	0.48	200	56	82.96
Reading comprehension & written composition	0.44	0.02	<.001	79.40	0.40	0.49	253	83	90.44
Reading comprehension & spelling	0.59	0.02	<.001	210	0.55	0.62	343	217	92.26
Text reading fluency & spelling	0.65	0.03	<.001	126	0.59	0.70	274	130	92.50
Text reading fluency & written composition	0.33	0.06	<.001	10.50	0.20	0.45	30	12	81.82

Note. Models were tested separately. b = Overall average effect size; SE = standard error; df = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval; k = number of effect sizes; levels = number of unique samples; I^2 = proportion of total variation between samples (heterogeneity).

11

whether the magnitudes of the relation for word reading and spelling versus reading comprehension and written composition are different, a meta-regression model was employed with the word reading and spelling relation as the reference relation (Table 2). To address the third research question about the relation between reading comprehension and written composition as a function of measurement, bivariate correlations were examined (Table 3), followed by meta-regression models (Tables 4 and 5) to test whether magnitudes are statistically significantly different. To address the fourth research question about the moderation by developmental phase, grade levels were entered as predictors for word reading-spelling and reading comprehension-dimensions of written composition outcomes in meta-regression models (Table 6). Grade level was converted into a categorical variable using the following developmental stages: primary grades (preschool to grade 2), upper elementary grades (grades 3–5), secondary grades (grades 6–12), and adult/university. Grouping studies into developmental stages is a common practice in meta-analysis (e.g., Florit & Cain, 2011; García & Cain, 2014; Petscher, 2010); it allows for studies that grouped grades together to not be excluded from analysis and gives more degrees of freedom or effect sizes for investigating development as a potential moderator.

Finally, sensitivity analysis was conducted considering the following: an alternative estimator *metafor*, publication status, reliability of reading and writing measures (meeting the high reliability threshold of .80 or not), normed nature of measures (normed or not), disaggregating primary grades (preschool to grade 2) into two categories—preschool and kindergarten and grades 1 and 2—and disaggregating secondary grades (grades 6–12) into grades 6–8 and grades 9–12, the nature of writing systems (alphabetic vs. morphosyllabic [Chinese]), and participants' disability status. Publication bias was examined using funnel plot asymmetry.

Results

Information and references of the studies included in the meta-analysis are found in the online supplemental materials due to the large number of included studies. As noted previously, *robumeta* was used in the present study to take into account the nested structure or dependency of effect sizes within each unique sample.

Research Question 1: Overall Magnitude of the Relation Between Reading and Writing

The final sample of 395 studies (646 unique samples, $N = 120,669$, $k = 2,265$) yielded an average correlation between reading (word reading, text reading fluency, and reading comprehension) and writing (spelling and written composition) of .72 ($p < .000$, 95% CI = [.70, .75]; Table 1).

Research Question 2: Variation of Relations as a Function of Linguistic Grain Size

The strength of the relation between reading and writing was estimated across different reading skills and writing skills. The relation between word reading and spelling was strongest ($r = .82$) whereas the relation between reading comprehension and written composition was moderate ($r = .44$; see Table 1), and they were statistically significantly different ($p < .001$). In fact, as shown in Table 2, word

TABLE 2

Meta-regression model for comparison of magnitudes

	b	SE	t	p	df	CI.LB	CI.UB
Intercept (word reading and spelling)	0.82	0.01	59.22	<.001	452.01	0.79	0.85
Reading comprehension & written composition	−0.35	0.03	−11.63	<.001	82.46	−0.41	−0.29
Reading comprehension & spelling	−0.23	0.02	−10.23	<.001	217.66	−0.28	−0.19
Text reading fluency & spelling	−0.16	0.03	−4.87	<.001	123.03	−0.23	−0.10
Text reading fluency & written composition	−0.55	0.05	−10.16	<.001	8.69	−0.67	−0.43
Word reading & written composition	−0.38	0.04	−9.69	<.001	48.67	−0.45	−0.30

Note. A total of 646 unique samples and 2,265 effect sizes. *SE* = Standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval.

13

TABLE 3

Bivariate correlations between reading comprehension tasks and writing quality

Written composition dimensions	Types of tasks in reading comprehension			
	Multiple choice	Open-ended	Oral retell[a]	Cloze
Writing quality	.47***	.48***	.55[b]**	.37***
Writing productivity	.23[b]	.28[b]	NA	.18*
Writing fluency	NA	NA	NA	.12[b]
Writing syntax	.46***	NA	NA	.21[b]
Writing vocabulary	.30[b]	NA	NA	.49[b]*
Writing conventions	.51***	NA	NA	NA

[a]Written retell was not included because only two effect sizes were identified (Jenkins et al., 2014).
[b]Less than 4 degrees of freedom and at least 2 effect sizes.
Significance level marked by asterisk to denote the following: *$p < .05$, **$p < .01$, ***$p < .001$.
NA: 1 or no effect sizes identified

TABLE 4

Meta-regression model for reading comprehension tasks (cloze task as the reference) and writing quality

	b	SE	df	p	CI.LB	CI.UB
Intercept	0.37	0.02	11.88	<.001	0.32	0.41
Multiple choice	0.11	0.04	19.50	.016	0.02	0.20
Open-ended	0.12	0.04	25.05	.012	0.03	0.21
Oral retell	0.18	0.07	5.91	.039	0.01	0.35

Note. SE = Standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval.

TABLE 5

Meta-regression model for the relation between reading comprehension and different dimensions of written composition (writing productivity as the reference)

	b	SE	t	p	df	CI.LB	CI.UB
Intercept	0.23	0.04	5.75	.001	6.45	0.14	0.33
Writing quality	0.23	0.05	4.90	.001	7.41	0.12	0.34
Writing fluency	−0.15	0.07	−1.99	.117	4.08	−0.35	0.06
Writing conventions	0.32	0.09	3.40	.007	9.48	0.11	0.53
Writing syntax	0.18	0.07	2.64	.019	14.68	0.03	0.32
Writing vocabulary	0.16	0.10	1.60	.154	6.75	−0.08	0.39

Note. A total of 83 unique samples and 253 effect sizes. *SE* = standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval.

TABLE 6

Meta-regression model where grade levels predict word reading-spelling relation and reading comprehension-written composition relations

	b	SE	t	p	df	CI.LB	CI.UB
Word reading and spelling							
Intercept	0.82	0.02	45.08	<.001	228.80	0.78	0.85
Grades 3–5	0.02	0.03	0.48	.633	149.40	−0.05	0.08
Grades 6–12	−0.06	0.05	−1.09	.284	40.20	−0.17	0.05
University/adult	−0.13	0.05	−2.75	.008	55.40	−0.23	−0.04
Reading comprehension and written composition (productivity as the reference)							
Intercept	0.21	0.05	4.48	.006	5.15	0.09	0.33
Writing quality	0.21	0.05	4.48	.004	5.86	0.10	0.33
Writing fluency	−0.12	0.10	−1.19	.335	2.49	−0.49	0.25
Writing conventions	0.26	0.10	2.59	.026	10.30	0.04	0.48
Writing syntax	0.17	0.08	2.29	.041	12.30	0.01	0.34
Writing vocabulary	0.12	0.10	1.13	.295	7.40	−0.13	0.36
Grades 3–5	0.08	0.05	1.47	.159	17.76	−0.03	0.19
Grades 6–12	0.10	0.06	1.69	.106	20.06	−0.02	0.23
University/Adult	−0.04	0.08	−0.55	.591	19.61	−0.22	0.13

Note. Reference group is primary grades that include prekindergarten to grade 2. Word reading and spelling had 395 unique samples and 896 effect sizes that included grade-level information. Reading comprehension and written composition had 68 unique samples and 203 effect sizes that included grade-level information. *SE* = Standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval.

reading and spelling had the strongest relation compared to all the other pairs of relations ($ps < .001$).

Research Question 3: Relation of Reading Comprehension and Written Composition as a Function of Measurement

Table 3 shows bivariate correlations between different types of reading comprehension tasks and different dimensions of written composition. Note that a full matrix could not be estimated due to lack of data (e.g., no studies have examined the relation between writing conventions and reading comprehension measured by an open-ended reading comprehension task). In addition, written retell was excluded because only one study included it (Jenkins et al., 2014). Overall, there was large variation in the magnitudes, ranging from .12 ($p = .284$) between reading comprehension measured by a cloze task and writing fluency to .55 ($p = .002$) between reading comprehension measured by oral retell and writing quality. It should be noted that several of the estimates had a less than ideal number of degrees of freedom and limited number of studies (see Table 3 for details). Therefore, in the meta-regression, we examined only the relation between writing quality and reading comprehension measured by cloze, multiple choice, open-ended question, and oral retell tasks. As shown in Table 4, writing quality had a

stronger relation with reading comprehension measured by multiple choice ($r =$ .48, $p = .016$), open-ended questions ($r = .49$, $p = .012$), and oral retell ($r = .55$, $p = .039$) than by the cloze task ($r = .37$, $p < .001$; the reference task in Table 4).

With regard to the relation between reading comprehension (average across tasks) and different dimensions of written composition, the relation between reading comprehension and writing productivity was used as the reference. As shown in Table 5, reading comprehension was more strongly related with writing quality (.46, $p = .001$), writing syntax (.41, $p = .019$), and writing conventions (.55, $p = .007$) than with writing productivity (.23, $p = .001$). The relation of reading comprehension with writing vocabulary (.39) was also descriptively stronger than with writing productivity, but the difference did not reach the conventional statistical significance level ($p = .154$). Writing fluency (.08) had a descriptively weaker relation than writing productivity, but it was not statistically significant ($p = .117$).

Research Question 4: Moderation by Grade Level

Regression model results are shown in Table 6. The reference group was primary graders in prekindergarten to grade 2. The word reading and spelling relation was statistically significantly stronger for primary grade students ($r = .82$, $p < .001$) than for university students and adults ($r = .69$, $p = .008$). The relation between reading comprehension with various dimensions of written composition did not differ by grade levels ($ps > .106$; see the bottom panel).

Sensitivity Analysis and Publication Bias

The analyses of overall effect size and relations by grain size (Table 1) were also conducted using *metafor* (Viechtbauer, 2010). The R package, *robumeta*, uses method of moments estimation while *metafor* uses restricted maximum likelihood (REML) estimation. *Metafor* showed highly similar effect sizes for all relations.

Second, publication status was examined to test for publication bias. Publication bias is when published studies have effects with significantly different magnitudes, typically larger and at a significant level, than unpublished (gray literature) studies (Thornton & Lee, 2000). Publication bias was tested for the overall relation between reading and writing and for the relations between reading and writing subskills (e.g., word reading, spelling). First, to test the overall relation, a funnel plot that showed effect size distribution around the mean was printed (Figure 2). Also, a mixed effects regression test for funnel plot asymmetry (Egger et al., 1997; Sterne & Egger, 2005) was run. The statistical model found evidence of publication bias ($z = 3.40$, $p = 0.00$). Next, a meta-regression was conducted to test whether the overall relation between reading and writing varied as a function of publication status. No statistical difference between published effects sizes ($k = 1,506$) and unpublished effects sizes ($k = 759$) was found ($p = .91$; see Table 7). Then publication bias was tested for the reading and writing relations by subskill. Funnel plots of effect size distribution for the six models were printed (Figure 3). Finally, mixed effects regression tests for funnel plot asymmetry were run. All tests showed no evidence of publication bias at the conventional significance level ($ps = .17–.82$).
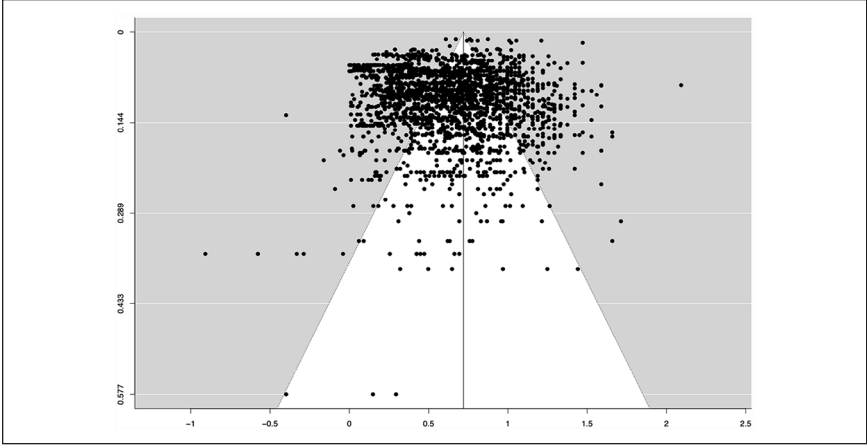
FIGURE 2. *Funnel plot of all effect sizes of the relation between reading and writing*

We also tested whether estimates were impacted by reliability of reading and writing measures (whether or not reliability met the threshold of .80, which is generally accepted as high reliability). A total of 1,244 effect sizes were available for reliability estimates for reading tasks, of which 199 were below .80, and a total of 1,298 effect sizes were available for reliability estimates for writing tasks, of which 176 were below .80. As shown in Table 7, both meta-regressions yielded statistically significantly weaker relations between reading and writing when a measure did not meet the .80 threshold ($p < .003$).

Sensitivity analysis was also conducted considering normed nature of tasks. Of 2,265 effect sizes, 1,202 effect sizes were based on normed reading measures and 885 effect sizes were based on normed writing measures. As shown in Table 7, normed measures yielded statistically significantly stronger relations (reading: non-normed: $r = .67, p < .001$; normed: $r = .78, p < .001$; writing: non-normed: $r = .67, p < .001$; normed: $r = .80, p < .001$).

Furthermore, we reran the overall analysis using only studies that met reliability estimates of .80 or above and that used normed measures. This subsample had sufficient effect sizes—364 effect sizes (126 unique samples)—and results were as follows: .79 ($p < .001$, 95% CI [.74, .84]) for the overall relation between reading and writing; .92 ($p < .001$, 95% CI [.86, .98]) for word reading and spelling; .56 ($p < .001$, 95% CI [.47, .66]) for word reading and written composition; .40 ($p < .001$, 95% CI [.36, .44]) for reading comprehension and written composition; .63 ($p < .000$, 95% CI [.57, .68]) for reading comprehension and spelling; and .78 ($p < .001$, 95% CI [.63, .93]) for text reading fluency and spelling. The relation between text reading fluency and written composition was not estimated due to lack of sufficient effect sizes for this analysis.

In addition, primary grades were disaggregated to preschool and kindergarten, and grades 1 and 2, and secondary school grades were disaggregated to grades 6–8 and grades 9–12 (see the bottom panels of Table 7). No

17

TABLE 7

Sensitivity analysis: Examining publication status, reliability, normed nature of reading and writing measures, and grade levels further disaggregated.

	b	SE	t	p	df	CI.LB	CI.UB
Publication status (unpublished as the reference)							
Intercept	0.72	0.02	32.41	<.001	172.00	0.68	0.77
Published	0.00	0.03	0.12	.908	301.00	−0.05	0.05
Reliability for reading measures (.80 or higher as the reference)							
Intercept	0.70	0.01	48.58	<.001	345.90	0.67	0.73
Less than .80	−0.16	0.05	−3.16	.003	39.30	−.27	−0.06
Reliability for writing measures (.80 or higher as the reference)							
Intercept	0.70	0.01	48.00	<.001	331.30	0.68	0.73
Less than .80	−0.21	0.05	−4.65	<.001	52.70	−0.30	−0.12
Reading measures normed or not (not normed as the reference)							
Intercept	0.67	0.02	40.89	<.001	306	0.63	0.70
Normed	0.11	0.02	4.64	<.001	615	0.06	0.15
Writing measures normed or not (not normed as the reference)							
Intercept	0.67	0.01	46.35	<.001	384	0.64	0.70
Normed	0.13	0.02	5.42	<.001	555	0.08	0.17
Primary grades (prekindergarten and kindergarten as the reference)							
Intercept	0.81	0.03	27.35	<.001	27.35	0.75	0.87
Grades 1–2	−0.05	0.04	−1.44	.154	81.80	−0.12	0.02
Higher grades (grades 6–8 as the reference)							
Intercept	0.61	0.04	14.91	<.001	35.3	0.53	0.69
Grades 9–12	−0.05	0.06	−0.80	.429	35.1	−0.18	0.08

Note. Publication status had 646 unique samples and 2,264 effect sizes. Reliability for reading measures had 375 unique samples and 1,241 effect sizes that included reliability estimates. Reliability for writing measures had 374 unique samples and 1,296 effect sizes that included reliability estimates. A total of 1,202 reading measures were normed and 885 writing measures were normed. Early elementary coding had 250 unique samples and 761 effect sizes that included reliability estimates; reference group is prekindergarten and kindergarten. *SE* = Standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval.

significant difference was found in the overall relation between reading and writing between preschool and kindergarten ($r = .81$, $p < .000$) and grades 1 and 2 ($r = .76$, $p = .154$) or between grades 6–8 ($r = .61$, $p < .000$) and grades 9–12 ($r = .56$, $p = .429$).

We also examined differences by writing systems, specifically alphabetic versus morphosyllabic (i.e., Chinese; see Table 8). For the overall relation between reading and writing, we found a significant difference between alphabetic languages ($r = .73$, $p < .001$) and Chinese ($r = .62$, $p = .009$). We also examined relations by grain size. This revealed a statistically significant difference for the relation between word reading and spelling between alphabetic languages ($r =$
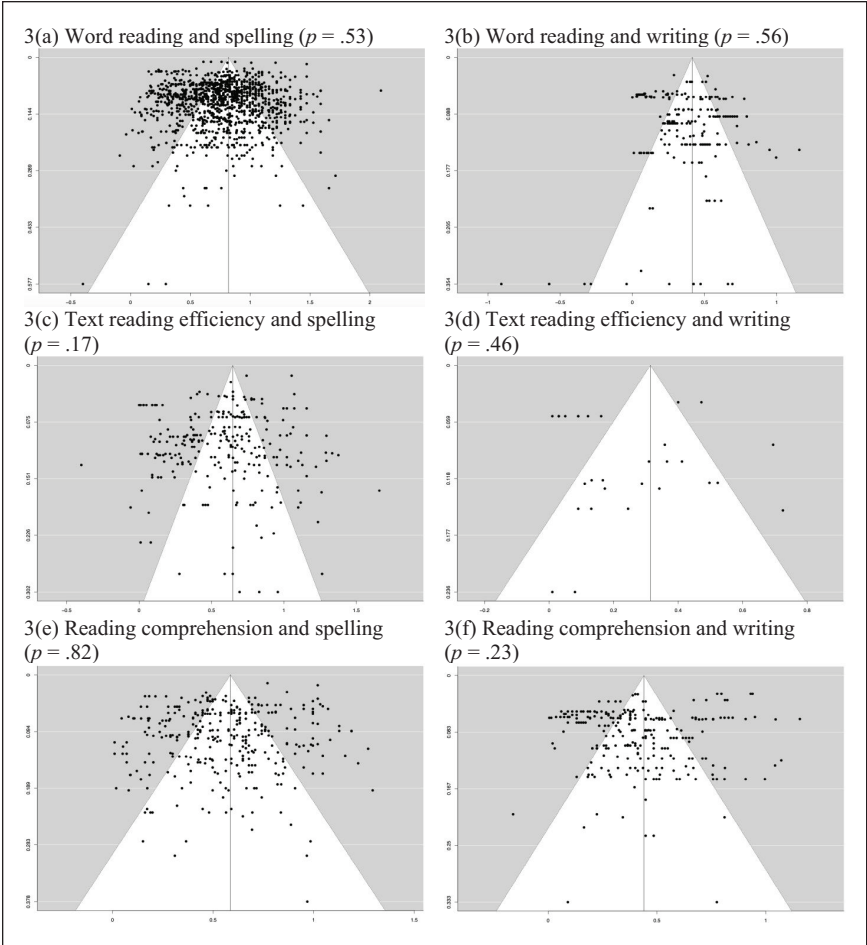
3(a) Word reading and spelling (*p* = .53)

3(b) Word reading and writing (*p* = .56)

3(c) Text reading efficiency and spelling (*p* = .17)

3(d) Text reading efficiency and writing (*p* = .46)

3(e) Reading comprehension and spelling (*p* = .82)

3(f) Reading comprehension and writing (*p* = .23)

FIGURE 3. *Funnel plots of relations between skills of reading and writing*

.83, $p < .001$) and Chinese ($r = .71$, $p = .023$). None of the other relations were statistically different.

Finally, we examined whether the magnitude of the relations differed by participants' disability status (see Table 9). We used a dichotomous variable (1 = at least 50% of the participants were reported as having a disability; 0 = less than 50% of participants were reported as having a disability). For the overall relation, we found no significant difference between typically developing participants ($r = .72$, $p < .001$) and participants with a disability ($r = .77$, $p = .152$). In addition, when the relation was examined by grain size, no statistical difference was found.

19

TABLE 8

Sensitivity analysis: Examining difference in effect sizes between Chinese and other languages (i.e., alphabetic orthographies, the reference group)

	b	SE	t	p	df	CI.LB	CI.UB
Overall relation							
Intercept	0.73	0.01	59.73	<.001	594.10	0.71	0.75
Chinese	−0.11	0.04	−2.75	.009	42.80	−0.20	−0.03
Word reading and spelling							
Intercept	0.83	0.01	60.8	<.001	478.30	0.80	0.85
Chinese	−0.12	0.05	−2.4	.023	31.10	−0.22	−0.02
Word reading and written composition							
Intercept	0.43	0.03	14.59	<.001	44.89	0.37	0.49
Chinese	−0.09	0.06	−1.53	.172	6.52	−0.23	0.05
Text reading fluency and spelling							
Intercept	0.67	0.03	24.40	<.001	118.80	0.62	0.72
Chinese	−0.24	0.15	−1.52	.199	4.24	−0.66	0.19
Text reading fluency and written composition							
Intercept	0.66	0.03	24.19	<.001	120.63	.60	0.71
Chinese	−0.23	0.16	−1.46	.215	4.23	−.65	.20
Reading comprehension and spelling							
Intercept	0.59	0.02	32.91	<.001	200.09	0.56	0.63
Chinese	−0.13	0.07	−1.96	.08	9.84	−0.28	0.02
Reading comprehension and written composition							
Intercept	0.44	0.03	17.47	<.001	67.60	0.39	0.49
Chinese	0.03	0.06	0.44	.665	15.40	−0.11	0.16

Note. The current meta-analysis included 121 effect sizes from Chinese speakers and 2,144 effect sizes from speakers of other languages. *SE* = Standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB = upper bound of 95% confidence interval.

Discussion

In the present study, we examined the relation between reading and writing to capture the overall relation but more importantly to capture variation of the relations, informed by the interactive dynamic literacy model (Kim, 2020a, 2022), depending on the linguistic grain size, measurement of reading comprehension (reading comprehension tasks) and written composition (dimensions of written composition), and developmental phase proxied by grade levels. After extensive and systematic search, a total of 395 studies that consisted of 612 unique samples, 2,265 effect sizes, and 120,669 individuals were included in the present meta-analysis.

Overall, we found that reading and writing skills are strongly related with a magnitude of $r = .72$. These results indicate that reading and writing are not independent skills. While reading-writing relations have been widely recognized (e.g., Fitzgerald & Shanahan, 2000; Kim, 2020a; Langer & Flihan, 2000; Shanahan,

TABLE 9

*Sensitivity analysis: Examining difference in effect sizes between students with and
without a disability*

	b	SE	t	p	df	CI.LB	CI.UB
Overall relation							
Intercept	0.72	0.01	57.17	<.001	559.60	0.69	0.74
Disability	0.05	0.04	1.45	.152	91.10	−0.02	0.12
Word reading and spelling							
Intercept	0.82	0.01	57.53	<.001	440.90	0.79	0.85
Disability	0.02	0.04	0.55	.582	84.90	−0.05	0.09
Word reading and written composition							
Intercept	0.41	0.03	14.51	<.001	44.77	0.35	0.47
Disability	0.11	0.06	1.77	.130	5.67	−0.04	0.27
Text reading fluency and spelling							
Intercept	0.65	0.03	22.73	<.001	113.50	0.59	0.71
Disability	0.14	0.10	1.47	.169	11.30	−0.07	0.35
Text reading fluency and written composition							
Intercept	0.64	0.03	22.56	<.001	115.3	0.58	0.69
Disability	0.16	0.10	1.61	.135	11.2	−0.06	0.37
Reading comprehension and spelling							
Intercept	0.58	0.02	31.86	<.001	185.50	0.54	0.61
Disability	0.09	0.06	1.35	.187	29.80	−.04	0.22
Reading comprehension and written composition							
Intercept	0.44	0.02	18.62	<.001	74.39	0.40	0.49
Disability	−0.01	0.09	−0.17	.875	4.83	−0.24	0.21

Note. If 50% or more of the sample had a disability, it was coded as "Disability" ($n = 212$); if less
than 50% of the sample had a disability, it was coded as "Typically developing" ($n = 2,053$). *SE* =
Standard error; *df* = degrees of freedom; CI.LB = lower bound of 95% confidence interval; CI.UB
= upper bound of 95% confidence interval.

2016; Shanahan & Lomax, 1986), to our knowledge, the magnitude of their rela-
tions has not been systematically captured before. The present findings are in line
with and explain recent meta-analytic findings that students who have reading
difficulties also have writing difficulties (Graham et al., 2021) and reading instruc-
tion impacts writing outcomes (Graham et al., 2018) and writing instruction influ-
ences reading outcomes (Graham & Hebert, 2010).

Beyond the overall relations, however, results revealed different magnitudes of
relations as a function of linguistic grain size. Theoretically based focal contrast
of the magnitude was between word reading and spelling versus reading compre-
hension and written composition, and the former is hypothesized to have a stron-
ger relation than the latter according to the interactive dynamic literacy model
(Kim, 2020a). This hypothesis was supported such that reading-writing relations
differed depending on the linguistic grain size—word reading and spelling have a
very strong relation (.82) whereas reading comprehension and written

composition have a moderate relation (.44). Although not focal contrasts in the present study, we also found moderate to fairly strong relations between other subskills of reading and writing. For example, spelling was fairly strongly related to text reading fluency (.65) and reading comprehension (.59), and word reading was moderately related to written composition (.42).

Differential relations were also found as a function of measurement of reading comprehension and written compositions. Reading comprehension is measured by various tasks, and studies have shown that tasks vary in the extent to which they tap into shallow (literal) and deep (inferential) comprehension (Francis et al., 2005; Keenan et al., 2008; Nation & Snowling, 1997). Similarly, written composition is evaluated in multiple dimensions, such as quality, productivity, fluency, vocabulary, syntax, and conventions (see earlier discussion). Because these dimensions differ in their focal construct and the skills they primarily tap into (Kim et al., 2014, 2015), the magnitude of their relations with reading comprehension was expected to differ (Kim, 2020a). We found support for the hypothesis. Writing quality had a stronger relation with reading comprehension measured by multiple choice and open-ended tasks than with reading comprehension measured by cloze tasks. In addition, across reading comprehension tasks, reading comprehension had moderate relations with writing quality, writing vocabulary, writing syntax, and writing conventions and had weak relations with writing productivity and writing fluency. These results indicate that reading comprehension–written composition relations are not uniform but instead multifaceted. Although the average relation between reading comprehension and written composition is moderate (.44), this is a summary of the relations across different reading comprehension tasks and different dimensions of written composition, and relations vary depending on focal dimensions of written composition and measurement of reading comprehension.

Findings of the present study also underscore the importance of measurement quality for reading-writing relations. Sensitivity analysis revealed that reading-writing relations were stronger when tasks met the criterion of high reliability (.80) and tasks were normed (vs. experimental). When reading-writing relations were examined using a subsample of effect sizes that met both criteria of high reliability and normed tasks, magnitudes were stronger. For example, the relation between word reading and spelling became .92 compared to .82, and the relation between word reading and written composition became .56 from .42. These results are in line with the fact that measurement error attenuates correlations (Spearman, 1904). Interestingly, the magnitude between reading comprehension and written composition showed a different pattern such that the relation did not become stronger: .40 when including effect sizes that met both criteria of high reliability and normed tasks versus .44 when including all effect sizes. Reasons for this is unclear and future studies are warranted to carefully examine why the patterns are different for the relation between reading comprehension and written composition. Overall, these findings indicate that reliability and normed nature of tasks influence the strength of correlations and therefore should be taken into consideration in understanding reading-writing relations.

Reading-writing relations also differed as a function of the developmental phase of reading and writing development. We hypothesized a different pattern

for word reading and spelling versus reading comprehension and written composition because developmental trajectories of these skills differ—majority of individuals reach asymptotes for lower-level reading and writing skills, word reading and spelling, at a later developmental phase whereas this is not expected for higher-order skills, reading comprehension and written composition. As such, word reading and spelling were expected to have a weaker relation in the later phase of development than in the beginning phase of development whereas no differences as a function of developmental phase were hypothesized for reading comprehension and written composition. This hypothesis was supported in that word reading and spelling had a stronger relation for primary grade students (.82) than that for university students and adults (.69), but the relation between reading comprehension and written composition did not differ as a function of developmental phase. These results revealed that the reading-writing relation is not constant across developmental phases, and therefore, development should be taken into account for understanding the reading-writing relation.

It is notable that the relation between word reading and spelling was stronger in languages with alphabetic writing systems than in Chinese, which employs the morphosyllabic writing system, indicating that reading/decoding and spelling/encoding words are not as strongly associated in Chinese compared to alphabetic languages. These results might be attributed to the visual complexity of Chinese characters compared to the vast majority of alphabetic languages. Chinese characters are composed of a complex array of strokes: Chinese characters can have up to 24 strokes although the majority of them have 6 to 13 strokes (Anderson et al., 2013). Visual complexity leads to difficulty in learning (Chang et al., 2016). In addition, strokes form different types of radicals such as phonological and sematic ones, and radicals have different positional characteristics and regularities and have variants (Ye & McBride, 2022). Therefore, accurate spelling/encoding of Chinese characters and words would require higher precision in visual and orthographic knowledge than for decoding to a greater extent than for alphabetic languages. In other words, the relative challenge of accurate spelling/encoding compared to decoding would be greater in Chinese than alphabetic languages. This, in turn, might result in more varied performance differences between word reading and spelling, which would lead to a weaker relation. Future studies are needed to investigate this speculation.

Overall, the findings of the present meta-analysis indicate the importance of recognizing the multifaceted nature of reading-writing relations. Although reading-writing relations have been widely recognized in previous work, its multifaceted nature of the relation has been rarely recognized. Overall, the results are in line with the hypothesis of the interactive dynamic literacy model that recognizes that reading-writing relations are not uniform, and they suggest our conceptualization of reading-writing relations should consider multiple aspects such as linguistic grain size, measurement, and developmental phase.

The results in the present study are correlational, and therefore, causal inference is limited. However, together with causal evidence on the relations between reading and writing from prior meta-analyses (Graham & Hebert, 2010; Graham et al., 2018, 2021), there are a few practical implications. One apparent

implication for reading-writing relations is that reading and writing acquisition will be enhanced when they are taught in an integrated manner rather than teaching them as separate entities (Graham, 2020; Kim, 2020a; Shanahan, 1988). Importantly, our findings of different magnitudes of relations by linguistic grain sizes offer a valuable nuanced and fine-grained picture. A very strong relation between word reading and spelling implies that for a majority of students, their word reading skill and spelling skill will be convergent such that those who are strong in word reading are highly likely to be strong in spelling, and those with weak word reading will highly likely experience weak spelling skill, and that only a small number of individuals show a discrepant profile in word reading and spelling skills. In contrast, the relation between reading comprehension and written composition is moderate such that there is greater divergence in individuals' skills in reading comprehension and written composition skills, and therefore, it is not uncommon for individuals with strong reading comprehension but weak written comprehension, and vice versa.

Another way of thinking about varying strengths of reading-writing relations is different degrees of transfer of skills. Stronger relations indicate higher likelihood of transfer of skills. This implies that what is learned in the context of word reading is highly likely to transfer to spelling, and vice versa, whereas degree of transfer is less likely between reading comprehension and written composition. Therefore, for individuals to benefit from reading-writing connections, reading-writing connections need to be made visible in teaching (Kim, 2022; Shanahan, 1988), and this is particularly important for reading comprehension and written composition. For example, students may not readily see how understanding authors' use of vocabulary, sentence structures, and rhetorical moves in comprehension applies to their own composition process. Therefore, instruction needs to make the reading-writing links explicit and visible, discussing how these aspects can inform and apply to students' own writing (e.g., relating analysis of an author's use of vocabulary for a specific audience to thinking about vocabulary choice in the students' own compositions).

The moderate relation between reading comprehension and written composition, compared to that between word reading and spelling, also implies a greater need for separate reading-focused instruction and writing-focused instruction for successful development of reading comprehension and written composition in addition to integrated instruction. As noted previously in the literature review, reading comprehension and written composition differ in the processes and the extent to which skills and knowledge are tapped, and therefore, instruction should address reading-specific and writing-specific needs.

Differential relations between reading comprehension and written composition as a function of tasks and focal dimensions point to the importance of using multiple tasks to accurately measure reading comprehension and written composition (Francis et al., 2005; Kim, 2020a, 2022; Stuhlmann et al., 1999). Use of multiple tasks is an important way to reduce measurement error and increase precision. This is especially important for high-stakes decisions that demand high precision such as determining one's reading and writing disability status. We recognize that it is not always feasible to use multiple tasks in practice because using multiple tasks per construct takes more resources and has practical

constraints (e.g., administering multiple tasks per construct takes longer assessment time, which is not always possible in settings such as school contexts). In these cases, one should be aware of the nature of tasks and dimensions and their differential tapping of skills and knowledge, acknowledge limitations associated with using a particular type of task or a single task, and need to exercise caution in interpretations.

Limitations and Future Directions

Future work is warranted to investigate the mechanisms of their relations. One example is directionality of reading-writing relations. Theoretically, reading-writing is hypothesized to have a bidirectional relation via reading and writing experiences (Kim, 2020a). For example, the experience of word reading is hypothesized to enhance spelling and the experience of spelling enhances word reading as word reading and spelling provide an opportunity to attend to phonological, orthographic, and morphological structure of words. Similar logic applies to reading comprehension and written composition. However, empirical evidence on bidirectionality of the relation is mixed (e.g., Ahmed et al., 2014; Kim et al., 2018). Another future direction is that if reading and writing are related, integrated reading-writing instruction should benefit both reading and writing outcomes (see Graham et al., 2017). However, the reading-writing relations are not perfect, and there are reading-specific aspects and writing-specific aspects that require instructional attention. This is particularly the case for reading comprehension and written composition. As noted previously, a moderate relation suggests that greater explicit instructional attention might be needed for students to see and benefit from the connection between reading comprehension and written composition. Future experimental work is necessary to expand our understanding of effective ways of integrating reading and writing instruction that promotes transfer between reading skills and writing skills (Shanahan, 2016), and effective instruction that addresses reading- and writing-specific aspects.

The present findings also indicate a need for systematically considering measurement of constructs and their implications to understand reading-writing relations more precisely (e.g., Kim, 2020a, 2022). As noted previously, measurement matters, especially for complex and multidimensional constructs such as reading comprehension and written composition. In future empirical work, it will be ideal if studies, across correlational and intervention work, investigate and report results for multiple dimensions of reading comprehension and written composition to further our nuanced understanding of reading-writing relations. Also needed is careful systematic attention to precision in measurement. Theoretical models typically assume perfect measurement properties (e.g., reliability), but, in reality, constructs are measured with measurement error. As noted previously, one way to address this is measuring a construct using multiple tasks and employing a latent variable approach in analysis to the extent feasible.

Conclusion

The results overall indicate that reading and writing skills are related. Beyond the overall average relation, the present results unpacked the nature of

reading-writing relations and extend our understanding of reading-writing relations by adding important nuances about their multidimensional nature.

Appendix: A Complete List of Search Terms

"Writ* Skills" OR "Written Compos*" OR "Compos*" OR "Writ* Process" OR "writ* routines" OR "writ* goals" OR "Writ* Tools" OR "writ* feedback" OR "writ* knowledge" OR spell* OR "CBM writ*" OR CIWS OR CWS OR WW OR "Word* written" OR "Word* spelled correctly" OR "Correct word sequences" OR "Incorrect word sequences" OR "Correct minus incorrect word sequences" OR dysgraphia OR "Writ* difficult*" OR "Writing disab*"

AND

"word read*" OR "word read* fluency" OR "list read* fluency" OR "read* fluency" OR "oral reading fluency" OR "text read* fluency" OR Decod* OR "Decod* fluency" OR "word attack" OR "Low read*" OR "low skill* read*" OR dyslexia OR DYSLE* OR "read* disab*" OR "read* comprehension" OR "comprehension" OR "special education" OR "tier 2" OR "tier 3" OR "struggl* read*" OR "learning disab*" OR "severe learning disab*" OR "specific learning disab*" OR "co-morbid*" OR "neurodevelopmental disorder" OR "struggling read*" OR "weak read*" OR "poor read*"

ORCID iD

Young-Suk Grace Kim [iD] https://orcid.org/0000-0002-4328-3843

References

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology*, *85*, 478–508.

Adams, M. A. (1990). *Beginning to read: Thinking and learning about print*. MIT Press.

Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology*, *106*, 419–434. https://doi.org/10.1037/a0035692

Anderson, R. C., Ku, Y. M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to see the patterns in Chinese characters. *Scientific Studies of Reading*, *17*(1), 41–56. https://doi.org/10.1080/10888438.2012.689789

Babayigit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive-linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology*, *103*(1), 169–189. https://doi.org/10.1037/a0021671

Bahr, R. H., Silliman, E. R., Berninger, V. W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1–9. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1587–1599. https://doi.org/10.1044/1092-4388(2012/10-0335)

Berman, R., & Verhoevan, L. (2002). Cross-linguistic perspectives on the development of text- production abilities. *Written Language and Literacy*, *5*, 1–43. https://doi.org/10.1075/wll.5.1.02ber

Berninger, V. W., & Abbott, R. D. (2010). Discourse-level oral language, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology*, *102*(3), 635–651. https://doi.org/10.1037/a0019319

Berninger, V. W., Abbott, R. D., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S.-J.(C.), Gould, L., Youngstrom, M., Shimada, S., & Amtmann, D. (2010). Relationship of word- and sentence-level working memory to reading and writing in second, fourth, and sixth grade. *Language, Speech, and Hearing Services in Schools*, *41*(2), 179–193. https://doi.org/10.1044/0161-1461(2009/08-0002)

Berninger, V. W., Nielson, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, *46*, 1–21. https://doi.org/10.1016/j.jsp.2006.11.008

Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). Guilford Press.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31–42. http://dx.doi.org/10.1037/0022-0663.96.1.31

Cao, Y., & Kim, Y.-S. G. (2021). Is retell a valid measure of reading comprehension? *Educational Research Review*, *32*, 100375. https://doi.org/10.1016/j.edurev.2020.100375

Chall, J. S. (1983). *Stages of reading development*. McGraw–Hill.

Chang, L.-Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading*, *20*, 64–85. https://doi.org/10.1080/10888438.2015.1104688

Coker, D. (2006). Impact of first-grade factors on the growth and outcomes of urban school children's primary-grade writing. *Journal of Educational Psychology*, *98*, 471–488. https://doi.org/10.1037/0022-0663.98.3.471

Coker, D. L., Ritchey, K. D, Uribe-Zarain, X., & Jennings, A. (2018). An analysis of first-grade writing profiles and their relationship to compositional quality. *Journal of Learning Disabilities*, *51*, 336–350. https://doi.org/10.1177/002221941770817

Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C., & Bivens-Taum, J. (2008). Cognitive models of writing: Writing proficiency as a complex integrated skill. *ETS Research Report Series, 2*, i–36. Downloaded from https://www.ets.org/Media/Research/pdf/RR-08-55.pdf.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 237–269). Lawrence Erlbaum Associates Publishers.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, *35*, 39–50. https://doi.org/10.1207/S15326985EP3501_5

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*(4), 553–576.

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 369–394). Erlbaum

Furnes, B., Elwér, Å., Samuelsson, S., Olson, R. K., & Byrne, B. (2019). Investigating the double-deficit hypothesis in more and less transparent orthographies: A longitudinal study from Preschool to Grade 2. *Scientific Studies of Reading*, *23*(6), 478–493. https://doi.org/10.1080/10888438.2019.1610410

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74–111.

Graham, S. (2020). The sciences of reading and writing must become more fully integrated. *Reading Research Quarterly*, *55*(S1), S35–S44. https://doi.org/10.1002/rrq.332

Graham, S., Aitken, A. A., Hebert, M., Camping, A., Santangelo, T., Harris, K. R., Eustice, K., Sweet, J. D., & Ng, C. (2021). Do children with reading difficulties experience writing difficulties? A meta-analysis. *Journal of Educational Psychology*, *113*(8), 1481–1506. https://doi.org/10.1037/edu0000643

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323–344). The Guilford Press.

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*, 669–686.

Graham, S., Harris, K., & Hebert, M. A. (2011). *Informing writing: The benefits of formative assessment. A Carnegie Corporation Time to Act report*. Alliance for Excellent Education.

Graham, S., & Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading. A Carnegie Corporation Time to Act Report*. Carnegie Corporation of New York.

Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2017). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Review Research Quarterly*, *53*, 279–304. https://doi.org/10.1002/rrq.194

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research*, *88*, 243–284. https://doi.org/10.3102/0034654317746927

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, *27*, 1703–1743. https://doi.org/10.1007/s11145-014-9517-0

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Lawrence Erlbaum Associates.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. https://doi.org/10.1002/jrsm.5

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

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

Hooper, S. R., Swartz, C. W., Wakely, M. B., de Kruif, R. E. L., & Montgomery, J. W. (2002). Executive function in elementary school children with and without problems in written expression. *Journal of Learning Disabilities*, *35*, 57–68.

Howell, K., & Nolet, V. (2000). *Curriculum-based evaluation: Teaching and decision-making* (3rd ed.). Wadsworth Thomson Learning.

Jenkins, J. R., Johnson, E., & Hileman, J. (2014). When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading*, *8*(2), 125–151. http://dx.doi.org/10.1207/s1532799xssr0802_2

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, *12*(3), 281–300. https://doi.org/10.1080/10888430802132279

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, *1*(1), 1–26.

Kim, Y.-S. G. (2020a). Interactive dynamic literacy model: An integrative theoretical framework for reading and writing relations. In R. Alves, T. Limpo, & M. Joshi (Eds.), *Reading-writing connections: Towards integrative literacy science* (pp. 11–34). Springer. https://doi.org/10.1007/978-3-030-38811-9_2.

Kim, Y.-S. G. (2020b). Structural relations of language and cognitive skills, and topic knowledge to written composition: A test of the direct and indirect effects model of writing (DIEW). *British Journal of Educational Psychology*, *90*(4), 910–932. https://doi.org/10.1111/bjep.12330

Kim, Y.-S. G. (2020c). Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER). *Journal of Educational Psychology*, *112*(4), 667–684. http://dx.doi.org/10.1037/edu0000407

Kim, Y.-S. G. (2022). Co-occurrence of reading and writing difficulties: The application of the interactive dynamic literacy model. *Journal of Learning Disabilities*, *55*(6), 447–464. https://doi.org/10.1177/00222194211060868

Kim, Y.-S., Al Otaiba, S., Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the dimensionality of first grade written composition. *Journal of Speech, Language, and Hearing Research*, 57, 199–211. https://doi.org/10.1044/1092-4388(2013/12-0152)

Kim, Y.-S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences*, *21*, 517–525. https://doi.org/10.1016/j.lindif.2011.06.004.

Kim, Y.-S., Al Otaiba, S., Wanzek, J., & Gatlin, B. (2015). Towards an understanding of dimension, predictors, and gender gaps in written composition. *Journal of Educational Psychology*, *107*, 79–95. https://doi.org/10.1037/a0037210

Kim, Y.-S. G., & Graham, S. (2022). Expanding the direct and indirect effects model of writing (DIEW): Dynamic relations of component skills to various writing outcomes. *Journal of Educational Psychology*, *114*(2), 215–238. https://doi.org/10.1037/edu0000564

Kim, Y.-S. G., Petscher, Y., Wanzek, J., & Al Otaiba, S. (2018). Relations between reading and writing: A longitudinal examination from Grades 3 to 6. *Reading and Writing: An Interdisciplinary Journal*, *31*, 1591–1618. https://doi.org/10.1007/s11145-018-9855-4

Kim, Y.-S. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology, 109*, 35–50. https://doi.org/10.1037/edu0000129

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Sage.

Langer, J. A. (1986). Reading, writing, and understanding: An analysis of the construction of meaning. *Written Communication*, *3*(2), 219–267.

Langer, J. A., & Applebee, A. N. (1987). *How writing shapes thinking: A study of teaching and learning*. National Council of Teachers of English.

Langer, J. A., & Flihan, S. (2000). Writing and reading relationships: Constructive tasks. In R. Indrisano & J. R. Squire (Eds.), *Writing: Research/theory/practice* (pp. 112–129). International Reading Association.

Mackie, C., & Dockrell, J. E. (2004). The nature of written language deficits in children with SLI. *Journal of Speech, Language, and Hearing Research*, *47*, 1469–1483. https://doi.org/10.1044/1092-4388(2004/109)

Malpique, A. A., Pino-Pasternak, D., & Roberto, M. S. (2020). Writing and reading performance in year 1 Australian classrooms: associations with handwriting automaticity and writing instruction. *Reading and Writing*, *33*, 783–805.

Mazany, T., Pimentel, S., Orr, C. S., & Crovo, M. (2015). Reading framework for the 2015 National Assessment of Educational Progress National Assessment Governing Board. *Assessment*. Retrieved from https://www.nagb.org/content/nagb/assets/documents/publications/frameworks/reading/2015-reading-framework.pdf

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education*, *41*, 68–84.

Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading*, *13*(5), 359–382. https://doi.org/10.1080/10888430903162878

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, *67*(3), 359–370. https://doi.org/10.1111/j.2044-8279.1997.tb01250.x

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). U.S. Government Printing Office.

Olinghouse, N. G., & Graham, S. (2009). The relationship between discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology*, *101*, 37–50. https://doi.org/10.1037/a0013462

Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing*, *22*, 545–565. https://doi.org/10.1007/s11145-008-9124-z

Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, *33*(4), 335–355.

Puranik, C. S., Lombardino, L., & Altmann, L. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech Language Pathology*, *17*, 107–120.

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review*, *28*(2), 225–265. https://doi.org/10.1007/s10648-015-9335-1

Shanahan, T. (1980). The impact of writing instruction on learning to read. *Literacy Research and Instruction*, *19*(4), 357–368. https://doi.org/10.1080/19388078009557556

Shanahan, T. (1988). The reading-writing relationship: Seven instructional principles. *The Reading Teacher*, *41*(7), 636–647.

Shanahan, T. (2016). Relationships between reading and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 194–207). Guilford Press.

Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading–writing relationship. *Journal of Educational Psychology*, *78*(2), 116–123. https://doi.org/10.1037/0022-0663.78.2.116

Smith, C. (2011). *Narrative writing in native English and ESL learners: Developmental trajectories and predictors* (Publication No. NR77944) [Doctoral dissertation, University of Toronto]. ProQuest Dissertations Publishing. https://search.proquest.com/docview/916424661?accountid=14509

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72–101.

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. Rothstein, A. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 99–110). John Wiley & Sons.

Stuhlmann, J., Daniel, C., Delinger, A., Denny, R. K., & Powers, T. (1999). A generalizability study of the effects of training on teachers' abilities to rate children's writing using a rubric. *Journal of Reading Psychology*, *20*, 107–127.

Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology*, *53*(2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, *20*(3), 375. https://doi.org/10.1037/met0000011

Torppa, M., Georgiou, G. K., Niemi, P., Lerkkanen, M.-K., & Poikkeus, A.-M. (2017). The precursors of double dissociation between reading and spelling in a transparent orthography. *Annals of Dyslexia*, *67*(1), 42–62. https://doi.org/10.1007/s11881-016-0131-5

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing*, *24*, 203–220. https://doi.org/10.1007/s11145-010-9266-7

Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*(2), 272–277. https://doi.org/10.1037/0022-0663.94.2.272

Ye, Y., & McBride, C. (2022). A dynamic interactive model of Chinese spelling development. *Educational Psychology Review*, *34*, 2897–2917. https://doi.org/10.1007/s10648-022-09684-3

Authors

YOUNG-SUK GRACE KIM, Ed.D., is a professor and the senior associate dean in the School of Education at University of California, Irvine; email: *Youngsk7@uci.edu*. Her research focuses on development and effective instruction of language, cognition, reading, and writing skills for children from diverse linguistic, cultural, and socioeconomic backgrounds.

ALISSA WOLTERS, MA, is a doctoral student at the University of California, Irvine; email: *awolters@uci.edu*. Her research focuses on language and literacy development for children from diverse backgrounds, especially emergent bilinguals and second language learners.

JOONG WON LEE, MA, is a doctoral student at University of California, Irvine; email: *joongwl@uci.edu*. His research focuses on the development and instruction of language and literacy skills such as morphological awareness, vocabulary, and reading comprehension.