education sciences

MDPI

Systematic Review

Writing Development in Chinese: A Meta-Analysis

Jr-An Lin [1], Li-Jen Kuo [1,*], Lu Chen [2], Yanbing Chen [2] and Haoran Shen [3]

[1] Department of Teaching, Learning, and Culture, College of Education and Human Development, Texas A&M University, College Station, TX 77843, USA; jrlin0908@tamu.edu
[2] Department of Educational Psychology, College of Education and Human Development, Texas A&M University, College Station, TX 77843, USA; lu.chen@tamu.edu (L.C.); chenyb0902@tamu.edu (Y.C.)
[3] School of International Letters and Cultures, The College of Liberal Arts and Sciences, Arizona State University, Tempe, AZ 85287, USA; hshen124@asu.edu
* Correspondence: lijenkuo@tamu.edu

Abstract

Chinese written composition draws on transcription and oral language skills, but the strength and comparability of their associations remain unclear. Guided by the Simple View of Writing in Chinese, this meta-analysis synthesized 170 correlations from 13 peer-reviewed publications grouped into 10 sample-source clusters. Correlated-effects robust variance estimation was used to estimate the overall association and examine moderation by skill domain, grade level, and script context. Componential skills showed a moderate positive association with Chinese composition ($r = 0.375$, 95% CI [0.293, 0.451], $p < 0.001$), accompanied by substantial heterogeneity ($I^2 = 85.63\%$). Transcription ($r = 0.457$) and oral language skills ($r = 0.340$) were both positively associated with writing, but the difference in strength of these associations was not statistically significant. Neither grade level nor script context significantly moderated the associations, and the exploratory Domain × Grade interaction was nonsignificant. The positive overall association remained stable across analyses addressing shared sample sources, pooled-grade correlations, partial correlations, and source influence. Examination of the evidence also revealed overlapping samples, pooled-grade reporting, heterogeneous skill and composition assessments, and a lack of transcription effects from Simplified Chinese contexts. These findings provide correlational evidence relevant to a componential account of Chinese writing while identifying methodological priorities for more informative developmental and contextual comparisons.

Keywords: Chinese; writing composition; Simple View of Writing; meta-analysis

Academic Editor: Meihua Liu

Received: 5 September 2026
Revised: 15 November 2026
Accepted: 20 November 2026
Published: 24 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Writing is a complex literacy process that requires multiple skill sets to operate together, including idea generation, thought organization, translating ideas into language, and producing written symbols under the constraints of attention and working memory (Hayes, 1996, 2012; McCutchen, 2000). Beyond language learning, writing provides a platform for learners to demonstrate knowledge, reasoning, and communication across academic domains (Graham, 2019; Klein & Boscolo, 2016; Langer & Applebee, 1987). Therefore, it is crucial to understand writing development by examining how different componential skills are associated with written composition.

The Simple View of Writing (SVW) provides a useful framework for examining these relationships by emphasizing transcription and higher-level language and idea-generation processes (Berninger et al., 2002; Berninger & Winn, 2006). Drawing on this componential perspective, the present meta-analysis uses transcription skills and oral language skills as two broad operational domains in Chinese writing. Transcription skills encompass character production and orthographic production skills, including copying, stroke-sequence knowledge, word spelling, and handwriting fluency. Oral language skills serve as an umbrella term for linguistic and metalinguistic abilities, including phonological awareness, vocabulary knowledge, morphological awareness, syntactic awareness, definitional skills, and oral narrative or ideation skills. These domains represent broad classifications of the assessed constructs rather than homogeneous or mutually exclusive processes, and not all measures within the oral language domain directly assess ideation.

Although both Chinese and alphabetic writing require the coordination of transcription, language, and discourse-level skills, the orthographic and linguistic characteristics of Chinese introduce specific demands for written composition. Chinese employs a morphosyllabic writing system in which characters generally represent morphemes associated with spoken syllables. Unlike alphabetic writing systems, Chinese character production requires learners to coordinate strokes and radicals organized within visually and spatially complex configurations, drawing on knowledge of character structures, radical positions, and conventional stroke sequences (Law et al., 1998; Wang et al., 2014; Yeung et al., 2017). Furthermore, the high density of homophones in Chinese means that a single spoken syllable may correspond to multiple characters with distinct meanings, making orthographic and morphological knowledge particularly relevant to selecting appropriate written forms (Kuo et al., 2026). These characteristics provide a rationale for examining how componential skills, particularly transcription and oral language skills, are associated with composition in Chinese rather than assuming that findings derived primarily from alphabetic writing systems generalize directly.

Previous research has identified associations between Chinese writing performance and transcription skills, including character writing, word spelling, and handwriting fluency (Wang et al., 2014; Yan et al., 2012; Yeung et al., 2013b, 2017). Similarly, oral language skills, including vocabulary knowledge, morphological awareness, syntactic awareness, and oral narrative ability, have been associated with Chinese written composition (Guan et al., 2014; Sun et al., 2018; Tong & McBride, 2016). Together, these findings suggest that both domains are relevant to Chinese writing development, although the strength of their associations and their variation across developmental and script contexts remain insufficiently synthesized.

Although previous research has examined various aspects of Chinese literacy instruction, including teachers' integration of technology (Lin et al., 2025), systematic synthesis of the associations between learners' componential skills and Chinese writing performance remains limited. Existing studies have often examined specific predictors, grade groups, or writing outcomes in isolation, making it difficult to assess the overall strength of the associations between componential skills and Chinese writing performance. In contrast, meta-analytic work on writing has more often focused on English or cross-linguistic samples, leaving Chinese writing underrepresented in broader evidence syntheses (Kent & Wanzek, 2016; Seoane et al., 2026).

Against this background, the present meta-analysis synthesizes the associations between transcription skills, oral language skills, and Chinese written composition and examines whether these associations vary across skill domains, grade levels, and script contexts. The following sections elaborate on the SVW and its application to Chinese

writing, followed by the theoretical rationale for examining skill domain, grade level, and script type as moderators.

1.1. A Simple View of Writing (SVW)

In English literacy research, the SVW has been used to conceptualize writing as a complex process involving the coordination of lower-level transcription skills and higher-level language and idea-generation processes (Berninger et al., 2002; Berninger & Winn, 2006). Within this framework, transcription encompasses lower-level procedural skills, particularly handwriting and spelling, required to translate linguistic representations into written text. In contrast, ideation refers to the generation and development of ideas and content, whereas text generation involves translating these ideas into linguistically organized expressions. Although theoretically distinct, these higher-level processes draw on linguistic and metalinguistic resources, including vocabulary knowledge, morphological awareness, syntactic awareness, and oral narrative skills (Kim et al., 2011; Kent & Wanzek, 2016). However, these linguistic resources should not be equated with ideation or text generation, as they may support written composition without directly assessing either process. Accordingly, the present meta-analysis groups these measures under the broader operational domain of oral language skills, recognizing that they represent different linguistic and metalinguistic abilities rather than interchangeable measures of ideation or text generation.

A central assumption of simple-view models is that writing is contingent upon the coordination of multiple componential skills rather than on any single skill alone. In this sense, transcription and oral language skills involving language formulation and ideation are both important for successful writing. A learner with strong handwriting and spelling but limited ideas, vocabulary, or syntactic resources may produce accurate but minimal or underdeveloped text. Conversely, a learner with well-developed oral language skills but weak transcription abilities may struggle to translate their thoughts into written text. Thus, the relationship among componential skills can be understood as constrained and interactive: weakness in one component may limit the extent to which strengths in another component can support writing performance.

Empirical studies in English writing have examined the associations between transcription, oral language, and writing performance. For example, Kim et al. (2011) found that oral language, spelling, and letter-writing fluency were uniquely related to beginning writing among kindergarten children. Similarly, Kent and Wanzek's (2016) meta-analysis showed that handwriting fluency, spelling, reading, and oral language were each positively related to writing quality and writing production. Research with older or struggling writers has also shown that transcription skills, including spelling accuracy and handwriting fluency, are associated with writing quality and productivity (Connelly et al., 2012). Extending this evidence to bilingual writing development, Ünal Gezer et al. (2025) conducted a three-year longitudinal study of Spanish–English emergent bilinguals and identified concurrent and longitudinal within- and cross-language associations between early literacy skills and writing outcomes. Together, these findings provide an empirical foundation for examining componential associations in writing across different linguistic contexts.

1.2. A Simple View of Writing in Chinese (SVWC)

The examination of Chinese composition has followed a similar componential path to English writing research, while paying attention to the unique linguistic and orthographic features of Chinese. Drawing on the SVW, Yeung et al. (2017) proposed the Simple View of Writing in Chinese (SVWC), arguing that Chinese writing quality is supported by multiple interacting components, particularly transcription and oral

language skills. In their model, transcription included stroke-sequence knowledge, word spelling, and handwriting fluency, whereas oral language included definitional, oral narrative, and syntactic skills. This framework is useful because it preserves the general logic of componential writing models while adapting the constructs to Chinese, a morphosyllabic and character-based writing system.

In Chinese, transcription involves the production of characters with complex orthographic structures, including conventional stroke sequences, radical configurations, and spatial arrangements. These characteristics highlight the relevance of character writing, copying skills, handwriting fluency, word spelling, and stroke-sequence knowledge to Chinese writing development (Law et al., 1998; Wang et al., 2014; Yan et al., 2012; Yeung et al., 2017). Beyond transcription, Chinese composition also draws on broader language resources that support meaning generation, sentence formulation, and discourse organization. Previous studies have identified associations between Chinese writing performance and vocabulary knowledge, morphological awareness, syntactic awareness, definitional skills, oral narrative skills, and other metalinguistic abilities (Guan et al., 2014; Sun et al., 2018; Tong & McBride, 2016; Yeung et al., 2017). Together, these findings highlight the relevance of both transcription and oral language skills to Chinese written composition.

1.3. Prominent Moderators in Chinese Composition Writing

Although the SVWC provides a useful framework for understanding Chinese writing development, the strength of the associations between componential skills and composition outcomes may vary across different contexts. Three moderators are especially relevant to the present meta-analysis: componential skill domain, grade level, and script type. Transcription and oral language skills may show different patterns of association with writing composition because they represent different dimensions of writing development. Furthermore, grade level may moderate the associations between componential skills and Chinese writing performance because the demands of transcription and higher-level language processes may vary across development. However, the extent to which their relative associations change across grades remains unclear (Ahmed et al., 2022; Kent & Wanzek, 2016). Last but not least, script type may also matter because Simplified and Traditional Chinese differ in visual-orthographic complexity, stroke density, and character forms, which may shape the transcription demands involved in learning to write Chinese (Hemelstrand et al., 2024; McBride et al., 2022). Therefore, the present meta-analysis examined whether the associations between componential skills and Chinese composition outcomes varied by skill domain, grade level, and script type.

1.3.1. Componential Skill Domain

The SVWC distinguishes between two broad componential skill domains: transcription skills and oral language skills. Although both domains have been identified as important to Chinese writing development, prior studies have often examined specific componential skills separately or within individual samples. As a result, it remains unclear whether transcription skills or oral language skills show stronger overall associations with Chinese composition. Comparing these two domains is therefore important for clarifying the relative roles of lower-level transcription processes and broader language resources in Chinese writing development.

1.3.2. Grade Levels

Grade level was examined as a moderator because the relative associations between componential skills and writing performance may shift across development. In early

writing, transcription skills may place heavy demands on learners because handwriting, copying, spelling, and character production are not yet fluent. When these lower-level skills require substantial attention, fewer cognitive resources may be available for planning, sentence construction, and idea elaboration (Ahmed et al., 2022; McCutchen, 2000). This issue may be especially relevant to Chinese writing because young learners must acquire character-specific transcription skills, including stroke-sequence knowledge, character copying, handwriting fluency, and word spelling (Law et al., 1998; Wang et al., 2014; Yeung et al., 2013b, 2017). However, these transcription demands may persist as learners progress through school, suggesting that transcription skills could maintain or strengthen their relative associations with Chinese writing performance across grades.

On the other hand, as transcription skills become more fluent, oral language skills may become relatively more strongly associated with writing performance, particularly as composition tasks require more elaborate vocabulary, greater syntactic complexity, and more effective discourse organization. Previous research suggests that the associations between oral language skills and writing performance may strengthen across grades (Seoane et al., 2026). However, this does not necessarily indicate that oral language skills become more strongly associated with writing than transcription skills across development. Moreover, existing studies have often focused on specific grade bands or individual developmental samples rather than systematically comparing these componential associations across grades. Given these competing developmental considerations and the limited comparative evidence, it remains unclear whether and how the relative associations between transcription and oral language skills vary across grade levels. Therefore, as for the present study, no directional hypothesis was specified for the exploratory Skill Domain × Grade interaction.

1.3.3. Script Types

Script type was examined as a moderator because Chinese writing is used in both Simplified and Traditional forms across different Chinese-speaking contexts. Simplified Chinese is commonly used in Mainland China, Singapore, and Malaysia, whereas Traditional Chinese is used in Taiwan, Hong Kong, and Macau (Chen, 1999; Taylor & Taylor, 2014). Although both scripts belong to the same morphosyllabic writing system, they differ in character forms, stroke density, and visual-orthographic complexity. Previous research has shown that character properties influence orthographic acquisition (Kuo et al., 2014, 2024), suggesting that these differences may shape the transcription demands associated with character production, copying, and handwriting fluency (Hemelstrand et al., 2024; McBride et al., 2022). However, most empirical studies have examined Chinese writing within a single script context, leaving it unclear whether the associations between componential skills and writing performance differ across Simplified and Traditional Chinese. In addition, script type is closely associated with geographical and educational contexts, including differences in curricula, participant characteristics, and assessment practices, making it difficult to isolate the influence of orthographic characteristics. Therefore, the present meta-analysis examined script type as a contextual moderator to explore potential variation in the associations between componential skills and Chinese writing performance, rather than as a direct measure of orthographic complexity.

2. The Present Study

Given the importance of writing development as a core indicator of literacy outcomes and the unique linguistic and orthographic features of Chinese writing, the present study addresses the limited systematic synthesis of this literature by conducting a componential meta-analysis. Guided by the SVWC, this study synthesizes existing correlational

evidence on the associations between componential skills and Chinese writing performance.

Beyond synthesizing correlational evidence relevant to the SVWC, the present study examines whether the associations between componential skills and Chinese writing performance vary across skill domains, grade levels, and script contexts. Although the findings may inform broader theories of writing development, the present meta-analysis does not directly compare Chinese writing with other writing systems. On the other hand, as discussed in Section 1.3.2, competing developmental considerations and limited comparative evidence did not provide a sufficient basis for specifying a directional hypothesis, so the interaction was examined exploratorily. The following research questions guided the study:

Research Questions

1. What are the overall pooled correlations between componential skills and Chinese writing performance?
2. Do the associations with Chinese writing performance differ significantly between transcription skills and oral language skills?
3. Does grade level moderate the associations between componential skills and Chinese writing performance?
4. Does script type, specifically Simplified Chinese versus Traditional Chinese, moderate the associations between componential skills and Chinese writing performance?

3. Method

3.1. Literature Search

Figure 1 summarizes the procedures of the literature search and study screening. A systematic literature search was conducted between 27 August and 1 September 2025, using four electronic databases: ERIC, APA PsycINFO, China National Knowledge Infrastructure (CNKI), and Education Source Ultimate. No publication-year restrictions were imposed due to the limited empirical literature on Chinese written composition. The search was restricted to peer-reviewed journal articles published in English. This restriction was informed by the emphasis on internationally indexed, predominantly English-language publications in academic evaluation systems across mainland China, Taiwan, and Hong Kong (Chou, 2014; Horta & Shen, 2020; Li & Flowerdew, 2009; Xu, 2020). Such publication incentives have made English-language journals important outlets for research from these contexts.

The search strategy combined three sets of English keywords related to (a) writing and composition, (b) correlational relationships, and (c) Chinese language and literacy. The same keyword groups were used across the four databases, with search terms entered through their respective interfaces. For the EBSCO-hosted databases, the following Boolean search expression was used: SU(writing OR composition OR "written expression" OR "writing achievement" OR "written composition") AND AB(correlat* OR "Pearson's *r*" OR "correlation matrix" OR intercorrelation) AND AB(Chinese OR Mandarin OR "Chinese language" OR "Chinese writing" OR "Chinese composition" OR "written Chinese" OR "Chinese literacy" OR "Chinese children" OR "Chinese students" OR "native Chinese" OR "Chinese L1" OR "first language Chinese").

For CNKI, the corresponding English keywords were entered through its search interface rather than using EBSCO-specific field codes. APA PsycINFO and ERIC were searched through the university library's integrated search interface, which returned a combined set of 350 records.

All retrieved records were subsequently imported into Covidence for duplicate removal and eligibility screening. Studies were screened based on their publication type, study design, participant characteristics, componential-skill measures, and writing outcomes. The first and fifth authors independently conducted title-and-abstract and full-text screening and were also responsible for coding study characteristics. Disagreements regarding study eligibility were discussed and resolved through consensus, resulting in 100% agreement on the final inclusion and exclusion decisions. Reasons for full-text exclusion were documented in Covidence and summarized in the PRISMA flow diagram (Figure 1).

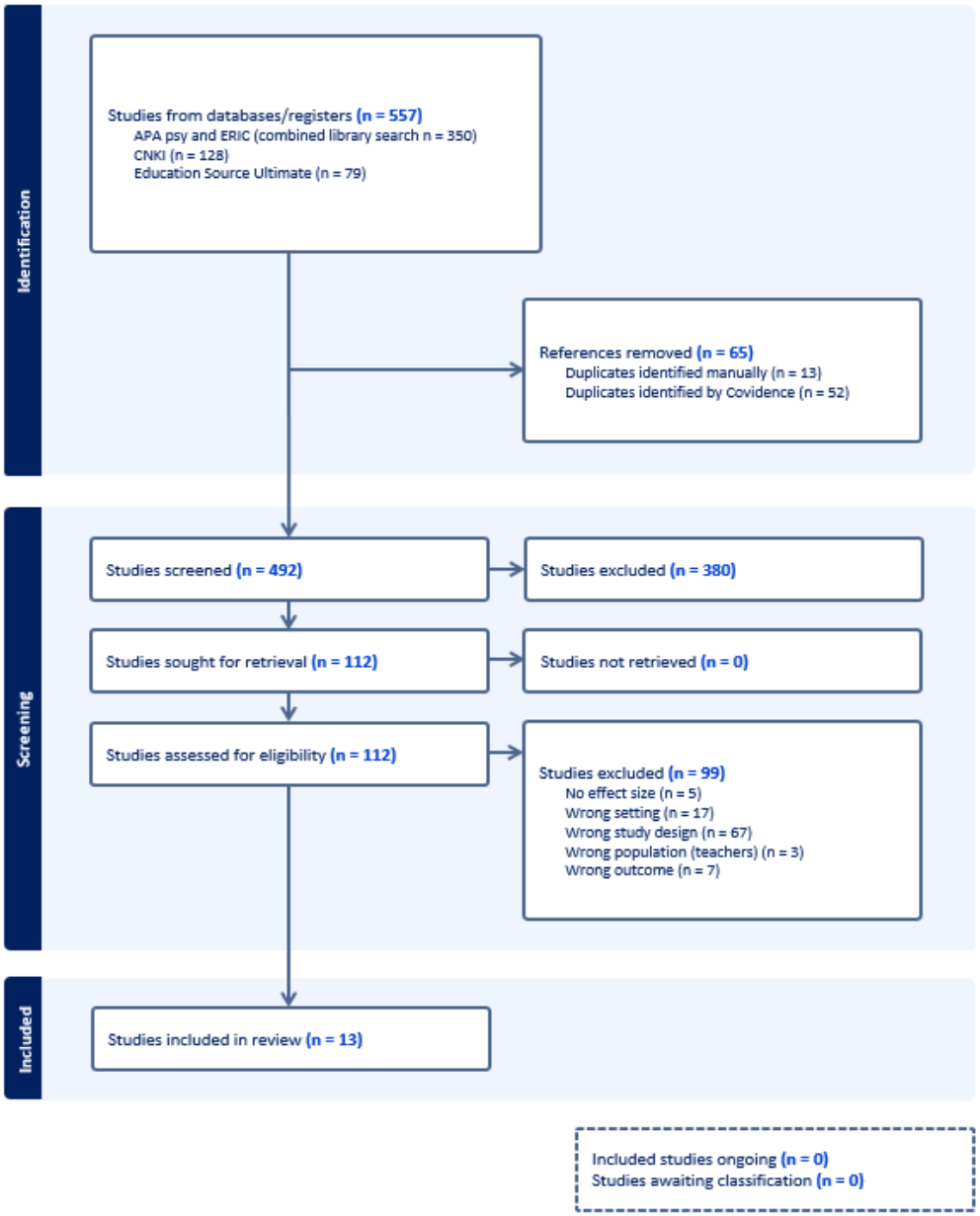


Figure 1. PRISMA figure of the search and inclusion of studies.

3.2. Inclusion and Exclusion Criteria

The studies identified through the initial search were assessed for inclusion in the meta-analysis according to the following criteria:

1. Empirical studies reporting extractable correlations between componential skills and Chinese text-level writing composition.

2. Studies published in peer-reviewed journals.
3. No publication-year limitations.
4. Studies involving monolingual or bilingual Chinese learners developing Chinese literacy in their home or regular educational contexts, rather than learners studying Chinese primarily as a foreign language.
5. Studies published in English.

The included studies involved elementary-school learners from Grades 1 to 6. Eligible writing outcomes included text-level composition measures assessing writing quality or performance, such as narrative, expository, and argumentative compositions. Studies focusing exclusively on character writing, word spelling, copying, or reading without a text-level composition outcome were excluded. When multiple eligible componential skills or composition outcomes were reported, the corresponding correlations were extracted separately. Writing outcomes were harmonized at the broader construct level by restricting inclusion to assessments of text-level Chinese composition. Although the included studies employed different writing genres and scoring procedures, all eligible outcomes involved the production and assessment of connected written text. The writing tasks, assessment procedures, and scoring criteria, where reported, are documented in Supplementary Data S1. However, differences in composition genres and scoring approaches were not examined as separate moderators due to the limited number of studies and uneven reporting across the included literature.

For participant eligibility, Chinese was not required to be the sole language of instruction but had to serve as a language of literacy development or formal literacy instruction. The same eligibility criteria were applied across geographical and educational contexts, and both monolingual and bilingual samples were included in the overall synthesis. For studies involving mixed populations, only data from eligible participants were extracted when reported separately. Language background was not examined as a moderator due to the limited number of studies and inconsistent reporting of participant language backgrounds.

Full-text exclusions were categorized according to the primary reason for ineligibility, including study design, writing outcomes, participant characteristics, data availability, and publication restrictions, as summarized in Figure 1. Studies involving clinical or special populations were excluded unless data from typically developing Chinese learners were reported separately. For example, only correlations from the typically developing comparison group in Yeung (2022) were included, whereas data from children with developmental dyslexia were excluded.

Following previous meta-analyses of literacy development published in high-impact journals, including those focused on Chinese (e.g., Ke, 2025; Kokkali & Antoniou, 2024; Suggate, 2016), only peer-reviewed journal articles were included to establish a clearly defined publication scope, focusing the synthesis on studies that had undergone formal peer review and provided accessible methodological and statistical reporting.

Adopting PRISMA guidelines (Page et al., 2021), the final sample comprised 13 publications contributing 170 effect sizes. However, the number of publications does not necessarily correspond to the number of independent participant samples, as some publications may involve overlapping cohorts that could not be fully distinguished.

3.3. Coding Procedures & Schemes

For each included study, we extracted the sample size, participant characteristics, study design, grade level, script type, componential-skill measures, writing composition outcomes, and corresponding correlation coefficients. When studies reported multiple eligible componential skills, composition outcomes, or assessment time points, the corresponding correlations were extracted separately. The fourth author extracted the

correlation coefficients, which were independently verified against the original publications by the first author. Study characteristics were independently coded by the first and fifth authors, while componential skills and their classifications were independently coded by the first and fourth authors. Initial intercoder agreement, calculated using percentage agreement, exceeded 95% for both coding processes. Discrepancies were resolved through discussion and two consensus meetings, during which the authors reviewed the original studies and agreed on the final coding decisions.

Additional audits of the extracted data were conducted to strengthen coding transparency and consistency. First, a measure-level classification audit reassessed componential skills according to their primary assessed constructs and task requirements, distinguishing production-based transcription from orthographic recognition or judgment tasks. This audit resulted in the separate identification of seven recognition-based effect sizes excluded from the two-domain analysis. Second, a descriptive subskill audit examined the distribution of effect sizes across transcription and oral language subskills to clarify the composition of the two domains. Third, a longitudinal measurement-timing audit separately documented predictor and writing-outcome assessment occasions and classified correlations as concurrent, prospective, or reverse-temporal. Fourth, the authors reviewed sample-source identifiers to document potentially dependent effect sizes across publications and verified whether extracted correlations were zero-order or partial correlations. These additional audits were conducted by the first and fifth authors, with coding decisions documented in Supplementary Data S1.

A brief descriptive methodological appraisal was also conducted by the first and fifth authors based on the coded study characteristics to examine the reporting of participant selection, measurement reliability, composition scoring procedures, and missing-data handling across the included studies. The findings are summarized in Supplementary Data S3. This appraisal was intended to identify methodological reporting gaps rather than provide a formal risk-of-bias assessment.

3.3.1. Componential Skill Domain Coding

Componential skills were classified into two broad domains, transcription and oral language skills, based on the primary construct assessed rather than solely on the administration or response modality. Transcription skills included character-writing and orthographic production skills, such as copying, word spelling and dictation, stroke-sequence knowledge, and handwriting fluency. The oral language domain was operationalized as a broad category encompassing linguistic, metalinguistic, and language-generation skills associated with written composition, rather than as a homogeneous construct of ideation.

To characterize the composition of this domain, we additionally conducted a descriptive subskill breakdown. Among the 141 oral language effects, syntactic skills (m = 65, 46.1%) and morphological awareness (m = 55, 39.0%) accounted for the majority of the observations. The remaining effects represented phonological awareness, vocabulary, oral narrative and ideation, orthographic–phonological naming, discourse organization, and definitional skills. These descriptive distributions are reported in Supplementary Data S1. Accordingly, the pooled oral-language estimate should be interpreted primarily in relation to the linguistic and metalinguistic skills represented in the available studies, rather than as an estimate of a single unified language or ideation construct.

Particular attention was given to assessments involving both linguistic and orthographic processes. For example, morphological-awareness tasks primarily assessed the ability to recognize, discriminate, or generate meaningful morphemes and compound words rather than directly measuring character-writing accuracy or handwriting fluency (Guan et al., 2014, 2019; Kuo et al., 2026; Sun et al., 2018; Yeung et al., 2013b). Therefore,

morphological awareness was classified as an oral language skill. On the other hand, orthographic measures were classified according to their primary assessed construct and actual task requirements. Measures requiring written production of characters, character components, or strokes were classified as transcription skills. In contrast, recognition- or judgment-based orthographic measures that did not require written production were distinguished from transcription. Following a re-examination of the original assessment procedures, three measures from Yeung et al. (2013b, 2016), comprising seven effect sizes, were classified separately as orthographic recognition/judgment. These measures involved identifying semantic or phonological relationships through recognition-based responses rather than producing written characters. They were not reassigned to oral language skills because their primary assessment demands involved orthographic processing. All seven effects were retained in the complete extracted dataset for transparency but excluded from the two-domain analysis. Thus, the analytical dataset comprised 170 effect sizes, including 29 transcription and 141 oral language effects.

3.3.2. Grade Level and Measurement Time Coding

Grade level was coded according to the grade at which writing performance was assessed, with three procedures applied based on how correlations were reported. First, when correlations were pooled across multiple grades, each correlation was retained as a single effect size and assigned the midpoint of the reported grade range. For example, pooled correlations across Grades 4–6 in Guan et al. (2013) were coded as Grade 5 without duplication. Second, when correlations were reported separately by grade, each correlation was assigned its corresponding grade. For example, correlations reported for Grades 4, 5, and 6 in Guan et al. (2014) were coded separately. Third, for longitudinal studies, correlations were coded according to the grade at which writing performance was assessed rather than the grade at which componential skills were measured. For example, correlations involving writing outcomes assessed in Grades 1, 2, and 4 in Yeung et al. (2013b) were coded as Grades 1, 2, and 4, respectively.

To further account for differences in measurement timing, we conducted an additional audit of the correlations extracted from longitudinal studies. Predictor and writing-outcome assessment occasions were coded separately, and correlations were classified as concurrent, prospective, or reverse-temporal based on their measurement sequence. Among the 109 effect sizes extracted from six longitudinal studies, 77 represented concurrent associations, 31 represented prospective associations, and one represented a reverse-temporal association. Following the exclusion of seven orthographic recognition/judgment effects, 102 correlations from longitudinal publications were retained in the primary analysis, including 74 concurrent, 27 prospective, and one reverse-temporal association.

These findings indicate that longitudinal study designs did not necessarily yield exclusively prospective correlations. Accordingly, the grade moderator represents the grade of writing assessment and should not be interpreted as directly capturing developmental changes in the relationships between componential skills and writing performance. The original grade information, measurement timing, and coding decisions are documented in Supplementary Data S1.

3.3.3. Script Type Coding

Script type was coded according to the writing system used in each study, distinguishing between Simplified and Traditional Chinese contexts. Because script type was closely associated with geographical and educational settings, it was treated as a contextual moderator rather than a direct indicator of character complexity. Differences in curricula, participant characteristics, and assessment procedures could not be

independently controlled within the available evidence. Therefore, associations between componential skills and writing performance across script contexts were interpreted cautiously.

3.4. Data Analysis

The 13 included publications in the meta-analysis yielded 170 effect sizes: 156 zero-order correlations and 14 partial correlations adjusted for age and nonverbal ability. The partial correlations were retained to include eligible evidence from publications that did not report corresponding zero-order estimates. The pooled estimate therefore summarizes a mixture of adjusted and unadjusted associations. Study characteristics and sample sizes were summarized in Table 1. Correlations were transformed to Fisher's z, and sampling variances were approximated by $1/(N - 3)$ for zero-order correlations and $1/(N - q - 3)$ for partial correlations, where q was the number of covariates ($q = 2$).

Table 1. Characteristics of the included Chinese writing studies.

Publication	Context; Script	Design	Writing Grade (Code)	*N* Used	Domains	*m*
Guan et al. (2013)	Mainland China; S	C	4–6 pooled (5)	156	O	12
Guan et al. (2014)	Mainland China; S	C	4, 5, 6	246, 242, 261	O	36
Sun et al. (2018)	China; Singapore; S	C	3	190; 390	O	6
Tong and McBride (2016)	Hong Kong; T	L	5, 6	129	O	8
Yan et al. (2012)	Hong Kong; T	L	3	153	T, O	11
Yeung (2022)	Hong Kong; T	C	3–5 pooled (4)	41	T, O	2
Yeung et al. (2013a)	Hong Kong; T	C	4	259	T, O	4
Yeung et al. (2013b)	Hong Kong; T	L	1, 2, 4	340	T, O	15
Yeung et al. (2016)	Hong Kong; T	L	2, 4	239	T, O	8
Yeung et al. (2017)	Hong Kong; T	C	1, 3, 5 pooled (3)	249	T, O	6
Yeung et al. (2020b)	Hong Kong; T	C	3–5 pooled (4)	388	T, O	2
Guan et al. (2019)	Mainland China; S	L	3, 4, 5, 6	246, 233, 227, 220	O	48
Yeung et al. (2020a)	Hong Kong; T	L	3, 4; 5, 6	67, 63; 62, 59	T, O	12

Note. C = cross-sectional; L = longitudinal; S = simplified script; T = traditional script in the script column and transcription in the domains column; O = oral language; *N* = sample size; *m* = number of effects. Pooled-grade codes are shown in parentheses.

To account for dependence among correlations from the same samples, analyses used correlated-effects robust variance estimation (RVE), implemented in R version 4.5.2 (R Core Team, 2025) with the robumeta package (version 2.1; Fisher et al., 2023). Publications sharing a data source were assigned to the same cluster, yielding 10 sample-source clusters (see Supplementary Data S2—Table S1). Three pairs reported shared data sources: Guan et al. (2013, 2014), Yeung et al. (2013a, 2016), and Yeung et al. (2017, 2020a). The remaining seven publications were assigned separate clusters because further overlap could not be verified. Models were fitted using the robu function, with an assumed within-cluster correlation of $\varrho = 0.80$ and small-sample corrections. Coefficient-specific effective degrees of freedom (df) were reported, and tests with df below 4 were interpreted cautiously.

An overall model estimated the pooled association between componential skills and Chinese writing (RQ1). Separate moderator models examined skill domain (RQ2), grade (RQ3), and script context (RQ4). A model containing all three main effects then estimated each moderator's coefficient after adjustment for the other two. Exploratory models examined whether grade-related patterns differed by domain or script context. The first included Domain × Grade with script adjustment; the second included Grade × Script with domain adjustment. Both retained all three main effects. Each interaction was fitted

separately because the 10 clusters provided limited information for estimating multiple interaction terms. Grade was modeled continuously and centered at Grade 4; transcription and simplified script were the reference categories. Domain × Script could not be estimated because no simplified-script transcription effects were available.

Association estimates and their 95% confidence intervals were back-transformed to Pearson's *r*. Moderator coefficients and statistical tests were reported on the Fisher-*z* scale. To aid interpretation, additional *r*-scale comparisons were calculated from model predictions, with approximate confidence intervals obtained using the delta method and the full CR2 covariance matrix in clubSandwich (version 0.6.1; Pustejovsky, 2025).

Sensitivity analyses examined whether the overall association and the domain-only, grade-only, script-only, and all-main-effects estimates depended on the handling of sample overlap, pooled-grade correlations, partial correlations, or within-cluster dependence. Both exploratory interaction models underwent the same checks. Sensitivity to overlapping samples was assessed by retaining only the publication with the larger sample from each shared-source pair (see Supplementary Data S2—Table S1). Pooled-grade correlations and partial correlations were excluded in separate analyses, and the assumed within-cluster correlation was varied across 0, 0.20, 0.40, 0.60, and 0.90. Each analysis changed one decision at a time. Leave-one-cluster-out analyses were conducted for the same models, omitting both publications together when a source was shared (see Supplementary Data S2—Table S2).

Heterogeneity was assessed using τ^2 and I^2 to summarize variation in correlations beyond sampling error. An approximate 95% prediction interval was calculated on the Fisher-*z* scale using an adaptation of the random-effects formula described by Deeks et al. (2026, Section 10.10.4.3). The pooled estimate was expanded in both directions by a margin equal to the *t* critical value multiplied by the square root of the sum of τ^2 and the squared robust standard error. To align this approximation with the fitted RVE model, the calculation used the robust standard error and coefficient-specific Satterthwaite degrees of freedom in place of the conventional standard error and $k - 1$ degrees of freedom. Both limits were back-transformed to Pearson's *r*.

Potential publication bias was examined using a funnel plot (Figure 2), an Egger-type regression test, a Kendall rank test, and trim-and-fill analysis. To account for dependence among effects, one generalized least-squares estimate was calculated per sample-source cluster, assuming a within-cluster correlation of 0.80. These estimates were analyzed in a separate random-effects model fitted by restricted maximum likelihood in metafor 5.0-1 (Viechtbauer, 2010). Neither the Egger-type test ($z = -0.635$, $p = 0.526$) nor the Kendall rank test ($\tau = -0.289$, $p = 0.291$) was statistically significant. Trim-and-fill imputed no additional aggregates and left the diagnostic pooled estimate unchanged. These findings were considered exploratory and did not rule out publication bias because only 10 sample-source clusters were available.

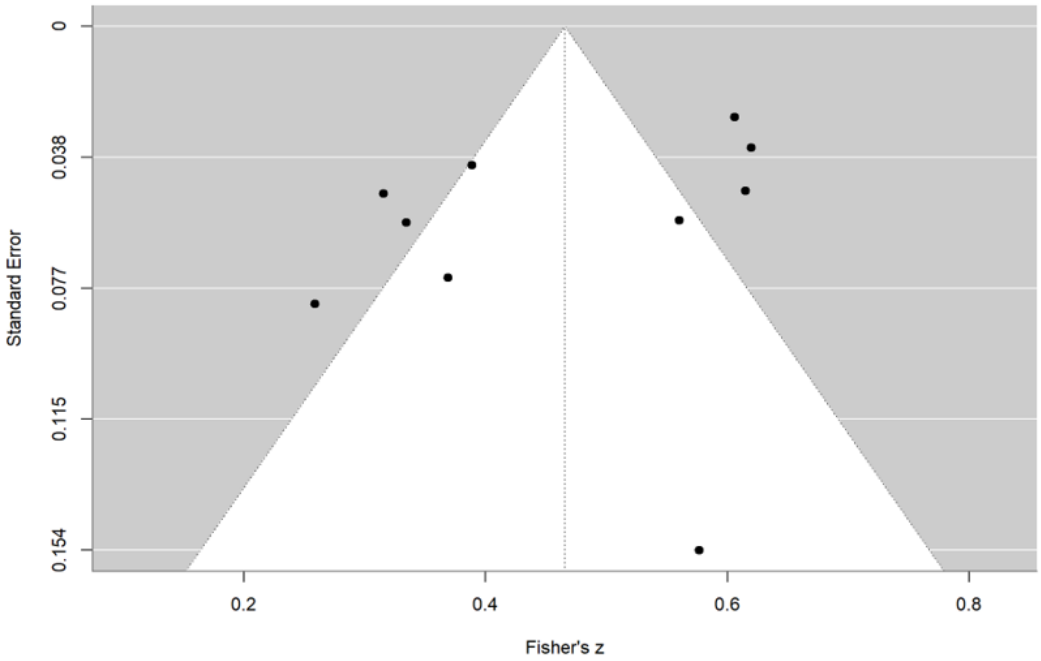


Figure 2. Funnel plot using sample-source-level aggregated effects. Each dot represents one sample-source cluster. The vertical dashed line indicates the pooled estimate from the random-effects model used for publication-bias diagnostics. The white funnel represents the approximate 95% pseudo-confidence region, while the grey shaded area represents values outside this region.

4. Results

The primary analysis included 170 effects from 13 publications grouped into 10 sample-source clusters. Study characteristics are summarized in Table 1, and model results are reported in Table 2. The overall association between the coded componential skills and Chinese writing was positive, $r = 0.375$, 95% CI [0.293, 0.451], $p < 0.001$, effective $df = 8.82$. Heterogeneity was substantial ($\tau^2 = 0.03051$ on the Fisher-z scale; $I^2 = 85.63\%$). The approximate 95% prediction interval was [−0.013, 0.664], indicating considerable variation among underlying associations despite the positive mean.

Separate moderator models examined whether the associations varied by skill domain, grade, or script context. In the domain-only model, estimated associations were $r = 0.457$, 95% CI [0.300, 0.589], for transcription skills and $r = 0.340$, 95% CI [0.275, 0.401], for oral language skills. The difference between domains was not statistically significant: $b = -0.139$, $SE = 0.059$, 95% CI [−0.280, 0.002], $df = 6.70$, $p = 0.052$. In the grade-only model, grade did not significantly moderate the association: $b = 0.005$, $SE = 0.021$, 95% CI [−0.049, 0.059], $df = 4.97$, $p = 0.829$. The estimated association at Grade 4, the centering point, was $r = 0.375$, 95% CI [0.290, 0.455]. In the script-only model, estimated associations were $r = 0.322$, 95% CI [0.203, 0.432], for simplified-script contexts and $r = 0.399$, 95% CI [0.275, 0.509], for traditional-script contexts. The difference between script contexts was not statistically significant: $b = 0.088$, $SE = 0.064$, 95% CI [−0.090, 0.266], $df = 4.01$, $p = 0.242$. The simplified-script estimate had a low effective df of 2.00.

When domain, grade, and script were entered together, none of the adjusted moderator coefficients was statistically significant. The estimates were $b = -0.134$, $SE = 0.061$, 95% CI [−0.285, 0.018], $df = 5.57$, $p = 0.073$ for domain, $b = 0.019$, $SE = 0.021$, 95% CI [−0.034, 0.072], $df = 5.20$, $p = 0.410$ for grade, and $b = 0.040$, $SE = 0.067$, 95% CI [−0.132, 0.213], $df = 5.03$, $p = 0.577$ for script. Adding the exploratory Domain × Grade interaction, while adjusting for script, yielded a nonsignificant interaction: $b = -0.066$, $SE = 0.044$, 95% CI [−0.196, 0.064], $df = 3.50$, $p = 0.218$. The separate Grade × Script interaction adjusted for

domain was also nonsignificant: $b = 0.017$, $SE = 0.046$, 95% CI [−0.120, 0.154], $df = 3.41$, $p = 0.731$. The low effective degrees of freedom for both interactions indicate limited independent information for comparing grade slopes.

Table 2. Main meta-analytic models and contextual estimates on the Pearson *r* scale.

Model	Model Component	Fisher-*z* Results				Estimates on Pearson's *r* Scale		
		z or *b*	95% CI (*z*)	*df* (*z*)	*p* (*z*)	*r*-Scale Estimate	95% CI (*r* Scale)	*df* (*r* CI)
Overall	Overall	0.394	[0.302, 0.486]	8.82	<0.001	0.375	[0.293, 0.451]	8.82
Domain only	Transcription	0.493	[0.310, 0.676]	4.75	0.001	0.457	[0.300, 0.589]	4.75
Domain only	Oral language	0.354	[0.283, 0.424]	7.65	<0.001	0.340	[0.275, 0.401]	7.65
Domain only	Oral vs. transcription	−0.139	[−0.280, 0.002]	6.70	0.052	−0.117	[−0.227, −0.007]	7.04
Grade only	Grade 4	0.395	[0.298, 0.491]	8.22	<0.001	0.375	[0.290, 0.455]	8.22
Grade only	Grade	0.005	[−0.049, 0.059]	4.97	0.829	0.004	[−0.042, 0.050]	4.97
Script only	Simplified	0.334	[0.206, 0.463]	2.00	0.008	0.322	[0.203, 0.432]	2.00
Script only	Traditional	0.422	[0.282, 0.562]	5.85	<0.001	0.399	[0.275, 0.509]	5.85
Script only	Traditional vs. simplified	0.088	[−0.090, 0.266]	4.01	0.242	0.076	[−0.079, 0.232]	3.77
All main effects	Transcription–Traditional	0.505	[0.333, 0.678]	5.10	<0.001	0.466	[0.321, 0.590]	5.10
All main effects	Oral–Simplified	0.331	[0.242, 0.421]	2.01	0.004	0.320	[0.237, 0.398]	2.01
All main effects	Oral–Traditional	0.371	[0.215, 0.528]	5.22	0.002	0.355	[0.212, 0.484]	5.22
All main effects	Oral vs. transcription	−0.134	[−0.285, 0.018]	5.57	0.073	−0.111	[−0.235, 0.013]	5.65
All main effects	Grade	0.019	[−0.034, 0.072]	5.20	0.410	0.014	[−0.026, 0.055]	5.11
All main effects	Traditional vs. simplified	0.040	[−0.132, 0.213]	5.03	0.577	0.036	[−0.117, 0.188]	4.91
Domain × Grade + script	Transcription–Traditional	0.536	[0.338, 0.733]	3.93	0.002	0.490	[0.326, 0.625]	3.93
Domain × Grade + script	Oral–Simplified	0.334	[0.201, 0.466]	2.01	0.008	0.322	[0.198, 0.435]	2.01
Domain × Grade + script	Oral–Traditional	0.370	[0.221, 0.519]	5.15	0.001	0.354	[0.217, 0.477]	5.15
Domain × Grade + script	Oral vs. transcription	−0.165	[−0.361, 0.030]	5.19	0.082	−0.136	[−0.292, 0.021]	5.37
Domain × Grade + script	Grade	0.070	[−0.058, 0.199]	2.64	0.168	0.052	[−0.033, 0.136]	2.51
Domain × Grade + script	Traditional vs. simplified	0.036	[−0.141, 0.214]	4.86	0.617	0.032	[−0.125, 0.189]	4.75
Domain × Grade + script	Domain × Grade	−0.066	[−0.196, 0.064]	3.50	0.218	−0.048	[−0.136, 0.040]	3.74
Grade × Script + domain	Transcription–Traditional	0.509	[0.334, 0.684]	4.99	<0.001	0.469	[0.322, 0.594]	4.99
Grade × Script + domain	Oral–Simplified	0.333	[0.161, 0.505]	1.93	0.015	0.321	[0.160, 0.466]	1.93
Grade × Script + domain	Oral–Traditional	0.372	[0.202, 0.542]	5.09	0.002	0.356	[0.200, 0.494]	5.09
Grade × Script + domain	Oral vs. transcription	−0.137	[−0.292, 0.019]	5.33	0.074	−0.113	[−0.241, 0.015]	5.42
Grade × Script + domain	Grade	0.007	[−0.152, 0.165]	1.88	0.867	0.018	[−0.055, 0.092]	2.92
Grade × Script + domain	Traditional vs. simplified	0.039	[−0.158, 0.236]	5.00	0.634	0.034	[−0.141, 0.209]	4.87
Grade × Script + domain	Grade × Script	0.017	[−0.120, 0.154]	3.41	0.731	0.015	[−0.106, 0.135]	3.30

Note. CI = confidence interval; *df* = effective degrees of freedom. Grade was centered at Grade 4; transcription and simplified script were the reference categories. *p*-values refer to Fisher-*z* coefficients. *r*-scale estimates represent predicted correlations or model-based differences.

Sensitivity analyses retained a positive overall association after retaining one publication per shared-source pair ($r = 0.384$, 95% CI [0.301, 0.461]), excluding pooled-grade correlations ($r = 0.334$, 95% CI [0.278, 0.387]), and excluding partial correlations ($r = 0.402$, 95% CI [0.310, 0.487]); all p values were below 0.001 (Table S2, Panel A). However, the domain contrast depended on the handling of shared sources. Retaining one publication per pair yielded stronger associations for transcription than oral language skills in both the domain-only model ($b = -0.172$, 95% CI [−0.300, −0.043], $df = 6.35$, $p = 0.017$) and the all-main-effects model ($b = -0.163$, 95% CI [−0.302, −0.025], $df = 5.42$, $p = 0.028$). These contrasts were nonsignificant in the primary analyses and in the checks excluding pooled-grade or partial correlations. Grade and script coefficients remained nonsignificant in the single-moderator and all-main-effects models, and both exploratory interactions remained nonsignificant throughout. Varying ϱ from 0 to 0.90 had negligible effects on the estimates.

Leave-one-cluster-out analyses yielded overall correlations ranging from 0.347 to 0.387, with all 95% confidence intervals excluding zero. The domain-only contrast reached nominal significance after omitting the Guan et al. (2013, 2014) cluster (p = 0.047) or the Yeung et al. (2013a, 2016) cluster (p = 0.036). Grade and script coefficients and all adjusted coefficients in the all-main-effects model remained nonsignificant. Both exploratory interactions also remained nonsignificant across all omissions (Table S2, Panel B).

The descriptive methodological appraisal (see Supplementary Data S3) indicated that participant selection, measurement reliability, and composition scoring procedures were generally documented across the included studies. Among the longitudinal studies, reporting of attrition and missing-data handling was less consistent. Furthermore, differences were observed in composition scoring criteria, transcription assistance, and the treatment of spelling errors. These findings reflect variation in methodological reporting rather than formal risk-of-bias judgments.

5. Discussion

The present meta-analysis provides quantitative evidence on the associations between componential skills and Chinese writing performance. Across the included studies, skills in both transcription and oral language domains were positively associated with Chinese composition, consistent with the componential perspective of the SVWC. However, these associations do not directly establish the independent contributions of componential skills proposed by the framework. Furthermore, the substantial heterogeneity across studies suggests that these relationships may vary across samples, assessments, and educational contexts. The discussion is organized around the research questions.

5.1. RQ1: What Are the Overall Pooled Correlations Between Componential Skills and Chinese Writing Performance?

The first research question examined the overall association between componential skills and Chinese writing performance. Grounded in the SVWC, Yeung et al. (2017) proposed and empirically supported a componential framework in which Chinese writing quality is related to multiple underlying skills, including those in the transcription domain, such as stroke-sequence knowledge, word spelling, and handwriting fluency, as well as oral language skills, such as definitional skills, oral narrative skills, and syntactic skills (Law et al., 1998; Wang et al., 2014; Yan et al., 2012; Yeung et al., 2017). The present meta-analysis extends this line of research by synthesizing correlational evidence across 13 eligible publications, which together contributed 170 effect sizes on the associations between componential skills and Chinese writing outcomes.

The overall association was positive and moderate in magnitude. However, substantial heterogeneity was observed across the included effects (I^2 = 85.63%), with an approximate 95% prediction interval ranging from −0.013 to 0.664. Although the pooled association was positive, this interval indicates that the underlying association in a comparable setting could be close to zero or potentially negative. Therefore, the average association should not be interpreted as a stable relationship across all populations and assessment conditions. Nevertheless, sensitivity analyses showed that the positive overall association remained statistically significant when accounting for potentially overlapping samples, excluding pooled-grade correlations, and excluding partial correlations. Leave-one-cluster-out analyses also yielded consistently positive estimates, suggesting that the overall finding was not driven by any single sample-source cluster. Although these findings provide correlational evidence relevant to the SVWC, they do not establish the independent or causal contributions of individual componential skills.

5.2. RQ2: Do the Associations with Chinese Writing Performance Differ Significantly Between Transcription Skills and Oral Language Skills?

The second research question examined whether the associations between the componential skills and Chinese writing performance differed by skill domain. Specifically, the skill domain was tested as a moderator to compare transcription skills and oral language skills. Prior research has provided evidence for the importance of both domains, but these skills have often been examined in separate studies or as distinct predictors within individual models. For example, previous studies have highlighted the roles of transcription skills, including character writing, copying, handwriting fluency, and spelling (Law et al., 1998; Wang et al., 2014; Yan et al., 2012; Yeung et al., 2017), as well as oral language and metalinguistic skills, including morphological and syntactic awareness, vocabulary, and oral narrative ability (Guan et al., 2014; Sun et al., 2018; Tong & McBride, 2016; Yeung et al., 2017).

The results of the primary meta-analysis showed that transcription skills had a numerically stronger association with Chinese writing performance than oral language skills (r = 0.457 and r = 0.340, respectively). However, the domain contrast did not reach statistical significance (p = 0.052), and the difference remained nonsignificant after accounting for grade level and script context (p = 0.073). These results did not establish a statistically reliable difference in the strengths of association between transcription and oral language skills. Nevertheless, the confidence intervals and limited independent evidence leave open the possibility of a meaningful difference between the domains.

It should be noted that the sensitivity analyses revealed that the domain comparison was influenced by how publications sharing sample sources were handled. When only one publication was retained from each shared-source pair, transcription skills showed significantly stronger associations with writing performance than oral language skills in both the domain-only model and the all-main-effects model. In contrast, the domain differences remained nonsignificant when pooled-grade or partial correlations were excluded. Furthermore, the domain-only contrast reached nominal significance in two of the ten leave-one-cluster-out analyses, whereas the adjusted domain contrast remained nonsignificant across all omissions. These findings suggest that the relative associations of transcription and oral language skills are sensitive to decisions regarding the inclusion of publications or sample-source clusters. Importantly, retaining one publication per shared-source pair also changes the composition of the available evidence. Therefore, the significant contrasts should not be interpreted as definitive evidence that transcription skills have a stronger association with Chinese writing performance.

Taken together, these findings highlight the challenges of comparing componential skill domains when the available evidence is unevenly distributed across skills, samples, and educational contexts. Moreover, the synthesized correlations do not account for the shared variance between transcription and oral language skills. Thus, the findings should be interpreted with caution and cannot be construed as establishing equivalent or independent contributions of the two domains. Future research examining both domains within the same participant samples, using comparable measures and clearly documented sampling procedures, is needed to better understand their relative associations with Chinese writing performance.

5.3. RQ3: Does Grade Level Moderate the Associations Between Componential Skills and Chinese Writing Performance?

Grade level was examined as a moderator because the role of componential skills may shift across the developmental spectrum. Previous research on English writing has suggested that transcription skills may place heavy demands on learners because handwriting, copying, spelling, and character production are not yet fluent. When these

lower-level skills require substantial attention, fewer cognitive resources may be available for planning, sentence construction, and idea elaboration (Ahmed et al., 2022; McCutchen, 2000). This issue may be especially relevant to Chinese writing because young learners must acquire character-specific transcription skills, including stroke-sequence knowledge, character copying, handwriting fluency, and word spelling (Law et al., 1998; Wang et al., 2014; Yeung et al., 2013a, 2017). At the same time, studies of Chinese writing have also shown that oral language skills, particularly syntactic skills and broader morphosyntactic abilities, are associated with writing development across elementary grades (Guan et al., 2014, 2019; Tong & McBride, 2016; Yeung et al., 2020a). However, existing evidence has not clearly established a consistent developmental pattern in the relative associations of transcription and oral language skills with writing performance. Some studies suggest developmental continuity or invariance across grades (Guan et al., 2014), while others show that syntactic skills remain important in upper elementary writing (Yeung et al., 2017). Therefore, it remains unclear whether the associations between componential skills and Chinese composition become stronger, weaker, or remain stable across grade levels.

Findings from the present study did not provide statistically significant evidence that grade level moderated the associations between componential skills and Chinese writing performance. The grade-only model showed no significant change in the strength of the associations across grades, and this pattern remained nonsignificant after accounting for skill domain and script context. Furthermore, the exploratory Skill Domain × Grade interaction, adjusted for script context, was not statistically significant. Therefore, the findings do not establish that the relative associations of transcription and oral language skills with Chinese writing performance change systematically across grade levels.

The absence of statistically significant moderation should not be interpreted as evidence that developmental differences are absent. Neither the exploratory Domain × Grade interaction ($b = -0.066$, $p = 0.218$) nor the separate Grade × Script interaction ($b = 0.017$, $p = 0.731$) was statistically significant after adjustment for the remaining moderators. However, the low effective degrees of freedom for these interaction estimates ($df = 3.50$ and 3.41, respectively) indicate limited independent information and considerable uncertainty regarding whether grade-related associations differ by skill domain or script context. Although the interactions remained nonsignificant across the sensitivity analyses, the estimates should be interpreted cautiously given the small number of independent sample-source clusters and the uneven distribution of the available evidence.

Evaluating how the associations between transcription and oral language skills and Chinese composition vary across development is challenging in several respects. First, studies differed in their composition assessments and scoring procedures, including whether students received assistance with unfamiliar characters or whether spelling errors were penalized during composition scoring (Yeung et al., 2013a, 2016, 2017). These differences may affect the observed associations between componential skills and writing performance, although their effects could not be directly examined in the present study. Second, some studies reported correlations pooled across multiple grade levels, limiting the precision of developmental comparisons (Guan et al., 2013; Yeung, 2022). Third, the longitudinal timing audit revealed that the extracted evidence included both concurrent and prospective associations, even within studies employing longitudinal designs. Because grade level was coded according to the writing outcome, correlations assigned to the same grade could involve componential skills assessed at different time points. Consequently, the grade moderator reflects variation in associations across writing-assessment grades rather than directly capturing developmental changes in the relationships between componential skills and writing performance. Therefore, additional primary empirical studies using comparable composition assessments, grade-specific correlations, and clearly documented measurement timing are needed to determine

whether the relative associations between componential skills and Chinese writing performance change across development.

5.4. RQ4: Does Script Type, Specifically Simplified Chinese Versus Traditional Chinese, Moderate the Associations Between Componential Skills and Chinese Writing Performance?

The examination of script type as a moderator is important because Simplified and Traditional Chinese are used across different Chinese-speaking contexts (Chen, 1999; Taylor & Taylor, 2014). Although the two scripts share the same morphosyllabic writing system, they differ in character forms, stroke density, and visual–orthographic complexity. Existing research suggests that orthographic structure may influence transcription demands, particularly in character writing, copying, spelling, and handwriting fluency (Hemelstrand et al., 2024; McBride et al., 2022). Therefore, it was reasonable to examine whether the associations between componential skills and Chinese writing performance varied across Simplified and Traditional Chinese contexts.

The present meta-analysis did not provide statistically significant evidence that script context moderated the associations between componential skills and Chinese writing performance. The script-only model showed a nonsignificant difference between Simplified and Traditional Chinese contexts, and the difference remained nonsignificant after accounting for skill domain and grade level. Furthermore, the Grade × Script interaction was not statistically significant. These results suggest that the available evidence does not establish systematic differences in the strength of the associations across script contexts.

These findings should be interpreted cautiously because script type was closely associated with geographical and educational contexts. For example, Simplified and Traditional Chinese are used across different regions where curricula, participant characteristics, language backgrounds, and assessment practices may also differ. Such contextual differences may influence how componential skills are developed, assessed, and associated with writing performance. Moreover, the distribution of componential skills was uneven across script contexts. In the two-domain dataset, all 29 production-based transcription effects originated from Traditional Chinese contexts, whereas the 102 effects from Simplified Chinese contexts exclusively assessed oral language skills. Consequently, the available evidence does not permit a complete comparison of transcription and oral language associations across both script contexts. Although the additive moderator model included skill domain and script type as separate predictors, statistical adjustment cannot fully address the absence of Simplified Chinese transcription observations. Consequently, the adjusted domain contrast was informed by comparisons between transcription and oral language skills within Traditional Chinese contexts, whereas the adjusted script contrast was informed by oral language effects across the two script contexts. These estimates therefore have limited generalizability beyond the observed combinations of skill domain and script context and should not be interpreted as independently established effects of domain or script type.

Therefore, the nonsignificant script differences should not be interpreted as evidence of an absence of script-related influences on Chinese writing development. Rather, the available studies do not provide sufficient independent and comparable evidence to isolate the role of script type. Future empirical studies directly comparing Simplified and Traditional Chinese learners using comparable componential-skill and composition assessments are needed to clarify these relationships.

5.5. Methodological Implications for Future Chinese Writing Research

Beyond synthesizing the associations between componential skills and Chinese writing performance, the present meta-analysis identified several methodological

limitations in the existing literature that constrain comparisons across componential domains and developmental contexts. First, while the synthesis included 170 effect sizes from 13 publications, these publications represented only 10 sample-source clusters. Several publications were derived from shared data sources, indicating that the number of publications does not necessarily reflect the amount of independent evidence available. The importance of this issue was further illustrated by the sensitivity analyses, in which retaining one publication per shared-source pair yielded significant domain contrasts favoring transcription skills, unlike the primary models. This variation demonstrates that domain-level comparisons may depend on the composition of the available evidence and the treatment of publications sharing sample sources. Future empirical studies should therefore clearly identify overlapping participant cohorts and report the relationships among publications derived from the same datasets.

Second, the reporting of grade-level correlations and measurement timing was inconsistent across studies. Some studies provided grade-specific correlations, whereas others reported correlations pooled across multiple grades. While midpoint coding permitted these pooled estimates to be retained, it could not recover developmental differences within the original samples. The sensitivity analysis excluding pooled-grade correlations also produced a different Domain × Grade estimate. Furthermore, the longitudinal timing audit revealed that some studies contributed both concurrent and prospective correlations. Future longitudinal studies should therefore provide complete correlation matrices and clearly distinguish the grades at which componential skills and writing outcomes were assessed.

Third, the available evidence was unevenly distributed across componential skill domains and script contexts. Oral language skills were more frequently examined than transcription skills, and no production-based transcription effects were available from Simplified Chinese contexts. Consequently, the available studies did not support a complete comparison of the two skill domains across both script contexts. More studies examining both skill domains within the same samples, as well as direct comparisons across script contexts using comparable assessments, are needed.

Finally, substantial variation was observed in the operationalization and reporting of componential skills and composition outcomes. Studies differed in the specific skills assessed, writing genres, scoring procedures, and the extent to which transcription demands were reduced during writing assessments. For example, some composition tasks provided assistance with unfamiliar characters or reduced penalties for character-writing errors, potentially lowering transcription demands. Similarly, variation in scoring criteria may have emphasized linguistic accuracy, content organization, or overall writing quality to different degrees. Some studies also reported partial correlations controlling for participant characteristics, whereas others provided zero-order correlations. These differences may affect the comparability of the synthesized associations, although the limited number of independent samples and incomplete reporting prevented a reliable examination of their effects through moderator analyses. The descriptive methodological appraisal further indicated that participant selection, measurement reliability, and composition scoring procedures were generally documented, whereas the reporting of missing-data handling was less consistent. Future research should therefore clearly document assessment constructs, writing prompts, scoring criteria, transcription-support procedures, missing-data handling, and statistical adjustments.

Taken together, these methodological gaps suggest that further progress in understanding Chinese writing development requires not only additional empirical studies but also more comparable assessments, transparent reporting, and independent developmental samples.

6. Conclusions

Writing development requires the coordination of linguistic, cognitive, and transcription processes. Grounded in the componential perspective of the SVWC, the present meta-analysis synthesized evidence on the associations between transcription skills, oral language skills, and Chinese writing performance. The findings revealed a moderate positive overall association that remained statistically significant across multiple sensitivity analyses. These results provide quantitative evidence consistent with the relevance of both componential skill domains to Chinese written composition. At the same time, the substantial heterogeneity across the included studies suggests that the strength of these associations varies considerably across samples and assessments.

Although the primary moderator and interaction analyses did not identify statistically significant differences by skill domain, grade level, or script context, the domain comparison was sensitive to the treatment of shared sample sources and the omission of individual clusters. These findings should not be interpreted as establishing equivalent associations across domains or as evidence that developmental and contextual differences are absent. Rather, the limited number of sample-source clusters, uneven distribution of componential skills, and incomplete coverage across script contexts constrained the precision and interpretability of these comparisons.

Several limitations of the present synthesis should be acknowledged. First, the search was restricted to peer-reviewed, English-language publications, potentially excluding relevant evidence from Chinese-language journals and grey literature. Furthermore, supplementary searches through Web of Science and Scopus and systematic citation and reference-list searches were not conducted. The use of correlation-related search terms may also have excluded studies reporting eligible correlations without explicitly identifying them in their abstracts. These restrictions may have limited the comprehensiveness of the evidence base. Second, despite the use of sample-source clustering and sensitivity analyses to address statistical dependencies, the relatively small number of independent sample-source clusters limited the precision of complex moderator and interaction models. Third, the synthesis included both zero-order and partial correlations, which differ in their statistical adjustments and may affect the comparability of the synthesized associations. Finally, although a brief descriptive methodological appraisal was conducted, it did not constitute a formal risk-of-bias assessment. Therefore, the extent to which potential methodological biases influenced the pooled estimates could not be fully determined. Importantly, because the underlying primary studies provided correlational evidence, the present meta-analysis cannot establish causal relationships or the unique, independent contributions of individual componential skills.

To advance understanding of the relationships among transcription, oral language, and written composition in Chinese, future empirical research should prioritize independent and clearly identified participant samples, grade-specific correlations, comparable measures of componential skills and written composition, and transparent reporting of assessment and scoring procedures. Studies examining both domains within the same samples and across Simplified and Traditional Chinese contexts would permit more direct comparisons and help distinguish script-related differences from broader educational and linguistic influences. Longitudinal and experimental research would provide additional evidence on developmental processes and the relationships among componential skills. As the evidence base expands, more comprehensive models incorporating executive functioning, working memory, motivation, and self-regulation may further refine the SVWC. Overall, the present study contributes a quantitative synthesis of the associations between componential skills and Chinese writing performance while identifying important methodological priorities for future research.

https://doi.org/10.3390/educsci16122017

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci16122017/s1, Supplementary Data S1: Coding Sheet, including study characteristics, extracted effect sizes, componential-skill classifications, domain subskill breakdowns, longitudinal measurement-timing audit, cleaned data, and coding guide. Supplementary Data S2: Supplementary Tables. Table S1: Publications Grouped by Sample Source; Table S2: Sensitivity and Influence Analyses (Panel A: Sensitivity Analyses; Panel B: Leave-One-Cluster-Out Analyses). Supplementary Data S3: Descriptive Methodological Appraisal of Included Studies.

Author Contributions: Conceptualization, L.-J.K. and J.-A.L.; methodology, L.C., J.-A.L. and L.-J.K.; software, L.C., and J.-A.L.; validation, J.-A.L., L.C. and H.S.; formal analysis, L.C., J.-A.L., Y.C. and H.S.; investigation, L.C., J.-A.L., Y.C. and H.S.; resources, L.-J.K.; data curation, L.C. and J.-A.L.; writing—original draft preparation, J.-A.L. and L.C.; writing—review and editing, J.-A.L., L.C. and L.-J.K.; visualization, J.-A.L. and L.C.; supervision, L.-J.K.; project administration, J.-A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

Ahmed, Y., Kent, S., Cirino, P. T., & Keller-Margulis, M. (2022). The not-so-simple view of writing in struggling readers/writers. *Reading & Writing Quarterly*, *38*(3), 272–296. https://doi.org/10.1080/10573569.2021.1948374.

Berninger, V. W., Vaughan, K. B., Graham, S., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., & Hawkins, J. M. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology*, *94*(2), 291–304. https://doi.org/10.1037/0022-0663.94.2.291.

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. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). Guilford Press.

Chen, P. (1999). *Modern Chinese: History and sociolinguistics*. Cambridge University Press.

Chou, C. P. (2014). The SSCI syndrome in Taiwan's academia. *Education Policy Analysis Archives*, *22*, 29. https://doi.org/10.14507/epaa.v22n29.2014.

Connelly, V., Dockrell, J. E., Walter, K., & Critten, S. (2012). Predicting the quality of composition and written language bursts from oral language, spelling, and handwriting skills in children with and without specific language impairment. *Written Communication*, *29*(3), 278–302. https://doi.org/10.1177/0741088312451109.

Deeks, J. J., Higgins, J. P. T., Altman, D. G., McKenzie, J. E., Veroniki, A. A., & Simmonds, M. (2026). Analysing data and undertaking meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.5.1, Chapter 10). Cochrane. Available online: https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-10 (accessed on 10 November 2026).

Fisher, Z., Tipton, E., & Hou, Z. (2023). *robumeta: Robust variance meta-regression* (Version 2.1) [R package]. Available online: https://cran.r-project.org/web/packages/robumeta/index.html (accessed on 10 November 2026).

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education*, *43*(1), 277–303. https://doi.org/10.3102/0091732X18821125.

Guan, C. Q., Ye, F., Meng, W., & Leong, C. K. (2013). Are poor Chinese text comprehenders also poor in written composition? *Annals of Dyslexia*, *63*, 217–238. https://doi.org/10.1007/s11881-013-0081-0.

Guan, C. Q., Ye, F., Wagner, R. K., Meng, W., & Leong, C. K. (2014). Text comprehension mediates morphological awareness, syntactic processing, and working memory in predicting Chinese written composition performance. *Journal of Educational Psychology*, *106*(3), 779–798. https://doi.org/10.1037/a0035984.

Guan, C. Q., Zhao, J., Kwok, R. K. W., & Wang, Y. (2019). How does morphosyntactic skill contribute to different genres of Chinese writing from grades 3 to 6? *Journal of Research in Reading*, *42*(2), 239–267. https://doi.org/10.1111/1467-9817.12239.

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.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*(3), 369–388. https://doi.org/10.1177/0741088312451260.

Hemelstrand, S., Wong, B. W.-L., McBride, C., Maurer, U., & Inoue, T. (2024). The impact of character complexity on Chinese literacy: A generalized additive modeling approach. *Scientific Studies of Reading*, *28*(1), 1–20. https://doi.org/10.1080/10888438.2023.2217967.

Horta, H., & Shen, W. (2020). Current and future challenges of the Chinese research system. *Journal of Higher Education Policy and Management*, *42*(2), 157–177. https://doi.org/10.1080/1360080X.2019.1632162.

Ke, S. (2025). Chinese morphological awareness assessment and its relation to reading acquisition: A cross-cultural meta-analysis. *Humanities and Social Sciences Communications*, *12*, 235. https://doi.org/10.1057/s41599-025-04531-6.

Kent, S. C., & Wanzek, J. (2016). The relationship between component skills and writing quality and production across developmental levels: A meta-analysis of the last 25 years. *Review of Educational Research*, *86*(2), 570–601. https://doi.org/10.3102/0034654315619491.

Kim, Y.-S. G., 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*(5), 517–525. https://doi.org/10.1016/j.lindif.2011.06.004.

Klein, P. D., & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research*, *7*(3), 311–350. https://doi.org/10.17239/jowr-2016.07.03.01.

Kokkali, V., & Antoniou, F. (2024). A meta-analysis of almost 40 years of research: Unreleasing the power of written expression in students with learning disabilities. *Educational Research Review*, *42*, 100592. https://doi.org/10.1016/j.edurev.2024.100592.

Kuo, L.-J., Ku, Y.-M., Chen, Z., & Shih, C.-Y. (2024). Acquisition of Chinese characters: The impact of character properties and the contribution of individual differences. *Applied Psycholinguistics*, *45*(6), 1114–1146. https://doi.org/10.1017/S0142716424000420.

Kuo, L.-J., Li, Y., Sadoski, M., & Kim, T.-J. (2014). Acquisition of Chinese characters: The effects of character properties and individual differences among learners. *Contemporary Educational Psychology*, *39*(4), 287–300. https://doi.org/10.1016/j.cedpsych.2014.07.001.

Kuo, L.-J., Lin, J.-A., Anderson, R. C., Hu, X., Chen, Z., & Ku, Y.-M. (2026). Vocabulary and morphological awareness and their relationship to reading comprehension in Chinese literacy development: Conceptual and methodological issues. *Applied Linguistics Review*, *17*(4), 1725–1781. https://doi.org/10.1515/applirev-2024-0285.

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

Law, N., Ki, W. W., Chung, A. L. S., Ko, P. Y., & Lam, H. C. (1998). Children's stroke sequence errors in writing Chinese characters. *Reading and Writing*, *10*(3–5), 267–292. https://doi.org/10.1023/A:1008091730338.

Li, Y., & Flowerdew, J. (2009). International engagement versus local commitment: Hong Kong academics in the humanities and social sciences writing for publication. *Journal of English for Academic Purposes*, *8*(4), 279–293. https://doi.org/10.1016/j.jeap.2009.05.002.

Lin, J.-A., Ku, Y.-M., Chen, L., Han, C., Chen, Y., Shen, H., & Kuo, L.-J. (2025). Technology integration in Chinese literacy instruction: A mixed-method study of in-service teachers. *Educational Research & Development Journal*, *28*(1), 28–59.

McBride, C., Pan, D. J., & Mohseni, F. (2022). Reading and writing words: A cross-linguistic perspective. *Scientific Studies of Reading*, *26*(2), 125–138. https://doi.org/10.1080/10888438.2021.1920595.

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, *35*(1), 13–23. https://doi.org/10.1207/S15326985EP3501_3.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ 372*, n71. https://doi.org/10.1136/bmj.n71.

Pustejovsky, J. E. (2025). *clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections* (Version 0.6.1) [R package]. Available online: https://CRAN.R-project.org/package=clubSandwich (accessed on 10 November 2026).

R Core Team. (2025). *R: A language and environment for statistical computing* (Version 4.5.2) [Computer software]. R Foundation for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 10 November 2026).

Seoane, R. C., Wang, J., Cao, Y., & Kim, Y.-S. G. (2026). Unpacking the relation between oral language and written composition: A meta-analysis. *Review of Educational Research*, *96*(3), 761–798. https://doi.org/10.3102/00346543251320359.

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540.

Sun, B., Hu, G., & Curdt-Christiansen, X. L. (2018). Metalinguistic contribution to writing competence: A study of monolingual children in China and bilingual children in Singapore. *Reading and Writing*, *31*(7), 1499–1523. https://doi.org/10.1007/s11145-018-9846-5.

Taylor, I., & Taylor, M. M. (2014). *Writing and literacy in Chinese, Korean and Japanese* (Rev. ed.). John Benjamins. https://doi.org/10.1075/swll.14.

Tong, X., & McBride, C. (2016). Reading comprehension mediates the relationship between syntactic awareness and writing composition in children: A longitudinal study. *Journal of Psycholinguistic Research*, *45*(6), 1265–1285. https://doi.org/10.1007/s10936-015-9401-3.

Ünal Gezer, M., Dixon, L. Q., Choi, Y., Wu, S., & Kuo, L.-J. (2025). Early literacy skills and L2 writing development in emergent bilinguals: A 3-year longitudinal study. *Early Education and Development*, *36*(6), 1298–1319. https://doi.org/10.1080/10409289.2025.2521895.

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.

Wang, Y., McBride-Chang, C., & Chan, S. F. (2014). Correlates of Chinese kindergarteners' word reading and writing: The unique role of copying skills? *Reading and Writing*, *27*(7), 1281–1302. https://doi.org/10.1007/s11145-013-9486-8.

Xu, X. (2020). Performing under 'the baton of administrative power'? Chinese academics' responses to incentives for international publications. *Research Evaluation*, *29*(1), 87–99. https://doi.org/10.1093/reseval/rvz028.

Yan, C. M. W., McBride-Chang, C., Wagner, R. K., Zhang, J., Wong, A. M. Y., & Shu, H. (2012). Writing quality in Chinese children: Speed and fluency matter. *Reading and Writing*, *25*(7), 1499–1521. https://doi.org/10.1007/s11145-011-9330-y.

Yeung, P. S. (2022). Writing motivation in Chinese children with developmental dyslexia. *Journal of Research in Reading*, *45*(2), 223–252. https://doi.org/10.1111/1467-9817.12403.

Yeung, P. S., Ho, C. S.-H., Chan, D. W.-O., & Chung, K. K.-H. (2013a). Contribution of oral language skills, linguistic skills, and transcription skills to Chinese written composition among fourth-grade students. *Discourse Processes*, *50*(7), 498–529. https://doi.org/10.1080/0163853X.2013.841070.

Yeung, P. S., Ho, C. S.-H., Chan, D. W.-O., & Chung, K. K.-H. (2013b). Modeling the relationships between cognitive-linguistic skills and writing in Chinese among elementary grades students. *Reading and Writing*, *26*(7), 1195–1221. https://doi.org/10.1007/s11145-012-9411-6.

Yeung, P. S., Ho, C. S.-H., Chan, D. W.-O., & Chung, K. K.-H. (2016). Orthographic skills important to Chinese literacy development: The role of radical representation and orthographic memory of radicals. *Reading and Writing*, *29*(9), 1935–1958. https://doi.org/10.1007/s11145-016-9659-3.

Yeung, P. S., Ho, C. S.-H., Chan, D. W.-O., & Chung, K. K.-H. (2017). A simple view of writing in Chinese. *Reading Research Quarterly*, *52*(3), 333–355. https://doi.org/10.1002/rrq.173.

Yeung, P. S., Ho, C. S.-H., Chan, D. W.-O., & Chung, K. K.-H. (2020a). Longitudinal relationships between syntactic skills and Chinese written composition in grades 3 to 6. *Journal of Research in Reading*, *43*(2), 201–228. https://doi.org/10.1111/1467-9817.12298.

Yeung, P. S., Ho, C. S.-H., Chan, D. W.-O., & Chung, K. K.-H. (2020b). Writing motivation and performance in Chinese children. *Reading and Writing*, *33*(2), 427–449. https://doi.org/10.1007/s11145-019-09969-0.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.