



Article

Investigation of Individual Differences in Single-Word Spelling Skill in Adult Speakers of Mandarin: An Orthographic Processing Explanation

Tianyi Wang  and Jackie Masterson * 

Department of Psychology and Human Development, IOE, UCL's Faculty of Education and Society,
20 Bedford Way, London WC1H 0AL, UK; tianyi.w@ucl.ac.uk

* Correspondence: j.masterson@ucl.ac.uk

Abstract

The examination of individual differences in spelling ability in adults is potentially theoretically informative but has attracted little research attention. Most of this work has involved speakers of English, and some indicated that orthographic processing may be an important factor that distinguishes better from poorer spellers. The current investigations were carried out in Mandarin, which has a logographic writing system. Because Chinese characters place particularly heavy demands on visual-orthographic processing, adult Mandarin spelling provides a strong test case for orthographic accounts of individual differences in spelling ability. Participants were 54 adult speakers of Mandarin. We used adaptations of the tasks that have previously been used in studies of spelling with English speakers, including assessments of phonological processing, visual memory, character report and reading fluency. We used hierarchical multiple regression to examine the unique contribution of these literacy-related variables to spelling accuracy after controlling for chronological age and single-character identification. We also conducted quantitative error analysis of the participants' spelling errors, using error categories that distinguish phonologically appropriate errors and orthographic errors. Finally, we conducted item-level logistic regression analyses to examine the influence of number of logographemes per character on spelling accuracy at the individual-participant level. The results of the hierarchical regression analyses examining the influence of the literacy-related variables showed that character report was the strongest predictor of spelling accuracy. Quantitative error analysis of the participants' spelling errors revealed that the rate of orthographic errors increased with a decrease in spelling accuracy, while the rate of phonological errors (homophone and semi-homophone substitutions) reduced. Item-based analyses based on individual-level logistic regression analyses showed that participants with lower spelling accuracy were more sensitive to the number of logographemes per character. The results are in line with those found in previous studies in English, indicating the importance of orthographic processes in determining spelling ability in adults. We discuss the findings in the light of theories that include orthographic working memory as a late stage in the process of single-word spelling.



Academic Editors: Jon
Andoni Duñabeitia and Ana
Paula Soares

Received: 27 May 2026

Revised: 14 September 2026

Accepted: 22 September 2026

Published: 29 September 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.

Keywords: Chinese spelling; adult spellers; orthographic processing; orthographic working memory; individual differences; phonological processing

1. Introduction

While a considerable body of research has focused on the cognitive and linguistic processes involved in printed word recognition and reading, comparatively less attention has been directed towards understanding the mechanisms underlying spelling. This gap is significant, as research into spelling processes not only contributes to more comprehensive theories of written language but also informs the development of effective interventions for individuals with spelling difficulties. Poor single-word spelling can negatively impact broader written expression, which in turn may hinder academic performance, personal communication, and occupational success (e.g., [Graham, 2006](#)).

Research into single-word spelling ability has primarily employed cross-sectional and longitudinal designs across different age groups of children, aiming to map the development of spelling and its relationship with various cognitive and linguistic skills, such as phonological awareness, orthographic processing, rapid naming, verbal memory, and vocabulary knowledge (e.g., [van den Boer et al., 2015](#); [Niolaki et al., 2020](#)). Another important line of inquiry has focused on identifying the underlying causes of spelling difficulty in individuals with dysgraphia—both children and adults—often through case studies and intervention-based research (e.g., [Brunsdon et al., 2006](#); [Kohnen et al., 2008](#); [Rapp & Kane, 2002](#)). A further avenue of investigation has examined the spelling performance of literate adults who do not report reading or spelling difficulties, in an effort to understand the cognitive mechanisms that contribute to both skilled spelling and spelling weaknesses (e.g., [Burden, 1992](#); [Holmes & Ng, 1993](#); [Holmes & Quinn, 2009](#); [Rastle, 2023](#)). The present study contributes to this final strand of research by focusing on individual differences in spelling ability in adults. The participants in the study were adult readers and spellers of written Chinese in simplified forms.

Studying spelling in adults is theoretically informative for several reasons. First, individual differences in adults reflect variation in the mature spelling system, disentangled from the developmental variability that complicates interpretation of child data. Second, as we will see in the next section, existing evidence for the role of orthographic processing in Chinese spelling comes from developing readers on the one hand (delayed copying in children) and from acquired dysgraphia on the other (case studies of Mandarin- and Cantonese-speaking patients); the question of whether orthographic processing also constrains variation within the intact adult population, that is, whether it functions as a dimension of individual differences rather than only as a locus of breakdown, has not been directly tested. Third, the increasing reliance of adult Mandarin speakers on Pinyin-based digital input, and the associated phenomenon of ‘character amnesia’ (提笔忘字) ([Langsford et al., 2025](#)), provides a specific context in which the maintenance of orthographic representations in adulthood is theoretically consequential. A number of researchers (e.g., [Frost, 2012](#); [Rastle, 2023](#); [Share, 2025](#)) have noted that previous research into literacy processes has been carried out largely in Western European languages (mostly in English), and there is a need to carry out research in non-alphabetic writing systems in order to uncover language universals.

Our study, then, aimed to examine cognitive processes underlying skilled spelling performance in adult speakers of Mandarin, and also whether factors that have been identified as discriminating better and poorer spellers of English also discriminate better from poorer spellers of Mandarin. In what follows, we discuss characteristics of the Chinese writing system, research on spelling in Chinese, and a recently proposed model for single-word spelling in Chinese. We then cover research carried out in English on single-word spelling ability and outline the current study.

1.1. Literature Review

Written Chinese employs a morpho-syllabic system that is visually complex. Several researchers (e.g., [H. Li et al., 2012](#); [X. Li & Ma, 2012](#)) have argued that reading and writing Chinese places greater demands on visual-orthographic processing compared to alphabetic systems, due to the need to recognise and reproduce detailed visual patterns within each character. In Mainland China, where the current study was conducted, simplified characters are used, in contrast to the traditional characters employed in Hong Kong. Simplified characters are usually less visually complex than traditional characters and have fewer strokes (e.g., the simplified character 药 (medicine) vs. the traditional character 藥). When children begin literacy instruction in Mainland China, they are taught the Pinyin alphabetic writing system and this is used alongside the teaching of written characters, through emphasis on stroke order in copying, and reading aloud.

1.1.1. Studies of Spelling in Chinese-Speaking Populations

Over the past two decades, research on Chinese character writing has been concerned with the relative roles of phonological, orthographic, and semantic processes in spelling development. Studies have employed two main methodological approaches: error analysis in dictation tasks and experimental tasks such as delayed copying. In a study with Mandarin-speaking children in school Grades 1 to 6, [Shen and Bear \(2000\)](#) conducted qualitative analysis of the errors made in spelling-to-dictation tasks. They observed that phonological errors, such as substitution of homophonic forms (e.g., writing 议 /yi4/ for 易 /yi4/) and use of alphabetic Pinyin predominated in the early grades, while there was a gradual progression to a predominance of orthographic errors (e.g., deletion or substitution of elements) as well as semantic errors in older children. This pattern was interpreted as evidence for an initial reliance on phonological coding, perhaps facilitated by transparent Pinyin instruction in the early years of school, followed by increasing involvement of orthographic and morphological structures. However, subsequent studies using similar methods in different instructional contexts have produced contrasting findings. [Tong et al. \(2009\)](#), studying Cantonese-speaking children in Hong Kong, found that even at age six, the overwhelming majority of errors were orthographic or meaning-related. Similarly, [Yeung et al. \(2013\)](#) reported that orthographic errors were the most frequent type of error across Grades 1 to 4 in children from Hong Kong.

These cross-regional differences suggest that the error profiles observed in Chinese children's spelling may be shaped by multiple co-occurring factors that differ between Mainland and Hong Kong contexts, including instructional environment (Pinyin-based instruction on the Mainland versus direct character-focused instruction in Hong Kong), spoken language (Mandarin vs. Cantonese), and script (simplified vs. traditional characters). Because these factors typically co-vary in cross-regional research designs, isolating the unique contribution of any single factor is difficult. We nevertheless highlight the potential role of instructional environment because Pinyin instruction represents a well-documented and comparatively well-characterised difference in early literacy experience between the two populations, and because prior work has argued for its influence on early phonological processing in Chinese children (e.g., [Lin et al., 2010](#)).

While error analysis provides a window into strategy use at the point of production, it captures the final written product rather than the real-time construction process. A second line of research has involved the delayed copying task, in which participants briefly view an unfamiliar character and then attempt to reproduce it from memory. Successful reproduction is thought to depend on knowledge of spatial configurations, radical placement, and internal component structure ([McBride-Chang et al., 2011](#); [Y. Wang et al., 2014](#)). [Y. Wang et al. \(2015\)](#) and [Ye et al. \(2021\)](#) reported that early delayed copying skills pre-

dicted later word spelling in children, even after accounting for a wide range of cognitive and linguistic covariates.

The delayed copying task has also been employed in studies involving Chinese-speaking adults with acquired dysgraphia. [Law and Leung \(2000\)](#) presented a case study of a Cantonese-speaking adult with acquired dysgraphia. The researchers conducted qualitative analysis of spelling errors, basing the analysis on logographemes—the smallest spatially separable units within a Chinese character that cannot be further decomposed (e.g., the character 吐 consists of two logographemes: 口 and 土). The participant's errors in delayed copying included substitution, omission, addition, and the transposition of logographemes. [Han et al. \(2007\)](#) investigated the spelling of another Mandarin-speaking dysgraphic adult, WLZ. The predominant errors made by WLZ in delayed copying involved logographeme substitution, followed by deletion. In addition, they noted that WLZ's spelling accuracy declined as the number of logographemes in a character increased. The authors proposed that the pattern shown by WLZ indicated a deficit in the graphemic output buffer, later referred to as orthographic working memory (hereinafter OWM), a processing component that is described in Section 1.1.2.

The findings from developmental studies reviewed above concern the acquisition of spelling in children. Their relevance to individual differences in adult spelling rests on the assumption that the cognitive components engaged in spelling, including orthographic processing and OWM, remain operative in the mature spelling system, and that variation in the efficiency of these components continues to produce variation in spelling accuracy in adulthood. Under this account, the same processing components that support successful acquisition in children also underlie skilled performance in adults, and individual differences in those components should predict individual differences in spelling accuracy at both developmental stages. This is the theoretical position adopted in the present study, and it motivates our prediction that measures of orthographic processing (specifically of OWM) will predict variation in spelling accuracy in a typical adult sample.

Two features of this work directly inform the current study: [Shen and Bear's \(2000\)](#) error-analysis approach shapes our quantitative analysis of spelling errors (RQ2), and [Han et al.'s \(2007\)](#) finding that spelling accuracy declined with the number of logographemes provides the empirical basis for our item-level analysis (RQ3).

1.1.2. Theoretical Framework: Orthographic Processing and Orthographic Working Memory in Chinese Spelling

In terms of a theoretical framework for single-word spelling in Chinese, [Mo \(2023\)](#) proposed an adaptation of the English dual route spelling model outlined by [McCloskey and Rapp \(2017\)](#), to account for the unique characteristics of Chinese orthography. This adapted framework is illustrated in Figure 1. According to the model, when spelling familiar words, activation progresses via the phonological input lexicon, and then either directly to the lexical orthographic store, or else via the lexical-semantic system and then to the lexical-orthographic store.

When spelling unfamiliar words, in the original model of [McCloskey and Rapp \(2017\)](#), as in other dual route models (e.g., [Barry, 1994](#)), the use of phoneme-grapheme correspondence knowledge was proposed. In [Mo's \(2023\)](#) adaptation for Chinese, however, the sublexical mechanism is grounded in the use of radical combinations. While most Chinese words are multi-character compounds—accounting for approximately 80% of words in written text—individual characters can also function as standalone words. As sublexical units, characters encode semantic and phonetic radical information, along with positional regularities (e.g., phonetic radicals most commonly appear on the right side of characters). Spellors may attempt to construct unfamiliar characters by drawing on their knowledge of radicals and their typical configurations. However, the information

conveyed by radicals—particularly phonological information—is often unreliable. Consequently, spelling attempts based solely on radical knowledge may not consistently yield accurate character representations.

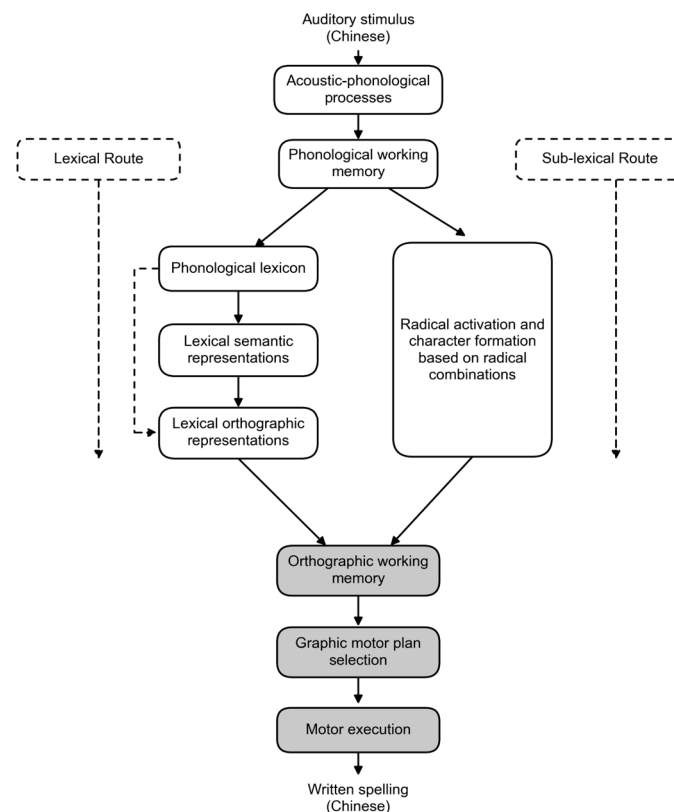


Figure 1. Model of single-word spelling in Chinese (Mo, 2023, adapted from McCloskey & Rapp, 2017).

According to the models of McCloskey and Rapp (2017) and Mo (2023), the stage of spelling production, before the selection of the graphic motor plan and motor execution, involves the engagement of OWM, which temporarily retains the output generated by either lexical or sublexical processing. Within OWM, the identity and position of orthographic elements are maintained long enough to support the execution of the graphic motor plan, enabling accurate written production. Patterns of spelling errors, involving transposition, substitution, addition and deletion of letters observed in English-speaking dysgraphic adults by, for example, Miceli et al. (1985), Caramazza et al. (1987), Hanley and Kay (1998), Kay and Hanley (1994) and Rapp and Kane (2002) were taken as support for including OWM in English models of single-word spelling. The observation that the spelling errors occurred with both words and nonwords and that accuracy decreased with an increase in letter string length were seen as additional support for an explanation in terms of an impairment of OWM.

Mo et al. (2018) argue that the visual complexity of Chinese characters means that OWM plays an important role in spelling development, serving as a platform for children to consolidate phonological, semantic and orthographic information when learning characters. Mo (2023) draws on a range of evidence in support of the role of OWM in Chinese spelling, citing the evidence, discussed above, from delayed copying in children and in adults with acquired dysgraphia. In the latter studies (Law & Leung, 2000; Han et al., 2007) error types were comparable to those reported in the studies with English-speaking dysgraphic adults, and accuracy decreased with an increase in word length (in terms of number of logographemes in the case of the studies with Mandarin-speaking adults with dysgraphia).

The Mo (2023) framework provides the theoretical basis for our central prediction: if OWM maintains orthographic representations during the transition from lexical or sublexical retrieval to graphic motor output, and if character report indexes this component, then character report should predict individual differences in spelling accuracy (RQ1).

1.1.3. Individual Differences in Adult Spelling: Evidence from English-Speaking Populations

Early research into spelling difficulties in English speakers attributed poor spelling primarily to phonological processing deficits, a view consistent with findings for individuals with dyslexia (Bruck & Waters, 1988; Campbell & Butterworth, 1985). Burden (1992) and Burt and Shrubsole (2000), for instance, reported that poor adult spellers demonstrated impaired nonword reading and that nonword reading accuracy was the strongest predictor of spelling after controlling for nonverbal intelligence. However, a growing body of evidence has indicated that a phonological deficit may not be the only underlying cause of spelling difficulty. Studies by Holmes and Quinn (2009) and Masterson et al. (2007) demonstrated that when poor spellers were matched with good spellers on reading ability, phonological weaknesses were neither consistently present nor sufficient to explain spelling difficulties. These findings shifted the focus towards orthographic processing as an influencing factor, a line of inquiry with its roots in the work of Frith (1980, 1985), who first observed that unexpectedly poor spellers showed difficulty detecting silent letters in a cancellation task, and proposed that they developed incomplete orthographic representations due to the use of ‘partial cues’ when reading.

Subsequent research has included two experimental approaches. First, behavioural studies using lexical decision tasks have shown that unexpectedly poor spellers exhibit slower and less accurate responses with certain kinds of stimuli. Holmes and Ng (1993) found that poor spellers were impaired in responding to long, idiosyncratic words and nonwords with internally misordered letters (e.g., *plaftrm*), suggesting premature termination of analysis during reading. Holmes and Castles (2001) extended this finding, reporting that unexpectedly poor spellers were slower in classifying long nonwords and nonwords. Holmes et al. (2008) reported slower and less accurate performance in lexical decision in unexpectedly poor spellers, and more errors with targets involving letter transpositions (e.g., *turlte*, *pilrgim*). Across these studies, the nature of poor spellers’ written errors in the spelling task, in terms of phonological plausibility, was qualitatively similar to that of good spellers—suggesting intact phoneme–grapheme knowledge—but speed and accuracy deficits in lexical decision pointed to poor orthographic processing.

A second line of evidence involves studies that have employed a letter report task to assess orthographic coding. Masterson et al. (2007) compared good and poor adult spellers matched on reading ability and found no group differences on phonological tasks, including spoonerisms and timed picture naming. However, the poorer spellers performed significantly worse than the better spellers in a letter report task involving briefly presented arrays. In addition, qualitative analysis of spelling errors revealed a higher proportion of phonologically implausible errors in the poorer spellers, and these involved letter additions, deletions, and substitutions (e.g., *signature* > SIGNIATURE, *financial* > FINANCAL). The authors suggested that the pattern indicated a weakness of orthographic processing. Evidence from a single-case study by Roncoli and Masterson (2016) of a 10-year-old boy, Alan, with a selective spelling impairment further substantiated this account. Alan demonstrated intact phonological processing but impaired letter report performance, alongside a pronounced word length effect on spelling accuracy. Spelling errors involved letter transposition, substitution, deletion, and addition—error types, as noted earlier, that have been documented in dysgraphia involving a deficit of OWM (Barisic et al., 2017; Caramazza et al., 1987; Kezilas et al., 2014; Kohnen et al., 2006).

Roncoli and Masterson (2016) suggested that the reason Alan had weak letter report performance was because OWM serves both reading and spelling, on the basis of evidence from people with acquired dysgraphia where the characteristics of the spelling impairment (errors involving transposition, substitution, addition and omission) were mirrored in reading (e.g., Caramazza & Miceli, 1990; Tainturier & Rapp, 2003).

The evidence suggests then that the effectiveness of processes involved in the encoding, storage, and retrieval of orthographic representations in orthographic working memory may distinguish better from poorer spellers in English. Whether the same account extends to individual differences in adult spelling in Mandarin, however, remains untested. Prior work on Chinese spelling has focused on developing readers (e.g., Shen & Bear, 2000; Yeung et al., 2013) or on adults with acquired dysgraphia (e.g., Han et al., 2007; Law & Leung, 2000), with no study to our knowledge examining individual differences within the typical adult population. This gap is theoretically consequential: without evidence from unimpaired adults, it remains unclear whether OWM operates as a dimension of individual variation within the mature spelling system or only as a locus of breakdown in impaired populations. It is also educationally consequential. In a context where digital, Pinyin-based input is increasingly replacing handwriting among adult Mandarin speakers, giving rise to widespread reports of the ‘character amnesia’ phenomenon (Langsford et al., 2025), identifying the processing components that underlie skilled adult spelling is a prerequisite for designing evidence-based literacy support for adults, a population whose needs are addressed neither by literacy interventions designed for children nor by clinical protocols designed for acquired dysgraphia.

Two features of this work directly inform the current study. The letter report paradigm used by Masterson et al. (2007) is the direct methodological ancestor of our character report task, and the individual-differences framing pioneered by Holmes and colleagues informs our analytic approach of treating spelling accuracy as a continuous dimension in an intact adult population. Taken together with the Chinese research on spelling reviewed in Section 1.1.1 and the theoretical framework in Section 1.1.2, these approaches converge on the prediction that variation in OWM accounts for individual differences in adult Mandarin spelling.

1.2. The Present Study

We aimed to investigate processes that may underly single-word spelling in Mandarin-speaking adults, and which may account for individual differences in spelling performance. We examined the effect on spelling performance of phonological ability and orthographic processing, in the form of a character report task, since these processes were identified as predictive of differences in spelling ability in the studies in English reviewed in the Introduction. We used a Spoonerisms task and a verbal short-term memory task to assess phonological ability. To assess orthographic processing, we used an array character report task, on the assumption that this task draws on the ability to actively maintain representations in OWM. In addition to character report, we employed a single-character identification task, in order to be sure that performance in character report was not constrained by single-character recognition ability. Visual short-term memory was also assessed for the same reason. We conducted quantitative error analysis of the participants’ spelling errors to examine whether different error patterns might be observed in better and poorer spellers. Finally, we examined the effect of the number of logographemes per character on spelling performance, in an item-based analysis.

RQ1. Among a set of theoretically motivated literacy-related predictors (phonological processing, visual short-term memory, character report, reading fluency, and single-

character identification), which would best predict individual differences in spelling accuracy in adult Mandarin speakers, after controlling for chronological age?

RQ2. Do better and poorer adult Mandarin spellers differ in the composition of the spelling errors they produce, specifically, in the relative proportions of orthographic errors and phonological errors?

RQ3. At the item level, does the number of logographemes per character predict spelling accuracy, and does this effect vary as a function of participants' overall spelling ability?

As for the research hypotheses, first, given the nature of Chinese script, successful spelling should be highly dependent on efficient orthographic processing (H. Li et al., 2012; X. Li & Ma, 2012; Mo et al., 2018). Sublexical processing is constrained due to the low predictability of phonetic radicals. While a range of tasks may potentially tap OWM (including visual STM and character recognition assessments), the character report task is potentially a sensitive OWM measure since it involves retention of orthographic units and distribution of attention across a multi-element array. Accordingly, we predicted the following:

Hypothesis 1. *Performance in the character report task could be a predictor of spelling accuracy in the participants in our study.*

We also sought to examine the nature of the spelling errors made by the participants in relation to their spelling ability. Within the theoretical framework outlined in the Introduction, errors arising from a weakness of OWM are characterised by phonologically implausible substitutions, additions, omissions, and transpositions. We therefore predicted the following:

Hypothesis 2. *Poorer spellers would be more likely to make orthographic errors involving character component substitutions, omissions, additions, and transpositions than better spellers.*

Finally, we conducted item-based analysis of the effect on spelling accuracy of the number of logographemes comprising each character in the spelling-to-dictation task. In the Introduction, it was noted that in the studies of dysgraphia involving a weakness of OWM, a pronounced effect of word length/number of logographemes was reported (e.g., Caramazza et al., 1987; Han et al., 2007). We expected the following:

Hypothesis 3. *The number of logographemes could have a greater influence on the spelling accuracy of poorer spellers than better spellers among our participants.*

2. Methods

2.1. Participants

Participants were recruited via a convenience-sampling procedure through advertisements posted on social media platforms between May 2022 and February 2023. Interested individuals contacted the first author and were screened for eligibility. Inclusion criteria were as follows: (i) age 18 years or older; (ii) native speaker of Mandarin; (iii) resident of Mainland China with all formal education completed in Mainland China; (iv) normal or corrected-to-normal vision; (v) no self-reported history of reading, spelling, or learning difficulties, and no history of neurological or psychiatric conditions.

Participants were 54 adults (24 female), aged between 19 and 51 years ($M = 26.92$, $SD = 6.43$; $Median = 25.10$). The age distribution was right-skewed, with the majority of participants aged in their twenties (19–29 years: $n = 45$, 83.3%) and smaller numbers in older age bands (30–39 years: $n = 5$, 9.3%; 40–49 years: $n = 3$, 5.6%; 50 years and above: $n = 1$, 1.9%). All were residents of Mainland China, right-handed, with middle to high socio-

economic status and were native speakers of Mandarin. All of the adults had normal or corrected-to-normal vision. The majority ($n = 46$) held a university degree (including 18 a bachelor's degree, 23 a master's degree and 5 a doctoral degree), and the remaining participants ($n = 8$) were graduates of junior colleges. All participants had therefore completed at least a post-secondary qualification. Using an independent-samples t -test, we compared spelling accuracy of participants with a junior-college diploma ($n = 8$, $M = 124$, $SD = 24.17$) and those with a university degree ($n = 46$, $M = 127.22$, $SD = 18.08$), showing no significant differences in spelling performance between the two groups, $t(52) = -0.44$, $p = 0.66$. Levene's test indicated equal variances ($W = 0.44$, $p = 0.51$).

All participants confirmed at recruitment that they had no history of reading, spelling, or learning difficulties, and had received all of their formal literacy education in Standard Mandarin (Putonghua) using the Pinyin phonetic system, in accordance with the national curriculum in Mainland China. Participants originated from a range of regions across Mainland China: Northeast China (e.g., Liaoning, Jilin, Heilongjiang; $n = 16$, 29.6%), South-Central China (e.g., Henan, Hubei, Hunan, Guangdong; $n = 16$, 29.6%), East China (e.g., Jiangsu, Zhejiang, Shanghai, Anhui; $n = 10$, 18.5%), North China (e.g., Beijing, Hebei, Shanxi; $n = 6$, 11.1%), Southwest China (e.g., Sichuan, Guizhou; $n = 3$, 5.6%), and Northwest China (e.g., Gansu, Xinjiang; $n = 3$, 5.6%). While participants therefore came from linguistically diverse regions and may have spoken different local dialects at home, their formal literacy instruction was uniformly conducted in Standard Mandarin. Home dialect background was not systematically recorded, and given that the province of origin does not reliably index the dialect(s) actually spoken at home, we cannot precisely characterise participants' pre-school language exposure. Parental education, language-use history (such as frequency of handwriting vs. digital input in adulthood), and frequency of Pinyin-based digital-input use were also not measured in the present study. These variables are noted as limitations of the sample characterisation (see Section Limitations and Implications).

Power analyses were conducted using G*Power 3.1 (Faul et al., 2009) to determine the sample size required for the planned analyses. Effect sizes reported in previous individual-differences studies of adult spelling in English have generally been in the large range (e.g., Holmes & Ng, 1993; Masterson et al., 2007). Adopting $f^2 = 0.35$ as the target effect size for the hierarchical multiple regression predicting spelling accuracy, with two predictors added on top of two control variables ($\alpha = 0.05$, 95% power), the required sample size was $N = 48$. Our final sample of $N = 54$ therefore exceeded this requirement.

2.2. Materials

The task that we used to assess spelling ability is described first. Then, assessments of nonverbal reasoning, vocabulary and reading fluency are outlined. Scores from these tasks were used as background variables in the statistical analyses since they have been found to be associated with spelling ability in past research. Finally, we outline the character report, single-character identification and visual memory tasks, as well as the phonological ability assessment.

Spelling. Spelling ability was assessed in a spelling-to-dictation task comprising 154 items. We note that no standardised spelling-to-dictation task exists for adult speakers of Mandarin using simplified Chinese characters, which precludes the use of an existing adult-normed instrument. We therefore drew on two published sources: the Masterson et al. (2008) list, originally developed for children but designed with progressive difficulty structure to differentiate individuals across a wide range of spelling proficiency, and the Chinese Scale and Assessment of Vocabulary (T. Wang & Tao, 1996) developed for dyslexia screening with developing readers, a standardised Chinese vocabulary assessment.

Combining these two sources allowed us to construct a dictation task that (i) covers a wide range of item difficulty, (ii) draws on validated materials rather than ad hoc item selection, and (iii) includes items appropriate for adult Mandarin speakers alongside items whose difficulty range was known to be sensitive to individual variation.

Sixty items were drawn from the spelling list developed by Masterson et al. (2008) for English speakers and subsequently translated into Mandarin by Wong (2017). In selecting Chinese translations, Wong (2017) referred to the Zhonghua Dictionary to ensure that each target item corresponded as closely as possible in meaning to its English original. The original 60-item English list was constructed to include words denoting concepts familiar to children aged 6 to 7 years and above, and to children from diverse linguistic and cultural backgrounds, and to span a range of familiarity levels so as to avoid floor and ceiling effects. The Mandarin translation of the list comprised 41 single-character words, 18 two-character words, and one three-character word. The remaining 94 items (single-character words) were drawn from T. Wang and Tao (1996). Across both sources, the final dictation task included 135 single-character words and 19 multi-character words in total, presented in random order.

Printed word Zipf frequency values for the 154 items were obtained from SUBTLEX-CH, a corpus of Chinese character frequencies compiled by Cai and Brysbaert (2010)¹. Frequencies ranged from 1.48 to 6.18 with a mean of 3.54 ($SD = 1.12$). The number of logographemes was ascertained for each of the 154 items. The mean number of logographemes was 3.56 ($SD = 1.29$) with a range of 2 to 11.

During administration of the spelling-to-dictation tasks, participants were asked to write down each word after it was read aloud by the researcher, who also provided a short sentence or phrase to contextualise the word. One point was awarded for each correctly spelled word. Reliability analyses based on the results for the participants in the current study yielded a Cronbach's alpha and a split-half reliability coefficient of 0.73, indicating acceptable internal consistency.

Each item was scored as a single unit, so that single-character and multi-character items contributed equally to the total. Two considerations motivated this decision. First, one point per correctly spelled item is the scoring convention of both source instruments (Masterson et al., 2008; T. Wang & Tao, 1996), and retaining it preserves comparability with those assessments. Second, single-character items were not scored per character in a sub-item sense: each single-character item is itself a Chinese word, and was dictated within a disambiguating phrase so that the target was unambiguously identified. This reflects a property of written Chinese that has no equivalent in alphabetic scripts, in which individual letters carry no lexical meaning. We acknowledge that multi-character items afford more opportunities for error than single-character items; they constitute, however, only 19 of the 154 items (12.3%), so any resulting distortion of the total score is small. We note also that the analysis of error types reported in Section 3.2 uses a finer unit than the accuracy score: errors there were coded at a sub-character level rather than as whole-item errors, so a single item could yield more than one error. The two units serve different purposes—item-level accuracy for the individual-differences analyses and sub-character coding for the characterisation of error types.

Reading fluency. Reading fluency was assessed with a task developed by Chan and Yeung (2020), but with the traditional characters used in the study of Chan and Yeung replaced by their simplified counterparts. The task consisted of 120 two-character words. Word frequency values for the simplified characters were obtained from SUBTLEX-CH (Cai & Brysbaert, 2010). The mean Zipf frequency was 4.47 ($SD = 0.86$). Participants were asked to read aloud as many words as possible within 45 seconds. The experimenter timed the session and recorded any reading errors. The fluency score was calcu-

lated as the number of words read correctly. If a participant completed all 120 words before the 45-second limit, their final score was adjusted: $(\text{Number of correctly read words} / \text{Time taken in seconds}) \times 60$. Based on the scores from the current study, the task demonstrated internal consistency with a Cronbach's alpha of 0.98. The split-half reliability was 0.68. Chan and Yeung (2020) reported a lower-bound reliability coefficient of 0.66 in their study.

Nonverbal reasoning. Nonverbal reasoning ability was assessed using Raven's Standard Progressive Matrices (Raven et al., 1996).

Vocabulary. Receptive vocabulary was assessed using the Chinese edition of the Peabody Picture Vocabulary Test (PPVT-C, Z. Gong & Guo, 1984), adapted from the original version developed by Dunn and Dunn (1997). In this task, the examiner presents a spoken word, and the participant is asked to select the corresponding image from four options. The task is discontinued when the participant makes more than five errors within a set of eight items. Accuracy scores were calculated, based on the ceiling item minus the total number of errors for each participant. In the current study, the task had a Cronbach's alpha of 0.74.

Visual short-term memory. We employed the visual memory task from Roncoli and Masterson (2016), that had been adapted from one originally used by Goulandris and Snowling (1991). Arabic letters (unfamiliar symbols for the participants) were the stimuli. There were 12 trials. Arrays of two, three, or four symbols were presented on the computer screen, for two seconds per symbol. The participants were then asked to select the symbols that were shown, in the correct order, from a test array of the symbols that were presented, intermixed with two distractor symbols. The test array was presented following a retention interval of one second for the first six trials and 10 seconds for the following six trials. Participants provided their responses by saying aloud the number tags under the relevant symbols in the test array. Cronbach's alpha was 0.68 on the basis of the data from the current study. The maximum score was 12.

Character report task. A letter report task, employed by Bosse et al. (2007), was adapted for Chinese-speaking participants by Chan and Yeung (2020) using 10 traditional single characters. In their study, the task demonstrated good reliability for a sample of adults (Cronbach's $\alpha = 0.83$). For the character report task in the present study, we used the same set of characters as Chan and Yeung, but in their simplified forms, as participants were Mandarin speakers. The characters were high frequency and commonly used: 雨 (/yǔ/, "rain"), 表 (/biǎo/, "watch"), 非 (/fēi/, "not"), 事 (/shì/, "thing"), 兔 (/tù/, "rabbit"), 京 (/jīng/, "Beijing"), 金 (/jīn/, "gold"), 门 (/mén/, "door"), 并 (/bìng/, "and"), and 东 (/dōng/, "east").

The task consisted of 20 trials involving six-character arrays. Each of the 10 characters appeared twice in each position across the trials, and each character was presented a total of 12 times in randomised combinations. Each trial began with a central fixation point displayed for 1000 ms, followed by a 50 ms blank screen, and then a centrally presented character array shown for 200 ms. Participants were asked to report as many characters as they could recall, regardless of their position in the array. Two practice trials were administered before the experimental trials began. The outcome measure was the total number of correctly reported characters, with a maximum possible score of 120. In the current study, Cronbach's alpha was 0.94 with a split-half reliability of 0.89.

Single-character identification. A character identification task was adapted from the letter identification task used by Bosse et al. (2007). The ten characters that were used in the character report task were presented individually at each of five presentation durations: 33, 50, 67, 84, and 101 milliseconds. A visual mask was displayed for 150 ms immediately after the character offset on each trial, to prevent afterimages. The order of characters and presentation durations was randomised, giving 50 trials in total. Participants were asked to report the

character after each presentation. Performance was scored using a weighted accuracy formula from Antzaka et al. (2018): (Accuracy at 33 ms $\times$ 5) + (Accuracy at 50 ms $\times$ 4) + (Accuracy at 67 ms $\times$ 3) + (Accuracy at 84 ms $\times$ 2) + (Accuracy at 101 ms). Thus, the maximum possible score was 150 (comprising $10 \times 5 + 10 \times 4 + 10 \times 3 + 10 \times 2 + 10 = 150$). Prior to the main set of trials, participants completed 10 practice trials—two for each presentation duration—with feedback provided. In the current study, the task demonstrated Cronbach's α of 0.62 and split-half reliability of 0.80.

Phonological ability. Scores from two assessments were used to constitute the measure of phonological ability. An assessment of phonological awareness developed by Hamilton (2007) in a study with Chinese-speaking adults was used, as well as an assessment of verbal short-term memory, the Digit Span Forwards subtest from the Revised Wechsler Adult Intelligence Scale in China (WAIS-RC, Y. X. Gong, 1983). For the onset-rime deletion task (Hamilton, 2007), participants were presented with one-, two-, or three-syllable words spoken by the tester, and were asked to say aloud the item without the initial, middle, or final sound (for example, 'please tell me what /guang1/ would be without the first sound', correct answer: /uang1/). The task comprised 33 trials for the main task and three practice items. Correct responses were awarded one point. Cronbach's alpha was 0.68 based on the scores from this study. For the verbal short-term memory task, the highest number of digits correctly repeated by the participant was the score. Cronbach's alpha was 0.72 on the basis of the data from this study.

Raven's Standard Progressive Matrices (Oxford Psychologists Press, Oxford, UK); the Peabody Picture Vocabulary Test (American Guidance Service, Circle Pines, MN, USA) and its Chinese edition (Xinhua Hospital affiliated to Shanghai Second Medical University, Shanghai, China); and the Chinese Scale and Assessment of Vocabulary (Chinese Educational Research Press, Beijing, China). The remaining measures were experimental tasks adapted from published studies, not commercial products, and are cited to their source publications.

2.3. Procedure

The study received ethical approval from the researchers' affiliated university. Prior to participation, all individuals provided written informed consent. Due to COVID-19 restrictions, data collection was conducted online between June 2022 and November 2023. All tasks were administered individually via video conferencing by the first author, a native Mandarin speaker. Each participant completed the assessments across three separate sessions. In the first session, nonverbal reasoning ability was assessed using Raven's Standard Progressive Matrices, which typically lasted 35 to 45 min. The second session, lasting approximately 30 to 40 min, included assessments of reading fluency, PA, single-character recognition and character report. The third session focused on the spelling-to-dictation task, which lasted around 30 min. This task involved the audio presentation of 154 items and their corresponding sentences, pre-recorded by the first author. Participants wrote their responses by hand and on completion of the task, they submitted a photo of their written script to the researcher. All experimental tasks involving limited presentation time were programmed in PsychoPy (version 2021.1.4; Peirce et al., 2019) and administered online via the Pavlovia platform (<https://pavlovia.org>), which provides validated stimulus timing precision comparable to laboratory-based experiment software. Participants in Mainland China used their personal laptops to complete the assessments, which included 23 Mac systems and 31 Windows systems, all operating with screen refresh rates ranging from 60 Hz to 75 Hz.

Because the experiments were administered on participants' own devices, stimulus timing was constrained by the refresh rate of each display. In the character report task,

the 200 ms array duration corresponds to a whole number of refresh cycles at both of the refresh rates present in the sample (12 cycles at 60 Hz, where one cycle is 16.7 ms, and 15 cycles at 75 Hz, where one cycle is 13.3 ms), so the intended duration was realisable exactly on every device, with a residual uncertainty of the order of a single refresh cycle, that is, at most 16.7 ms at 60 Hz and 13.3 ms at 75 Hz, or approximately 8% of the intended duration. In the single-character identification task, the five presentation durations (33, 50, 67, 84 and 101 ms) correspond to two, three, four, five and six refresh cycles on a 60 Hz display; on the 75 Hz displays, each duration could only be realised as the nearest whole number of refresh cycles, giving a maximum departure from the intended duration of approximately 7 ms. Timing variation in this magnitude is unrelated to participants' orthographic processing ability, and therefore contributes within-participant measurement noise rather than systematic bias; its expected effect is to attenuate rather than to inflate the observed associations between the experimental measures and spelling accuracy.

Before the assessments began, participants were asked to turn on their laptop cameras so the researcher could verify that they were seated approximately 50 cm from the screen. The functionality of the laptop microphones was also checked. The researcher used screen sharing to introduce each task and provide instructions. For tasks requiring visual stimuli—character identification, reading fluency and character report—the researcher ensured that participants understood the instructions and completed any necessary practice trials. Once this was confirmed, screen sharing was turned off to minimise distraction during the task. For tasks that did not require visual input, such as the phonological awareness (PA) task, the shared screen was also turned off after the initial instructions were given. Rest breaks were provided between tasks to reduce fatigue and maintain engagement.

2.4. Data Analysis

Correlation and regression analyses were used to examine the effect of the orthographic, phonological and background variables on participants' spelling ability.

Hierarchical regression was used because, unlike simultaneous regression, it allows the researcher to evaluate the incremental contribution of each variable by entering predictors in a motivated sequence and observing the change in explained variance at each step. This was particularly appropriate in the present study, given the exploratory nature of the research: as there is no established precedent for spelling ability research in adult speakers of Mandarin, there was no strong a priori basis for assuming that all predictors would contribute equally or independently. Correlation and hierarchical regression analyses were conducted in R (version 4.5.2; [R Core Team, 2025](#)) via RStudio (version 2026.01.0+392; [Posit Team, 2026](#)) using the car (version 3.1.5; [Fox & Weisberg, 2018](#)) and lm.beta (version 1.7-3; [Behrendt, 2014](#)) packages.

Variables that showed a significant bivariate correlation with spelling accuracy were entered into the hierarchical regression as predictors, allowing us to focus the model on associations that were empirically supported and to limit the ratio of predictors to observations in our modest sample ([Green, 1991](#)). The three predictors of interest, namely single-character identification, visual short-term memory, and character report, are like measures of orthographic processing, and each is theoretically related to orthographic working memory (OWM). They differ, however, in the complexity of the OWM demands they place on the participant: single-character identification requires the recognition and oral production of individual characters with minimal short-term retention load; visual short-term memory requires brief retention of visual arrays without a reading or lexical component; and character report requires the simultaneous recognition and short-term retention of multiple characters, thus combining the demands tapped separately by the other two tasks. This graded structure motivated our hierarchical entry order. Single-character

identification was entered in Step 1 (alongside chronological age as a demographic control) to remove variance associated with the simplest orthographic-recognition demands. Visual short-term memory and character report were then entered in Step 2, allowing us to evaluate whether more demanding orthographic processing and, in particular, the multi-character retention required for character report account for spelling variance over and above the character recognition variance captured in Step 1.

Quantitative analyses of the participants' errors from the spelling-to-dictation task were carried out using the categories of error from [Shen and Bear \(2000\)](#) and [Han et al. \(2007\)](#). For each participant, the proportion of orthographic errors, phonological errors, and blank responses was calculated as a proportion of total error, and we examined the relationship between these rates of error and participants' spelling accuracy using regression analyses.

For the item-based analysis examining the influence on spelling accuracy of the number of logographemes per character, individual-level binary logistic regressions were conducted for each of the participants. Accuracy for each item in the spelling task (coded as 0 = correct, 1 = incorrect) served as the dependent variable, and the log-transformation of number of logographemes was the predictor variable. The resulting B coefficient from each participant's regression was taken as an index of the individual's sensitivity to character length/complexity. These B coefficients were then analysed in relation to participants' spelling ability (accuracy) using partial correlation, controlling for chronological age. The partial correlation analyses and logistic regression analyses were conducted in Python (version 3.12.7; [Python Software Foundation, 2024](#)) via Spyder (version 5.5.1; [Raybaut, 2009](#)) using the statsmodels (version 0.14.2), pandas (version 2.2.2), NumPy (version 1.26.4), SciPy (version 1.13.1) and matplotlib (version 3.9.2) libraries.

3. Results

This section reports the results in three parts, corresponding to the three analytical strands of the study. We present first descriptive statistics for the scores from the assessments of nonverbal ability, vocabulary, reading fluency, spelling, phonological awareness, verbal short-term memory, visual short-term memory, character report and character identification (See Table 1). We then report the correlation and hierarchical regression analyses examining the association between these measures and spelling accuracy, followed by the quantitative analysis of spelling errors and the item-level analysis of the effect of number of logographemes per character.

Table 1. Descriptive statistics for all variables.

Variables	Mean (SD)	Range
Nonverbal ability (/60)	45.10 (7.19)	24–59
Vocabulary (/120)	112.48 (4.14)	100–118
Reading fluency	97.99 (16.14)	58–128.57
Phonological awareness (/33)	28.17 (4.00)	16–33
VerbalSTM (/12)	10.02 (1.19)	8–12
VisualSTM (/12)	10.52 (1.63)	5–12
Character Identification (/150)	105.00 (21.02)	37–134
Character report (/120)	81.28 (13.78)	46–110
Spelling accuracy (/154)	126.74 (18.87)	70–150

Note. VisualSTM: visual short-term memory; VerbalSTM: verbal short-term memory.

Prior to conducting the correlation and regression analyses, exploratory data analysis was performed to assess the distributional properties of each variable. Scores for single-character identification and character report met the assumption of normality; however, the distributions of the other variables were non-normal and these variables were log-

transformed. The data were analysed twice, once using the transformed scores and once using the untransformed scores, and the results of the analyses did not differ. The findings reported below are therefore based on the untransformed scores.

3.1. Association of Spelling Accuracy and Literacy-Related Variables

Prior to conducting correlation analyses, we created a composite measure of phonological ability from the scores for the phonological awareness task and the verbal STM task. There was a moderate degree of association between the two measures (Pearson's $r = 0.39$, $p = 0.004$). Z-scores for these tasks were summed for each participant. The results of Pearson's correlation for all the variables as well as chronological age are presented in Table 2.

Table 2. Correlation analysis for scores in all tasks and chronological age.

		1	2	3	4	5	6	7	8	9
1	Chron. Age	-								
2	Nonverbal Abil.	−0.04	-							
3	Vocabulary	0.18	0.42 **	-						
4	VisualSTM	−0.21	0.24	0.24	-					
5	Phon. Ability	0.09	0.02	0.28 *	−0.10	-				
6	Reading fluency	−0.13	0.01	0.21	0.16	0.42 **	-			
7	Char. Ident.	−0.49 ***	−0.05	−0.13	0.01	−0.04	−0.12	-		
8	Character report	−0.15	−0.15	0.11	0.11	0.28 *	0.35 *	0.29 *	-	
9	Spelling	−0.27 *	−0.02	0.17	0.34 *	0.06	0.13	0.18	0.44 ***	-

Note. Chron. Age: Chronological Age; Nonverbal Abil.: Nonverbal Ability; VisualSTM: visual short-term memory; Phon. Ability: phonological ability (z); Char. Ident.: character identification. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Inspection of the table reveals that spelling accuracy was significantly associated with chronological age, visual short-term memory and character report scores. Hierarchical regression analysis was used to examine the unique and incremental effects of the predictor variables on spelling accuracy. Based on the results from the correlation analyses, chronological age, visual short-term memory and character report were entered as independent variables. Single-character identification was also entered, in Step 1: although it was not significantly correlated with spelling accuracy, it indexes the least demanding of the orthographic processing measures used here, and including it as a control allows the incremental contribution of the more demanding character report measure in Step 2 to be assessed over and above basic character recognition. The results are reported in Table 3.

Table 3. Hierarchical regression analysis with spelling accuracy as the dependent variable.

Step	Predictor	B	β	t	p	R ²	ΔR^2	f ² Δ	Tol	VIF
1	Chron. Age	−0.78	−0.24	−1.55	0.128	0.08	0.08	0.09	0.76	1.31
	Char. Ident.	0.05	0.06	0.38	0.707				0.76	1.31
2	Chron. Age	−0.57	−0.18	−1.23	0.224	0.30	0.23 **	0.33	0.71	1.41
	Char. Ident.	−0.03	−0.03	−0.19	0.852				0.69	1.45
	VisualSTM	2.93	0.25	2.01	0.050				0.92	1.09
	Character report	0.55	0.40	3.11 **	0.003				0.90	1.11

Note. Chron. Age: chronological age; Char. Ident.: character identification; VisualSTM: visual short-term memory; R² = cumulative variance explained. ΔR^2 = change in R² at each step. f² Δ = Cohen's f² for the change in R² (small = 0.02, medium = 0.15, large = 0.35; Cohen, 2013). Adjusted R² for Steps 1 and 2 were 0.04 and 0.24, respectively. ** $p < 0.01$.

In Step 1, chronological age and single-character identification were entered as control variables. Age and single-character identification were not significant predictors (chronological age: $p = 0.128$; single-character identification: $p = 0.707$). The control model ac-

counted for 8% of the variance in spelling accuracy (adjusted $R^2 = 0.04$), with an effect size of $f^2 = 0.09$.

In Step 2, visual short-term memory and character report were added. Character report was a significant predictor ($B = 0.55$, $\beta = 0.40$, $t = 3.11$, $p = 0.003$), while visual short-term memory was a marginally significant predictor ($B = 2.93$, $\beta = 0.25$, $t = 2.01$, $p = 0.050$). Age and single-character identification remained non-significant. The addition of visual short-term memory and character report significantly increased the variance explained, $\Delta R^2 = 0.23$, $f^2 \Delta = 0.33$, a medium-to-large effect. The full model explained 30% of the variance in spelling accuracy (adjusted $R^2 = 0.24$), with a large overall effect size (Cohen's $f^2 = 0.43$).

All VIF values stayed well below two, confirming no multicollinearity concerns across both steps. Overall, the final model identifies character report as the strongest predictor of spelling accuracy ($\beta = 0.40$), after controlling for age and single-character identification, with visual short-term memory also making a marginally significant independent contribution.

In order to examine the potential influence of participants' educational level on the assessments in this study, we conducted a supplementary analysis in which educational level (1 = junior-college diploma, 0 = university degree) was entered as an additional covariate in Step 1 of the hierarchical regression analysis. Educational level was not a significant predictor of spelling accuracy ($B = -0.40$, $p = 0.958$), and the character report effect remained robust ($B = 0.55$, $p = 0.003$). Full results are reported in Table A1 in Appendix A.

3.2. Quantitative Analysis of Spelling Errors

Following the error categories used by Shen and Bear (2000), spelling errors were categorised into phonological and orthographic types. Phonological errors included homophone and semi-homophone substitutions. For example, the character 议 (/yì/ or /yi4/, "discuss") in the target word 议会 (/yì huì/ or /yi4 hui4/, "council") was incorrectly replaced with phonologically similar characters such as 意 (/yì/ or /yi4/, "meaning") or 衣 (/yī/ or /yi1/, "clothes"). Orthographic errors were analysed based on the classification of Han et al. (2007), and corresponded to transpositions, substitutions, additions and deletions. Transpositions were errors involving the incorrect spatial arrangement of components. For example, the character 地 (/dì/ or /di4/, "place") was written as the non-character 也土, reversing the configuration. Deletions were errors where a component or stroke was omitted, for instance, 日 (/rì/ or /ri4/, "day") was written as 口, omitting the internal stroke. Additions involve the insertion of strokes (e.g., the character 猴 (/hóu/ or /hou2/, "monkey") written with an additional short slash in the left semantic radical as 猴). Substitutions were errors where a character was replaced with another that was visually similar. These included real-word substitutions (e.g., 河 (/hé/ or /he2/, "river") spelled 苛 (/kē/ or /ke1/, "severe")), and nonword substitutions (e.g., 鼻 (/bí/ or /bi2/, "nose") written as a non-existent character 𪗇). Blank responses were those where the participant did not provide any response.

A random selection of 20% of the spelling errors was independently coded by two trained raters. Inter-rater reliability was calculated using Cohen's kappa, yielding a coefficient of 0.85 (95% CI [0.78, 0.91]), indicating good agreement according to Landis and Koch's (1977) criteria. Following this, the remaining data were coded by one of the raters. Spelling errors were coded according to strokes as sub-character unit rather than counting whole-word errors, which means that more than one error may have been made in a single word.

To look for a potential relationship between spelling ability and error types, the proportion of each type was calculated for each participant. Across the 54 participants, the mean proportion of orthographic errors was 0.40 ($SD = 0.15$), the mean proportion of phonological errors was 0.39 ($SD = 0.19$), and the mean proportion of blank responses was

0.21 ($SD = 0.18$). Three separate regression analyses were conducted to examine whether spelling accuracy was predicted by the preponderance of each error type. Participants' chronological age was entered as the first step in these analyses since it was found to be significantly associated with spelling ability in the correlation analyses presented earlier (please see Table 2). In each analysis, the error proportion was entered in Step 2 and spelling accuracy was the dependent variable. In the regression analysis with the proportion of orthographic errors, chronological age was not a significant predictor ($\beta = -0.21, p = 0.116$), while the proportion of orthographic errors was significant ($\beta = -0.29, p = 0.031$). In the regression analysis with the proportion of phonological errors, chronological age was not significant ($\beta = -0.13, p = 0.281$), while the proportion of phonological errors was ($\beta = 0.52, p < 0.001$). In the regression analysis with the proportion of blank responses, chronological age was not significant ($\beta = -0.24, p = 0.070$), while the proportion of blank responses was significant ($\beta = -0.28, p = 0.038$). The results are plotted in Figure 2.

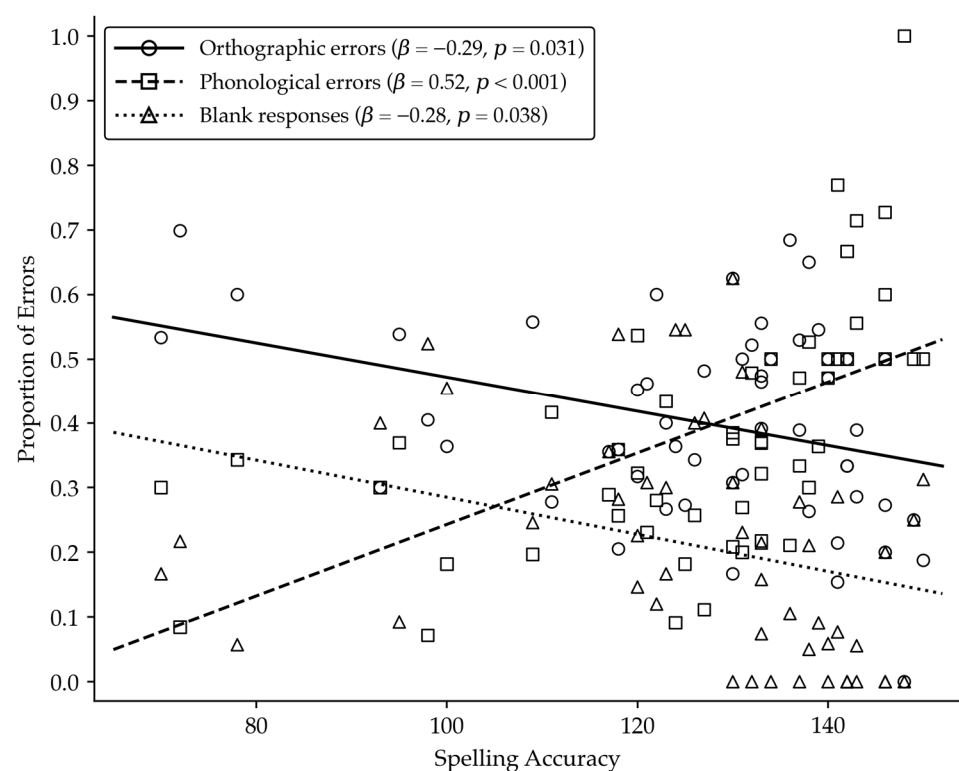


Figure 2. Proportion of orthographic errors, phonological errors, and blank responses as a function of spelling accuracy. Lines represent best-fit linear regressions.

Figure 2 shows a relatively low proportion of phonological errors for spellers with lower levels of accuracy, while orthographic errors and blank responses predominate. With an increase in spelling accuracy, there is a concomitant increase in phonological errors and decrease in orthographic and blank errors.

3.3. Item-Based Analyses of the Effect of Number of Logographemes on Spelling Accuracy

Individual-level logistic regression analyses were conducted. For each of the participants, a binary logistic regression was fitted with accuracy for each item in the spelling-to-dictation task (0 = correct, 1 = incorrect) as the dependent variable, and the log-transformed number of logographemes per character as the predictor. The resulting B coefficient for each participant indexes the strength of the effect of number of logographemes on the individual's spelling performance, with a larger B coefficient indicating greater sensitivity. We

next explored the association between the B coefficients and participants' level of spelling accuracy. The data are plotted in Figure 3.

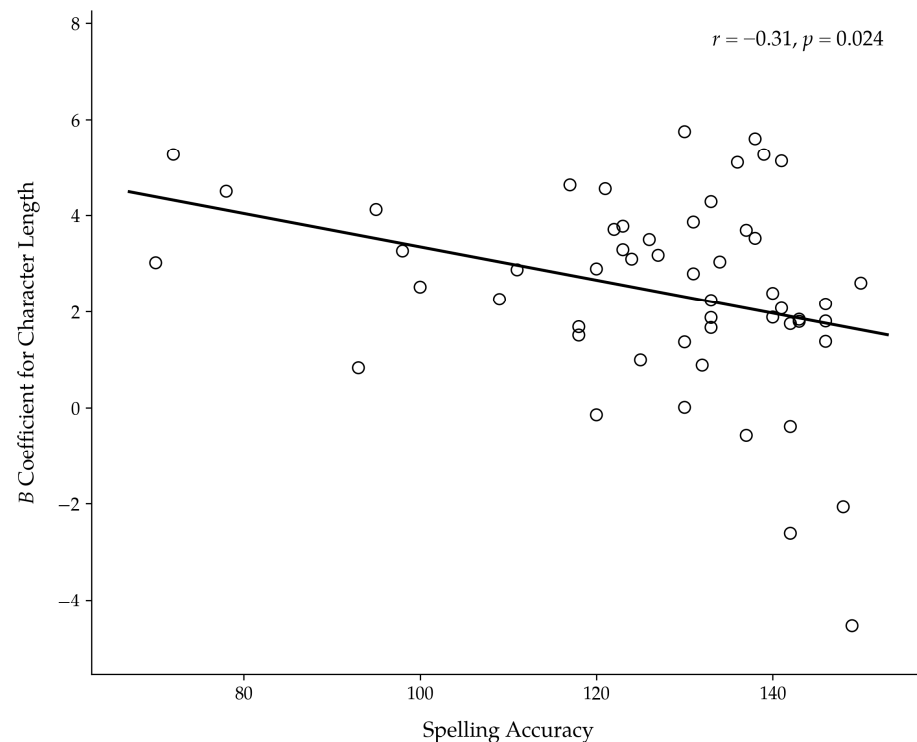


Figure 3. B coefficients from logistic regressions examining the relationship between spelling accuracy and the number of logographemes for each participant. The solid line represents the best-fit linear regression.

Analysis with partial correlation was conducted using the B coefficients from the logistic regression analyses and participants' spelling accuracy as variables. We controlled for chronological age since this variable had been found to be significantly associated with spelling accuracy in the first set of analyses (see Table 2). The results revealed a significant association, as shown in Figure 3, indicating a stronger effect of the number of logographemes on spelling accuracy for poorer spellers.

4. Discussion

This study investigated spelling-to-dictation in adult speakers of Mandarin and represents, to our knowledge, the first empirical investigation of individual differences in single-word spelling within this population. We employed tasks tapping a range of literacy-related variables that have been shown to be associated with spelling ability in past studies with adults. We aimed to examine the factors that might differentiate better from poorer spellers in our participants, and whether these would be the same or different to those identified in studies with alphabetic writing systems such as English.

The results revealed that scores for character report, but not phonological processing, were significantly associated with accuracy in the spelling-to-dictation task. Visual short-term memory was also found to be associated with spelling accuracy, although it was only a marginally significant predictor when entered together with character report in the second step of the hierarchical regression analysis, where participants' chronological age and single-character identification scores were entered in the first step. The finding aligns with previous results from studies with English-speaking adults (Holmes & Castles, 2001; Holmes & Ng, 1993; Holmes & Quinn, 2009), indicating the importance of orthographic

processing in differentiating better and poorer spellers. The findings from the character report task can be compared to those in the study of Masterson et al. (2007) where letter report was found to be worse in poorer than better adult spellers. Indeed, letter report has been found to be a significant predictor of spelling ability in studies with children in a range of languages (Chinese: T. Wang et al., 2025; Dutch: van den Boer et al., 2015; English: Niolaki et al., 2020; French: Valdois et al., 2025; Greek: Niolaki et al., 2024), although the theoretical interpretation of the association in these studies has been in terms of the visual attention span hypothesis. According to this, greater visual attention capacity results in more efficient orthographic learning since multiple letters in a printed word can be processed in parallel.

In addition, although the composite phonological ability measure showed substantial variability across participants (range: -4.10 to 2.83), it was not associated with spelling accuracy. This null finding is consistent with the studies of English-speaking adults by Holmes and Quinn (2009) and Masterson et al. (2007), who also found that phonological skills did not distinguish better from poorer spellers when reading ability was comparable. In the present study, all participants were competent readers, suggesting that phonological processing may be sufficiently developed in literate adults and that it no longer constrains spelling performance. As a cross-sectional study, our findings suggest an association between OWM and spelling ability but cannot definitively establish a causal relationship. Longitudinal or training studies would be required to better disentangle the causal directions of this relationship.

Within the framework adopted in the current study, the finding is consistent with the involvement of OWM in spelling, as in Mo (2023), whereby the speller must temporarily retain the identity and spatial arrangement of character components long enough to support graphic motor execution. Single-character identification taps stored orthographic representations in long-term memory. It was not found to be a significant predictor of spelling accuracy. The character report task engages the maintenance of orthographic identity in short-term memory units (in order to keep track of which have been reported during verbal report in the task). The finding that character report uniquely predicted spelling accuracy suggests that individuals with poorer spelling can have less efficient OWM.

We should also note that we found a significant negative association between chronological age and spelling accuracy in the correlation analysis. This may reflect the well-documented phenomenon of ‘character amnesia’ (提笔忘字) among Chinese adults, whereby handwriting accuracy deteriorates as reliance on phonology-based Pinyin digital-input methods increases over time (Langsford et al., 2025).

The analysis of spelling error proportions revealed a crossover pattern, depicted in Figure 2. We found that the proportion of orthographic errors increased with decreasing levels of spelling accuracy, while the proportion of phonological errors decreased, and blank responses increased. We suggest that the orthographic errors, which in the main part involved omission of components of characters, are due to OWM constraints, as suggested in previous studies with English-speaking spellers (e.g., Barisic et al., 2017; Caramazza et al., 1987; Kezilas et al., 2014; Kohnen et al., 2006; Roncoli & Masterson, 2016). Production of phonologically plausible substitutions may be due to failure to select the appropriate character from multiple characters activated by the spoken word (since Mandarin has very high homophone density), or else as a fallback strategy when a speller cannot recall a complete orthographic representation or, possibly, both of these reasons. Since production of a phonologically plausible substitute is resource-demanding, the poorer spellers may resort instead to blank responses which could account for the crossover pattern that we observed. These potential explanations could be investigated further with

alternative paradigms, such as masked priming and asking participants to judge the veracity of their spellings.

The item-based analyses revealed that spelling accuracy of the poorer spellers tended to be more negatively affected than the better spellers by the number of logographemes in a character. A pronounced effect of letter length in spelling was reported in English-speaking adults with dysgraphia that seemed to involve a deficit of OWM (e.g., Caramazza et al., 1987; Caramazza & Miceli, 1990; Kay & Hanley, 1994). This was also observed for Chinese adults with dysgraphia in terms of the number of logographemes per character and using delayed copying (Law & Leung, 2000; Han et al., 2007).

Taken together, the findings in our study that character report predicted spelling accuracy, plus the higher rate of orthographic errors and disproportionate effect of number of logographemes among the poorer spellers, mirror the findings interpreted as being due to OWM processes in previous research, which is extended here to individual differences in adult spellers of Mandarin. If the results are supported by further studies, this information could be used to develop targeted interventions. Sage and Ellis (2006) found that a programme involving letter and word search tasks resulted in improved spelling performance in an English-speaking adult with impaired OWM, and Roncoli and Masterson (2016) found improved spelling in an English-speaking child following whole-word delayed copying training. Extending this to Chinese, targeted interventions could be designed to enhance OWM efficiency/capacity. If OWM is a core deficit, then training this cognitive function should lead to measurable gains in spelling accuracy, particularly for longer words. Beyond adopting these paradigms (Makaure & Wilsenach, 2026), interventions for adult Mandarin spellers would benefit from being tailored to the specific orthographic characteristics of Chinese, for instance, targeting the maintenance of logographeme identity and position within complex character structures, rather than being adapted directly from intervention protocols developed for alphabetic writing systems.

Limitations and Implications

In terms of limitations, we only used a character report task as the OWM measure. We may have found different results with other types of tasks and stimuli. Future research may consider incorporating diverse tasks. Secondly, the participants in the current study could be considered to be homogeneous in terms of spelling ability. It would be informative to carry out future studies with participants with identified literacy problems.

A further limitation concerns the characterisation of the sample. We did not collect data on parental education, or handwriting-versus-digital-input habits. All participants had completed tertiary education (from junior-college diploma through to doctoral degree), which likely restricts the sample to a middle-to-upper educational stratum within the adult Mandarin-speaking population. Future studies should include a broader distribution of educational and socio-economic backgrounds, and should collect quantitative measures of adult handwriting versus digital-input practices, to test whether the current findings generalise to less formally educated adults and to those with more variable handwriting exposure.

5. Conclusions

This study investigated individual differences in single-word spelling among adult Mandarin speakers. The converging findings implicate OWM as a source of individual variation in adult Mandarin spelling, extending an account derived from research with English spellers and Chinese speakers with acquired dysgraphia. In a context where digital Pinyin-based input is displacing handwriting, OWM may be a productive target for adult

literacy support. Future work should use converging OWM measures, broader samples, and longitudinal or training designs to test the causal direction of this relationship.

Author Contributions: Conceptualization, T.W. and J.M.; methodology, T.W. and J.M.; software, T.W.; validation, T.W. and J.M.; formal analysis, T.W.; investigation, T.W.; resources, T.W.; data curation, T.W.; writing—original draft preparation, T.W.; writing—review and editing, T.W. and J.M.; visualisation, T.W.; supervision, J.M.; project administration, T.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The research project was approved by the Research Ethics Committee of University College London, IOE (Registration number: Z6364106/2021/12/56) on 27 May 2022. The study was carried out in accordance with the relevant guidelines and regulations.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: Request to access the datasets for this research should be emailed to the corresponding author.

Acknowledgments: The authors thank all the people who took part in the study. The authors also thank Liory Fern-Pollak for her guidance in experiment design. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

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

Abbreviations

The following abbreviations are used in this manuscript:

OWM	Orthographic working memory
Visual STM	Visual short-term memory
Verbal STM	Verbal short-term memory

Appendix A

Table A1. Hierarchical regression analysis with spelling accuracy as the dependent variable after controlling for educational level.

Step	Predictor	<i>B</i>	β	<i>t</i>	<i>p</i>	<i>R</i> ²	ΔR^2	<i>f</i> ² Δ	Tol	VIF
1	Chron. Age	−0.77	−0.24	−1.46	0.151	0.08	0.08	0.09	0.71	1.42
	Char. Ident.	0.05	0.06	0.38	0.707				0.75	1.34
	Edu_Level	−0.40	−0.01	−0.05	0.958				0.92	1.08
2	Chron. Age	−0.59	−0.18	−1.21	0.234	0.30	0.23 **	0.33	0.66	1.52
	Char. Ident.	−0.03	−0.03	−0.20	0.843				0.68	1.48
	Edu_Level	0.73	0.01	0.11	0.914				0.92	1.08
	VisualSTM	2.93	0.25	1.99	0.052				0.92	1.09
	Character report	0.55	0.40	3.08 **	0.003				0.90	1.11

Note. Chron. Age: chronological age; Char. Ident.: character identification; Edu_Level: educational level (1 = junior-college diploma, 0 = university degree); VisualSTM: visual short-term memory; *R*² = cumulative variance explained. ΔR^2 = change in *R*² at each step. *f*² Δ = Cohen's *f*² for the change in *R*² (small = 0.02, medium = 0.15, large = 0.35; Cohen, 2013). Adjusted *R*² for Steps 1 and 2 were 0.02 and 0.23, respectively. ** *p* < 0.01.

Note

- ¹ Zipf frequency is a standardised frequency measure that was introduced by van Heuven et al. (2014). It is calculated as log10 (frequency per million words) + 3 and has the advantage over raw frequency values that its interpretation does not depend on the size of an individual corpus. The Zipf scale is a logarithmic scale and ranges from 1 to 7 (with 7 being the highest frequency).

References

- Antzaka, A., Martin, C., Caffarra, S., Schleger, F., Carreiras, M., & Lallier, M. (2018). The effect of orthographic depth on letter string processing: The case of Visual Attention Span. *Reading and Writing*, 31(3), 565–581. [CrossRef]
- Barisic, K., Kohnen, S., & Nickels, L. (2017). Developmental graphemic buffer dysgraphia in English: A single case study. In *Developmental dysgraphia* (pp. 30–54). Routledge. [CrossRef] [PubMed]
- Barry, C. (1994). Spelling routes (or roots or rutes). In G. D. A. Brown, & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27–49). John Wiley.
- Behrendt, S. (2014). *lm.beta: Add standardized regression coefficients to linear-model-objects* (R package version 1.7-3) [Computer software]. Available online: <https://CRAN.R-project.org/package=lm.beta> (accessed on 26 April 2026).
- Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, 104(2), 198–230. [CrossRef] [PubMed]
- Bruck, M., & Waters, G. (1988). An analysis of the spelling errors of children who differ in their reading and spelling skills. *Applied Psycholinguistics*, 9(1), 77–92. [CrossRef]
- Brunsdon, R., Coltheart, M., & Nickels, L. (2006). Severe developmental letter-processing impairment: A treatment case study. *Cognitive Neuropsychology*, 23(6), 795–821. [CrossRef] [PubMed]
- Burden, V. (1992). Why are some “normal” readers such poor spellers? In C. M. Sterling, & C. Robson (Eds.), *Psychology, spelling and education* (pp. 200–214). Multilingual Matters.
- Burt, J. S., & Shrubsole, C. S. (2000). Processing of phonological representations and adult spelling proficiency. *Australian Journal of Psychology*, 52(2), 100–109. [CrossRef]
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, 5(6), e10729. [CrossRef] [PubMed]
- Campbell, R., & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: A developmental case with associated deficits of phonemic processing and awareness. *The Quarterly Journal of Experimental Psychology*, 37(3), 435–475. [CrossRef] [PubMed]
- Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, 37(3), 243–297. [CrossRef] [PubMed]
- Caramazza, A., Miceli, G., Villa, G., & Romani, C. (1987). The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition*, 26(1), 59–85. [CrossRef] [PubMed]
- Chan, K. S. C., & Yeung, P. S. (2020). Prediction of Chinese reading fluency by verbal and non-verbal visual attention span measures. *Frontiers in Psychology*, 10, 3049. [CrossRef] [PubMed]
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.
- Dunn, L., & Dunn, L. (1997). *Peabody picture vocabulary test* (3rd ed.). American Guidance Service.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. [CrossRef] [PubMed]
- Fox, J., & Weisberg, S. (2018). *An R companion to applied regression* (3rd ed.). Sage Publications.
- Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 495–516). Academic Press.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301–330). Lawrence Erlbaum Associates.
- Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, 35(5), 263–279. [CrossRef] [PubMed]
- Gong, Y. X. (1983). Revision of wechsler adult intelligence scale in China. *Acta Psychologica Sinica*, 15(3), 362–370.
- Gong, Z., & Guo, D. (1984). *Peabody picture vocabulary test—Chinese edition*. Xinhua Hospital affiliated to Shanghai Second Medical University.
- Goulondris, N. K., & Snowling, M. (1991). Visual memory deficits: A plausible cause of developmental dyslexia? Evidence from a single case study. *Cognitive Neuropsychology*, 8(2), 127–154. [CrossRef]
- Graham, S. (2006). Writing. In P. Alexander, & P. Winne (Eds.), *Handbook of educational psychology* (pp. 457–478). Erlbaum.
- Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research*, 26(3), 499–510. [CrossRef] [PubMed]
- Hamilton, L. G. (2007). *Literacy from pictures: The influence of character transparency and phonological awareness on learning to read Chinese* [Unpublished doctoral dissertation]. University of Oxford.
- Han, Z., Zhang, Y., Shu, H., & Bi, Y. (2007). The orthographic buffer in writing Chinese characters: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, 24(4), 431–450. [CrossRef] [PubMed]
- Hanley, J. R., & Kay, J. (1998). Does the graphemic buffer play a role in reading? *Cognitive Neuropsychology*, 15(3), 313–318. [CrossRef] [PubMed]
- Holmes, V. M., & Castles, A. E. (2001). Unexpectedly poor spelling in university students. *Scientific Studies of Reading*, 5(4), 319–350. [CrossRef] [PubMed]

- Holmes, V. M., Malone, A. M., & Redenbach, H. (2008). Orthographic processing and visual sequential memory in unexpectedly poor spellers. *Journal of Research in Reading*, 31(1), 136–156. [CrossRef]
- Holmes, V. M., & Ng, E. (1993). Word-specific knowledge, word-recognition strategies, and spelling ability. *Journal of Memory and Language*, 32(2), 230–257. [CrossRef]
- Holmes, V. M., & Quinn, L. (2009). Unexpectedly poor spelling and phonological-processing skill. *Scientific Studies of Reading*, 13(4), 295–317. [CrossRef]
- Kay, J., & Hanley, J. R. (1994). Peripheral disorders of spelling: The role of the graphemic buffer. In G. D. A. Brown, & N. C. Ellis (Eds.), *Handbook of spelling, theory, process and intervention* (pp. 295–315). Wiley.
- Kezilas, Y., Kohnen, S., McKague, M., & Castles, A. (2014). The locus of impairment in English developmental letter position dyslexia. *Frontiers in Human Neuroscience*, 8, 356. [CrossRef] [PubMed]
- Kohnen, S., Nickels, L., Brunsdon, R., & Coltheart, M. (2008). Patterns of generalisation after treating sub-lexical spelling deficits in a child with mixed dysgraphia. *Journal of Research in Reading*, 31(1), 157–177. [CrossRef]
- Kohnen, S., Nickels, L., Coltheart, M., & Brunsdon, R. (2006). Predicting generalisation in the training of irregular word spelling: Treating lexical spelling deficits in a child with mixed dysgraphia. *Australian Journal of Psychology*, 58(S1), 81.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. [CrossRef]
- Langsford, S., Xu, Z., & Cai, Z. G. (2025). Constructing a 30-item test for character amnesia in Chinese. *Reading and Writing*, 38(1), 121–141. [CrossRef]
- Law, S. P., & Leung, M. T. (2000). Structural representations of characters in Chinese writing: Evidence from a case of acquired dysgraphia. *Psychologia*, 43(1), 67–83.
- Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. *Journal of Research in Reading*, 35(3), 287–307. [CrossRef]
- Li, X., & Ma, G. (2012). Word boundaries affect visual attention in Chinese reading. *PLoS ONE*, 7(11), e48905. [CrossRef] [PubMed]
- Lin, D., McBride-Chang, C., Shu, H., Zhang, Y., Li, H., Zhang, J., Aram, D., & Levin, I. (2010). Small wins big: Analytic Pinyin skills promote Chinese word reading. *Psychological Science*, 21(8), 1117–1122. [CrossRef] [PubMed]
- Makaure, P., & Wilsenach, C. (2026). A linguistic analysis of spelling errors in Northern Sotho-speaking Grade 3 learners: A cross-sectional approach. *Perspectives in Education*, 44(2), 103–120. [CrossRef]
- Masterson, J., Colombo, L., Spencer, K., Ftika, K., & Syntili, A. (2008). Factors affecting spelling of English- and Greek-speaking Grade 1 and 2 children. In *Presented at the 1st writing systems conference*. Institute of Education.
- Masterson, J., Laxon, V., Lovejoy, S., & Morris, V. (2007). Phonological skill, lexical decision and letter report performance in good and poor adult spellers. *Journal of Research in Reading*, 30(4), 429–442. [CrossRef]
- McBride-Chang, C., Chung, K. K., & Tong, X. (2011). Copying skills in relation to word reading and writing in Chinese children with and without dyslexia. *Journal of Experimental Child Psychology*, 110(3), 422–433. [CrossRef] [PubMed]
- McCloskey, M., & Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive neuropsychology*, 34(3–4), 65–82. [CrossRef] [PubMed]
- Miceli, G., Silveri, M. C., & Caramazza, A. (1985). Cognitive analysis of a case of pure dysgraphia. *Brain and language*, 25(2), 187–212. [CrossRef] [PubMed]
- Mo, J. (2023). Does orthographic working memory matter in Chinese spelling? In *Routledge international handbook of visual-motor skills, handwriting, and spelling* (pp. 123–135). Routledge. [CrossRef]
- Mo, J., McBride, C., & Yip, L. (2018). Identifying the unique role of orthographic working memory in a componential model of Hong Kong kindergarteners' Chinese written spelling. *Reading and Writing*, 31(5), 1083–1108. [CrossRef]
- Niolaki, G. Z., Papadimitriou, V., Terzopoulos, A. R., & Masterson, J. (2024). Greek-spelling predictors; an investigation of literacy-and cognitive-related factors. *Journal of Cognitive Psychology*, 36(3), 378–393. [CrossRef]
- Niolaki, G. Z., Vousden, J., Terzopoulos, A. R., Taylor, L. M., Sephton, S., & Masterson, J. (2020). Predictors of single word spelling in English speaking children: A cross sectional study. *Journal of Research in Reading*, 43(4), 577–596. [CrossRef]
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. [CrossRef] [PubMed]
- Posit Team. (2026). *RStudio: Integrated development environment for R* (Version 2026.01.0+392) [Computer software]. Posit Software, PBC. Available online: <https://posit.co/> (accessed on 26 April 2026).
- Python Software Foundation. (2024). *Python language reference* (Version 3.12.7) [Computer software]. Available online: <https://www.python.org/> (accessed on 23 May 2026).
- Rapp, B., & Kane, A. (2002). Remediation of deficits affecting different components of the spelling process. *Aphasiology*, 16(4–6), 439–454. [CrossRef]
- Rastle, K. (2023). Understanding reading, understanding writing. *The Cognitive Psychology Bulletin*, (8). Available online: https://www.researchgate.net/profile/Kathleen-Rastle/publication/369171296_Understanding_reading_understanding_writing/links/642150b192cfd54f84323da8/Understanding-reading-understanding-writing.pdf (accessed on 31 March 2026).

- Raven, J. C., Court, J. H., & Raven, J. (1996). *Standard progressive matrices*. Oxford Psychologists Press.
- Raybaut, P. (2009). *Spyder documentation* (Version 5.5.1) [Computer software]. Available online: <https://www.spyder-ide.org/> (accessed on 23 May 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 26 April 2026).
- Roncoli, S., & Masterson, J. (2016). 'Unexpected' spelling difficulty in a 10-year-old child with good reading skills: An intervention case study. *Writing Systems Research*, 8(2), 143–166. [CrossRef]
- Sage, K., & Ellis, A. (2006). Using orthographic neighbours to treat a case of graphemic buffer disorder. *Aphasiology*, 20(9), 851–870. [CrossRef]
- Share, D. L. (2025). Anglocentrism in reading research and practice: The dialectical interface of language, orthography and literacy. *Language Teaching*, 58(1), 1–29. [CrossRef]
- Shen, H. H., & Bear, D. R. (2000). Development of orthographic skills in Chinese children. *Reading and Writing*, 13, 197–236. [CrossRef]
- Tainturier, M. J., & Rapp, B. (2003). Is a single graphemic buffer used in reading and spelling? *Aphasiology*, 17(6–7), 537–562. [CrossRef]
- Tong, X., McBride-Chang, C., Shu, H., & Wong, A. M. (2009). Morphological awareness, orthographic knowledge, and spelling errors: Keys to understanding early Chinese literacy acquisition. *Scientific Studies of Reading*, 13(5), 426–452. [CrossRef]
- Valdois, S., Zaher, A., Meyer, S., Diard, J., Mandin, S., & Bosse, M. L. (2025). Effectiveness of visual attention span training on learning to read and spell: A digital-game-based intervention in classrooms. *Reading Research Quarterly*, 60(1), e576. [CrossRef]
- van den Boer, M., van Bergen, E., & de Jong, P. F. (2015). The specific relation of visual attention span with reading and spelling in Dutch. *Learning and Individual Differences*, 39, 141–149. [CrossRef]
- van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, 67(6), 1176–1190. [CrossRef] [PubMed]
- Wang, T., Fern-Pollak, L., & Masterson, J. (2025). Investigation of the influence of visual attention span and a range of literacy-related variables on writing (spelling) to dictation in Mandarin-speaking children aged 10 to 12 years. *Acta Psychologica*, 261, 105908. [CrossRef] [PubMed]
- Wang, T., & Tao, B. (1996). *The scale and assessment of vocabulary*. Chinese Educational Research Press. (In Chinese)
- 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. [CrossRef]
- Wang, Y., Yin, L., & McBride, C. (2015). Unique predictors of early reading and writing: A one-year longitudinal study of Chinese kindergarteners. *Early Childhood Research Quarterly*, 32, 51–59. [CrossRef]
- Wong, H. L. S. (2017). *An investigation into processes involved in spelling in Mandarin speaking children aged 7–9* [Unpublished master's thesis]. Institute of Education, University College London.
- Ye, Y., Tong, C. K. Y., McBride, C., Yin, L., & Tse, C. Y. (2021). Concurrent and longitudinal predictors of beginning writing in Chinese: The unique role of copying skills. *Reading and Writing*, 34, 2417–2437. [CrossRef]
- Yeung, P. S., Ho, C. S. H., Wong, Y. K., Chan, D. W. O., Chung, K. K. H., & Lo, L. Y. (2013). Longitudinal predictors of Chinese word reading and spelling among elementary grade students. *Applied Psycholinguistics*, 34(6), 1245–1277. [CrossRef]

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.