Author's final manuscript

To cite:
Sun-A Kim, Kiel Christianson, & Jerome Packard (2015). Working memory in L2 character processing: The case of learning to read Chinese. In Edward Zhisheng Wen, Mailce Mota, & Arthur McNeill (Eds.), *Working Memory in Second Language Acquisition and Processing* (pp. 85-104). Bristol: Multilingual Matters.

Working Memory in L2 Character Processing: The Case of Learning to Read Chinese

Sun-A Kim, Kiel Christianson, and Jerome Packard

Author Note

Sun-A Kim, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University; Kiel Christianson, Department of Educational Psychology and Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign; Jerome Packard, Department of East Asian Languages and Cultures and Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign.

The manuscript was supported by a research grant from the Hong Kong Polytechnic University (G-U950) to the first author. The experiment reported here is a portion of the first author's doctoral dissertation which was funded in part by *Jiede* Empirical Research Grant from the Chinese Language Teachers Association.

Correspondence concerning this article should be addressed to Sun-A Kim, AG514, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: ctskim@polyu.edu.hk

Abstract

This study examines the influence of individual working memory (WM) on learning to read L2 Chinese in a character learning experiment. Two research questions are addressed. First, how do visual features of characters affect learning to read and WM involvement? Second, which domain of WM plays a more important role in learning to read Chinese? Seventy American college students enrolled in beginning/intermediate Chinese classes participated in a character learning experiment, a visuospatial and verbal WM task, and a character knowledge measure. In the character learning experiment, participants learned to read 18 unfamiliar simple Chinese characters divided into three levels of visual distinctiveness: distinctive, normal, and similar. In the distinctive set, one stroke of each character was artificially enhanced to make the characters visually distinctive. The normal set consisted of normal characters, and the similar set consisted of character pairs whose members were visually similar.

Results revealed that participants with higher levels of individual WM were better able to learn to read simple Chinese characters. Participants who had higher visuospatial WM spans were better able to learn the visually enhanced characters, while those with higher verbal WM capacities performed better in learning the regular Chinese characters belonging to the normal and the similar sets. The results suggest that learning ordinary Chinese characters is essentially a linguistic task for L2 character learners, which inevitably involves verbal WM. Visuospatial and verbal WM in Chinese beginning literacy acquisition and the domain-specificity/-generality of WM are discussed.

Keywords: working memory, Chinese characters, reading, Chinese as a second language

Working Memory in L2 Character Processing:

The Case of Learning to Read Chinese

Working memory (WM) is “a dedicated system to maintain and store information in the short term” (Baddely, 2003: 829). According to the best-known view of WM (Baddeley & Hitch, 1974), WM comprises three components: a higher-level control system with limited capacity of attention called the central executive, and two storage subsystems called the phonological loop and the visuospatial sketchpad. The central executive plays a role in coordinating and supervising information taken in through the phonological loop and the visuospatial sketchpad. The phonological loop temporarily maintains verbal information including sound and language, while the visuospatial sketchpad temporarily holds visual and spatial information. In other words, the WM system has two memory components in the verbal and the visuospatial domains and one attention-control component, which coordinates and regulates these two memory domains.

Various measures have been used to assess individual WM, grouped into short-term memory (STM) span tasks and WM span tasks based on the complexity of the tasks. STM span tasks are simpler than WM tasks, because STM tasks require participants to maintain target items in memory (the storage component). In contrast, WM tasks require participants to remember targets while performing another task (Miyake et al., 2001). In short, STM tasks consist of the storage requirement, whereas WM tasks have both storage and processing requirements. Digit span tasks and non-word span tasks are examples of verbal STM tasks, in that participants try to remember and recall as many digits or non-words as possible. On the other hand, reading span tasks are examples of verbal WM tasks, in that participants read aloud sentences and judge the plausibility of sentences (the processing component) while attempting to remember the word or letter at the end of each sentence (the storage component) for later recall.

Previous studies suggest that the components of WM reflected by STM and WM span tasks differ in the verbal and visuospatial domains. In the verbal domain, simple STM tasks reflect the storage component (i.e., the phonological loop) and complex WM tasks reflect both storage and central executive components (Engle et al., 1999). In the visuospatial domain, on the other hand, both STM and WM tasks are posited to involve the central executive component, but their degree of involvement in the central executive has not yet been clearly established (Miyake et al., 2001; Shah & Miyake, 1996). In general, maintaining visual images is more challenging than memorizing a sequence of verbal stimulus, because in the visuospatial domain a rehearsal mechanism --- such as the phonological loop in the verbal domain --- does not seem to exist (Miyake et al., 2001).

It has been documented that WM plays an important role in learning new words in another language (Cheung, 1996; Martin & Ellis, 2012; Service, 1992). However, most of these studies have been done in languages with alphabetic writing systems using verbal STM tasks and have focused on the role of verbal WM in learning new words. Verbal WM tasks have been used more for second language (L2) learning experiments requiring 'higher cognitive resources', such as grammar learning, reading comprehension and reasoning, and tend not to be used for L2 experiments requiring vocabulary learning and reading. Even when WM is occasionally included in L2 word learning, only the verbal domain of WM is examined (Cheung, 1996; Service, 1992).

To our knowledge, no empirical research in L2 acquisition has examined how the two domains of WM (i.e., verbal WM and visuospatial WM) influence learning L2 words. If one wants to investigate how visuospatial WM contributes to learning new L2 words, Chinese reading is an ideal candidate. Chinese script is visually more complex than alphabetic writing systems, because many strokes are condensed into the space occupied by a character (e.g., 繁 ‘complex’ and 懿 ‘beautiful’), and a difference even in the slant or length of a stroke will sometimes distinguish two characters (e.g., 千 ‘thousand’ vs. 干 ‘shield’; 天 ‘heaven’ vs. 夫 ‘husband’). Also, written Chinese is phonologically opaque, with pronunciation not explicitly revealed in the script. Therefore, the Chinese writing system is an ideal candidate for examining the role of verbal WM and visuospatial WM in L2 word learning due to its visual complexity and phonological opacity. If visuospatial WM plays a role in L2 word learning, then it is reasonable to expect that it would be manifested more clearly in learning to read a language like Chinese.

Although the results of research on the use of visual skills in reading Chinese are still not conclusive (e.g., McBride-Chang, 2004), it has been reported that first language Chinese readers tend to use visual strategies longer in the beginning stage of literacy and have better visual skills than native readers of alphabetic writing systems, ostensibly due to the Chinese readers' experience with visually complex characters. Four-year-old L1 Chinese children who are able to read 15 characters on average are found to still use visual strategies, although L1 English readers of an equivalent age who read about four words on average are reported not to use visual reading strategies but do use phonology-based strategies (Chen, 2004). This is further supported by another study in which five-year-old L1 Chinese children outperformed their Western counterparts in a visual task (McBride-Chang et al., 2011).

The current study aimed to examine the influence of individual visuospatial WM and verbal WM on adult L2 learners’ reading of simple Chinese characters with different visual features. Specifically, this study attempts to answer two questions. First, how do the visual features of characters affect learning to read? It has been reported that a visually enhanced feature in a word or a character helps less experienced readers read better (Chen, 2004; Ehri & Wilce, 1985). To examine how visual properties of characters influence reading characters by adult beginning L2 Chinese learners, visual features of simple characters were varied in three levels: distinctive, normal, and similar sets. In the distinctive set, one stroke of each character was artificially enhanced to make the characters visually distinctive. The normal set consisted of normal characters, and the similar set consisted of character pairs whose members were visually similar to each other. The second question is whether both visuospatial and verbal WM play a role in learning to read Chinese, and if so, whether they predict distinct learning behaviors. Individual differences in ability to process visual-orthographic information measured by visuospatial WM is thought to be important in learning to read Chinese because of the visual complexity of the Chinese script, and verbal WM may be necessary to hold phonological forms in memory while setting up a lexical entry.

Methods

Participants

Seventy students (28 females; 18-27 years old; mean age: 20 years old) enrolled in Chinese language classes at the University of Illinois participated and received monetary compensation for their participation. Forty-five students were native speakers of English, and 25 were not native speakers of English, but their L2 was in all cases English. There were 14 Korean, 2 Thai, 1 Spanish, 1 German, 1 Russian, and 6 Chinese dialect (Taishanese, Cantonese, and Fuzhouese) speakers. Forty-three students were enrolled in first semester (beginning) Chinese classes and 27 in third semester (intermediate) classes. At the time of participation in the experiment, the beginning learners had completed in-class instruction for

5 weeks (around 30 hours of in-class instruction) and learned approximately 60 characters in class, while most of the intermediate learners had completed 32 weeks (around 160 hours of instruction) of a two-semester beginning Chinese course which used a textbook containing 840 words and 745 characters.

Individual Differences Measures

WM tasks. Both a letter rotation task and a reading span task were used to assess individual participants' visuospatial and verbal WM capacity, respectively. Since both tasks measure WM, which involves not only information storage but also concurrent processing of additional information, the tasks consisted of two components: the presentation of to-be-remembered target stimuli, such as the spatial orientation of the top of a letter or a letter at the end of the sentence, and the completion of a secondary processing task, such as answering whether an image was normal or mirrored or judging the semantic plausibility of a sentence.

The letter rotation and reading span tasks had the same overall structure. Each task consisted of 42 items divided into 12 sets with a single set consisting of 2, 3, 4, or 5 items. Each set size appeared three times in each WM task. Each participant viewed the same items in the same predetermined single random order. WM tasks assess individual participants' differences in WM capacity, so the tasks needed to be implemented for each participant under the exact same conditions and order. In each WM task, participants were told to answer a question about an item while simultaneously retaining another piece of information about the item in memory, and were then asked to write down the remembered information on an answer sheet.

Letter rotation task. The letter rotation task used in the current study was modified from the task of the same name by Miyake et al. (2001) and was implemented using E-Prime software by the first author. The capital letters of F, J, L, P, or R were used as items in this task. Two manipulations were made of the position of the letters. In the first manipulation, the letters either remained as normal or were flipped along their vertical axis, making them either normal or mirror images. In the second manipulation, both normal and mirror images of each letter were rotated at multiples of 45 degrees, yielding the 8 possible positions of 45, 90, 135, 180, 225, 270, 315, and 360 degrees of rotation. These two manipulations required participants to identify both flipped and rotated orientations of letter images. Each letter had 16 possible normal or mirror-image orientations in total, yielding 80 possible images for all 5 letters. This task contained 42 random items among 80 possibilities, assigned to 1 of 12 sets. Each set contained sizes of two, three, four or five items and appeared three times in the entire task. The set started with the two-item presentation, with the same number of items presented three times. Then the number of letters per set was increased stepwise to 3, 4, and then to 5 letters per set, which yielded 12 sets in total.

Participants were instructed to say aloud “Regular” for normal letters or “Flipped” for mirrored letters immediately after seeing each letter on the computer screen and to remember the spatial orientation of the letter (i.e., where the top of the letter was pointing) simultaneously. After each set of items, participants were asked to report the orientation of each letter’s top surface in the correct serial order of presentation by writing numbers on an answer sheet containing a grid representing the eight possible positions of the letter’s top surface. Figure 1 displays the procedure of a set size of two. Participants started with a practice trial with three sets of two items and continued the practice trials until they became comfortable with the task. The task was designed to be fast-paced, moving to the next item with the exprimenter as soon as participants answered or reached the maximum time interval of three seconds between items. However, after each set when participants reported the orientation of the letters, they were given as much time as they wanted. This took 10-15 minutes, including instruction, practice, and the main task.

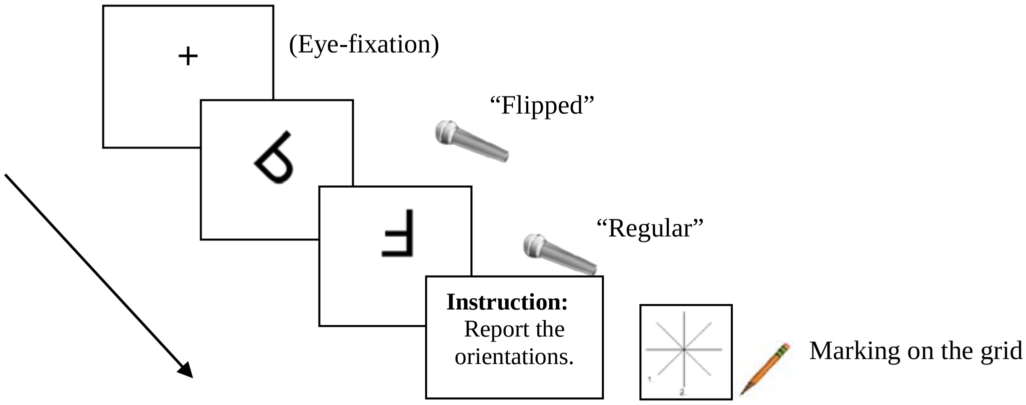


Figure 1. Example of a letter rotation task with a two-item set

Scoring the letter rotation task data followed the methods used by Gupta (2003), Miyake et al. (2001), and Shah & Miyake (1996). The memory portion of the test consisted of participants’ recalled report of the letter’s top surface orientation. Participants’ vocal judgments on normal or mirrored-images were recorded by the experimenter but were not included in the data analysis. In the WM tasks, the processing performance (in our case, image report) usually serves to ensure that participants paid attention to the secondary task and correlates positively with performance on the storage component, which indicates there is no trade-off between processing and recall accuracy (Conway et al., 2005; Kane et al., 2004; Shah & Miyake, 1996).

Individual written reports on the orientation of the top of the letter were collected, scored by the experimenter, and included in the data analysis as a measure of participants’ visuospatial WM capacity. Items that were correctly recalled in the correct serial position were counted as correct and given one point. The maximum score was 42 points.

Reading span task. The reading span task required participants to memorize the letter at the end of each sentence while judging the semantic plausibility of the sentence. Inefficient comprehension of sentences can influence holding the memorized letters in the correct order, so conducting the reading span task in the participant's native language is the ideal. Among the total of 70 participants, because the majority of participants were native speakers of English and Korean (64.3% and 20% respectively), the reading span task was conducted in those two languages. For the participants whose native language was not Korean, the English version was used.

The English version of the task used in this study was the task taken from Kane et al. (2004) developed by the study authors. The Korean version of the task was from Kim (2008) based on the original Kane et al. (2004) task. For the English version, each sentence consisted of 10-15 English words and was either semantically plausible or implausible. The total number of sentences in the task was 42. Out of 42 sentences, 19 sentences were semantically plausible, and the other 23 were implausible. At the end of each sentence, an alphabetic letter was presented. A sentence and a letter comprised a single item for the task, and a set consisted of two, three, four, or five items. Each set size occurred three times in the entire task, which yielded 12 trials total. The set size did not progressively increase but was randomly assigned to prevent the participants from strategically focusing on the to-be-remembered final letters.

The Korean version was created by matching the total number of items, the number of items in each set and the order of the 12 sets appearing in the task. The sentences and recall-syllable at the end of the items were presented in Korean, and were not direct translations of the English version but were matched with the English version for the order and number of ‘yes’ and ‘no’ answers to the semantic plausibility judgment subtask. The

Korean task was shown to be comparable with the English version by Kim (2008), who conducted both the English and Korean reading span tasks on 32 Korean-English late bilinguals, yielding a significant correlation of 0.71 ($p < 0.01$).

Participants required 10-15 minutes to complete the reading span task. They were required to read aloud the presented sentence, answer “Yes” or “No” prompted by a single question mark (?) regarding whether the sentence was semantically plausible, say aloud the letter after the question mark and remember the letters until they were told to write them down. Immediately after a set of two to five items ended, three question marks (???) appeared, which signaled the participant to write the letters down in serial order on the answer sheet. After that, participants moved to the next set by pressing the spacebar. Figure 2 displays the procedure of a set size of two.

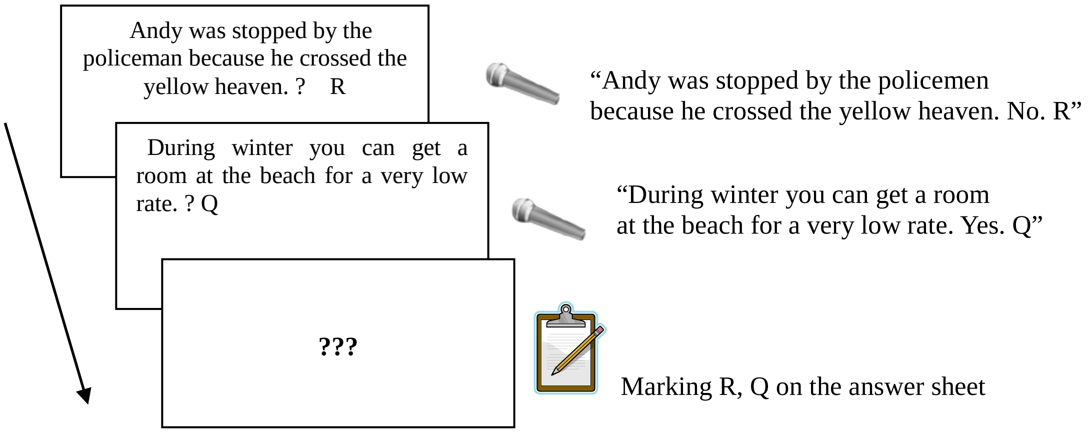


Figure 2. Example of the reading span task with a set size of two

The reading span task was scored using “All-or-nothing Unit Scoring” (Conway et al., 2005). Each item of the set received one point if all the items of the set were correct and in the correct serial order. If any item of the set was incorrect, then no credit was given. The maximum score for the reading span task was 42.

Chinese character test. This test was created by the lead author based on the Chinese textbooks used at the research site. The test consisted of 60 Chinese words, and required participants to write the pronunciation (*pinyin*) and meaning (English) of each word. Twenty words were selected from the first semester, another 20 from the second semester, and the last 20 drawn from the third and fourth semesters. Of the 60 words, 34 were one-syllable, and 26 were two-syllable, totaling 86 characters altogether. Partial points were given for the *pinyin* portion of the test, with 0.3 point given, respectively, for the correct initial, final, and tone. However, no partial point was given for the meaning portion. The maximum score on the character test was 146.

Questionnaires. At the beginning and the end of the experimental session, participants were asked to fill out two questionnaires, a Language Background Questionnaire and a Post-experiment Questionnaire. The Language Background Questionnaire obtained a detailed linguistic profile of the participants, such as their native language, foreign languages, and the length and the setting of Chinese language learning. The Post-experiment Questionnaire confirmed that it was the first time for the participants to learn the characters provided in the experimental session, with no participant reporting that there was a previously-known character in the study.

Character Learning Experiment

Materials. Eighteen characters were used as stimuli in the experiment. Only structurally simple characters without any phonetic component were selected to test for visual strategies. Also, all characters were ancient or extremely uncommon in order to ensure that even though all were real characters, none of the participants would have had any prior experience with them, as was confirmed by the Post-experiment Questionnaire. All character stimuli were selected from *Comprehensive Chinese Character Dictionary* (漢語大字典, 1993), and had identical forms in the traditional and simplified character types. The complete stimulus set for the experiment appears in Appendix A. Character pronunciations were recorded by a female native speaker of Chinese.

The 18 stimuli consisted of 3 character-type sets with 6 characters in each set. The three types were visually distinctive, normal, and similar. The mean number of character strokes in each type was 4.8, 4.8, and 5, respectively. The distinctive set contained characters with a visually enhanced feature, created by exaggerating the width or length of one stroke of a normal character (Chen, 2004). This is similar to methods used in previous studies on English orthography, in which the height and width of alphabetic components within an English word were varied (Ehri & Wilce, 1985). As an example of a distinctive character, a stroke in the center of the character  was exaggerated to form the character . The normal set consisted of six regular (albeit ancient or uncommon) Chinese characters without enhancement, e.g., . The similar type consisted of three pairs of characters that were visually similar to each other, differing only in one or two character components, e.g.,  vs. . The characters used in the distinctive and normal types were counter-balanced by participant by constructing two lists. For the similar set, counter-balancing was not applied. Each participant was randomly assigned to one of two lists.

Procedure. Participants were tested individually, seated in front of a laptop computer and wearing a headset in a quiet lab. Participants were instructed that they would first learn 18 novel Chinese characters one at a time and then be asked to name the learned characters. The experimenter sat behind the participant to monitor, record and score performance.

Each character appeared individually on the computer screen with its pronunciation simultaneously presented over the headset. The entire set of 18 characters was presented in random order by *E-Prime* software, starting with a fixation point displayed for 700 ms. After the fixation point, each character was presented for 5000 ms with its pronunciation played twice over the headphones. As soon as participants heard the pronunciation, they repeated aloud what they heard two times for each character. Immediately after learning the characters, all 18 were presented randomly one at a time on the computer screen, and participants were asked to name them. The learning and test phases were repeated three times.

Results

Descriptive Statistics and Correlations

Descriptive statistics for individual differences measures are presented in Table 1. Figure 3 shows the mean proportion correct in naming accuracy for character types by trial. There was a significant effect of trial from a mixed logit analysis (listed in Table 3), with significantly higher naming accuracy rate in Trial 2 and Trial 3 than in Trial 1, which indicates that participants were gradually able to learn to read the characters over three trials. Although the average proportion correct is not high even in Trial 3, there was a substantial amount of learning within less than 20 minutes.

Table 1

Descriptive statistics for the individual differences measures

Variable	M	SD	Min	Max
Letter Rotation	19.86	6.85	6	33
Reading Span	18.24	9.66	2	42
Character Test	60.92	9.66	24.50	129
Age	20.14	1.82	18	27

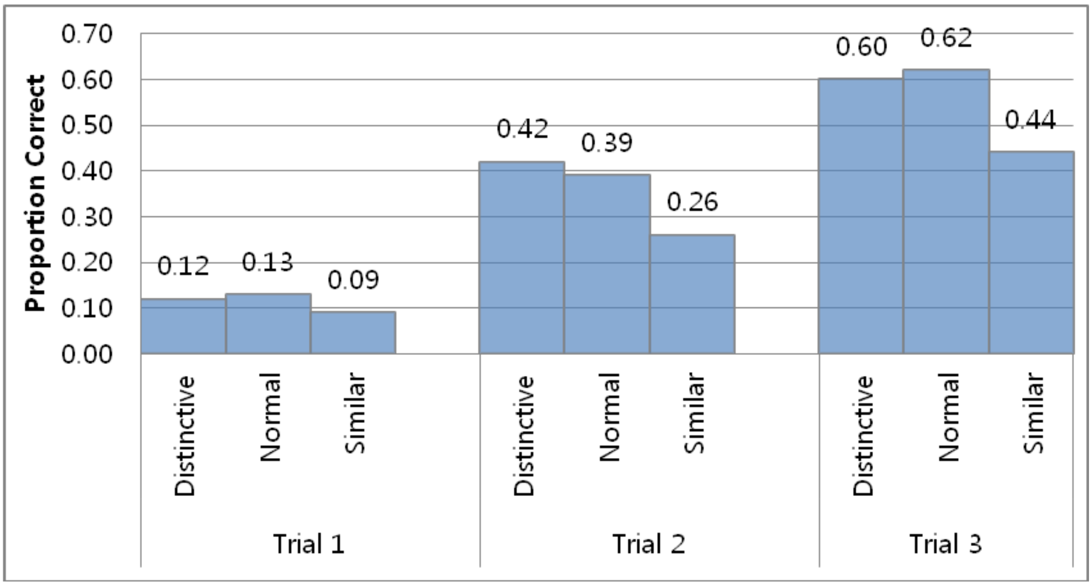


Figure 3. Proportion correct in naming accuracy for character types by trial

Correlations among individual differences measures, i.e. letter rotation, reading span, and character test are listed in Table 2. A significant correlation was found between letter rotation and reading span (r = 0.31, $p < 0.05$) but other measures did not correlate. This suggests that the two WM measures share some cognitive resources, but they are independent of the Chinese reading proficiency measure, i.e. the character test.

Table 2

Correlations among individual differences measures

Variable	Letter Rotation	Reading Span	Character Test
Letter Rotation	–	–	–
Reading Span	.31*	–	–
Character Test	.12	.12	–

*p < .05

Mixed Logit Analysis

A character was scored as *correct* if participants named an entire syllable (excluding tone) correctly, and as *incorrect* otherwise. If participants produced two or more syllables, the last syllable was taken as their final answer. Because character naming results from the experiment were dichotomous data, statistical analyses were carried out using a mixed logit model (Jaeger, 2008) in R (R Development Core Team, 2009).[1] Naming accuracy was the dependent variable, with discrete predictors (trial, character type, list, and participants' L1) and continuous predictors (character test scores, reading span task scores and letter rotation task scores) as fixed-effect factors, and participants and items as random-effect factors. To obtain the most parsimonious model with the best fit using the minimum number of predictors, all predictors and all possible interactions were first entered into the model, with the least contributing predictors eliminated one by one using likelihood ratio tests. In what follows, the coefficients and significance levels for those predictors remaining in the minimal, best-fitted model will be presented in the tables.

No significant effects were found for list, character knowledge or participants' L1. There was a significant effect of trial, with significantly higher naming accuracy rates in the second and the third trials than in the first trial, which indicates that participants gradually progressed in learning to read characters over the learning trials. Table 3 displays naming performance on all three character types combined over all three trials, and naming performance is seen as a coefficient (log-odds estimate) whose positive value indicates performance over a set baseline. In the predictor column, the naming accuracy of the first trial of learning (Trial 1) serves as the intercept and therefore the baseline for the other

[1] Mixed logit models are a generalization of logistic regression for binomially distributed outcomes which accounts for random subject and item effects in a single analysis. Analysis of Variance (ANOVA) is inappropriate to analyze dichotomous data (Dixon, 2008; Jaeger, 2008).

predictors, shown as the differences from the Trial 1 (intercept) baseline. The positive estimate and significance values for Trial 2 and Trial 3 mean that naming was more accurate in Trial 2 and Trial 3 than in Trial 1, indicating that participants improved in learning to read characters trial by trial, following our general expectation. The estimate for reading span is also positive and significant, meaning that participants whose reading span test scores were higher performed better in learning to read characters over all trials.

Table 3

Naming Accuracy on All Three Character Types in All Three Naming Trials

Predictor	Estimate	SE	z-value	Pr(>|z|)
(Intercept)	-3.155	0.319	-9.889	< 0.001 ***
Trial 2	1.798	0.119	15.050	< 0.001 ***
Trial 3	2.848	0.122	23.302	< 0.001 ***
Reading Span	0.032	0.013	2.497	0.013 *

*$p < .05$; ***$p < .001$

Table 4 shows the results of naming accuracy by trial. The row for Trial 1 shows no significant predictor, meaning that no variable significantly predicted participants' ability to learn characters in the first trial. The column for Trial 2 lists three predictors, and only similar character type and reading span were statistically significant. Since character type is a categorical and not a numerical variable, the intercept represents the naming accuracy for the default level of the character type, which here is the distinctive character type. The negative coefficient value for the 'similar' type indicates that naming performance for the similar character type was worse than the baseline (distinctive) type by an estimate (log-odds) of -0.866. The positive coefficient for reading span indicates that participants with higher reading span scores performed significantly better than those with lower reading span scores. In Trial

3, only reading span scores were significant, indicating that participants with higher reading span scores learned to read characters significantly better than those with lower reading span scores. Taken together, these results indicate that individual WM capacity contributed to participants' character learning more than the character types, because the reading span effect was found in Trial 2 and Trial 3, but the character type effect was only found for the similar type in Trial 2.

Table 4

Naming Accuracy in Each Naming Trial

Trial	Predictor	Estimate	SE	z-value	Pr(>|z|)
1	(Intercept)	-2.416	0.219	-11.01	< 0.001 ***
2	(Intercept)	-1.002	0.317	-3.166	0.002 **
	Normal	-0.114	0.151	-0.755	0.450
	Similar	-0.866	0.389	-2.228	0.026 *
	Reading Span	0.032	0.011	2.900	0.004 **
3	(Intercept)	-0.408	0.336	-1.216	0.224
	Reading Span	0.038	0.013	3.061	0.002 **

*$p < .05$; **$p < .01$; ***$p < .001$

Table 5 shows the results of separate analysis of each character type by each trial. In Trial 1, no variable contributed significantly to naming performance for any character type. From Trial 2, however, significant contributors to learning started to emerge. Letter rotation scores significantly predicted participants' accuracy in reading distinctive characters, while reading span scores significantly predicted accuracy in reading normal characters. This pattern for distinctive and normal types was also observed in Trial 3. For similar characters, no predictor was found in Trial 2, but reading span score was a significant predictor in Trial 3.

The similar type showed these effects later than the distinctive and the normal types, possibly because in Trial 2 the accuracy rate for the similar type was too low to show any significance compared to the normal type. In Trial 3, however, when the similar type reached a sufficiently high level of naming accuracy, verbal WM emerged as a significant factor. It is notable that there was no interaction between letter rotation and reading span scores. The absence of an interaction suggests that visuospatial WM and verbal WM resources are independent.

Table 5

Naming Accuracy on Each Character Type in Each Trial of the Training Phase

Trial	Character Type	Predictors	Estimate	SE	z-value	Pr(>|z|)
1	Distinctive	(Intercept)	-2.128	0.198	-10.770	< 0.001 ***
	Normal	(Intercept)	-2.281	0.270	-8.448	< 0.001 ***
	Similar	(Intercept)	-2.786	0.429	-6.490	< 0.001 ***
2	Distinctive	(Intercept)	-1.747	0.515	-3.391	< 0.001 ***
		Letter Rotation	0.066	0.023	2.822	0.005 **
	Normal	(Intercept)	-1.068	0.329	-3.250	0.001 **
		Reading Span	0.031	0.011	2.674	0.007 **
	Similar	(Intercept)	-1.434	0.260	-5.505	< 0.001 ***
		Reading Span	0.013	0.022	0.592	0.554
3	Distinctive	(Intercept)	-0.757	0.571	-1.325	0.185
		Letter Rotation	0.066	0.025	2.615	0.009 **
	Normal	(Intercept)	-0.039	0.385	-0.101	0.920
		Reading Span	0.037	0.015	2.488	0.013 *
	Similar	(Intercept)	-0.889	0.408	-2.178	0.029 *
		Reading Span	0.033	0.014	2.432	0.015 *

* $p < .05$; ** $p < .01$; *** $p < .001$

The results can be summarized as follows. First, participants significantly improved in learning characters from one trial to the next. Second, different levels of visual distinctiveness did not affect participants' learning to read characters, as seen in Table 4. Third, higher levels of individual WM resources increased participants' ability to read characters. Specifically, visuospatial WM as measured by the letter rotation task and verbal WM as measured by the reading span task contributed differently to learning different types of characters. Stronger visuospatial WM helped participants learn visually distinctive characters, whereas stronger verbal WM helped in learning regular characters. This interesting finding may be related to the nature of the different types of WM, to be further discussed below. Finally, in the first trial when participants were exposed to the characters for the first time, not only was their learning performance not good, but their individual differences in WM also had no effect. However, the effects of WM resources started to become noticeable as early as the second trial resulting in better learning, which then continued to the third trial.

Discussion

The present study of WM in learning to read L2 Chinese characters had two research questions. The first question was how different visual features of characters might interact with WM and affect the learning of those characters. The second question was whether visuospatial WM or verbal WM would play a more important role in learning Chinese characters. For the first question, the results indicate that higher levels of individual WM resources increased participants' ability to learn simple Chinese characters, and that visuospatial properties represented by an exaggerated stroke in a character taxed participants' visuospatial WM, while regular characters taxed their verbal WM. In particular, visuospatial WM as measured by the letter rotation task and verbal WM as measured by the reading span task contributed to learning different types of characters. Participants who had higher visuospatial WM capacity were better able to learn visually enhanced characters, whereas participants with higher verbal WM capacities performed better in learning regular Chinese characters belonging to the normal and the similar sets. It is intriguing that visuospatial WM facilitated learning visually enhanced characters, whereas verbal WM was helpful in learning regular Chinese characters. For the second question, the results suggest that verbal WM plays a more important role than visuospatial WM even in learning to read ordinary Chinese characters. Although visuospatial WM effects were found in reading visually distinctive characters, the visually distinctive characters were artificially created, and not ordinary characters.

The fact that visuospatial WM was implicated in learning the visually distinct characters implies that participants used the bold strokes as cues for recall. In other words, participants linked the position of the bold stroke to the rest of the character, and the visually exaggerated stroke was not recognized as being a linguistic cue even by novices. Using the position of the visually enhanced stroke to recall a character is obviously a visuospatial task, which is what the letter rotation task measures and predicts.

Altogether, this result suggests that learning ordinary Chinese characters is essentially a linguistic task for L2 character learners, as shown in the reading span effect (i.e., verbal WM) on performance for the normal and similar types. On the other hand, the manipulation of providing a distinctive, exaggerated bold stroke in a character may be considered non-linguistic, so the visuospatial WM factor only correlated with this distinctive type with its visual manipulation, and not with any of the other types.

Visuospatial and Verbal Working Memory in Learning to Read Chinese

Although the conditions of the experiment were devised with three degrees of visual distinctiveness of the characters, the results may be seen as applying to the two domains of WM. Visuospatial WM was linked to the characters with a visually distinctive stroke (distinctive type), whereas verbal WM was related to normal, conventional characters (normal and the similar types) regardless of whether there were visually similar characters in the pool of the characters to be learned. It is important to note that a visually salient stroke was artificially made for the experiment and is not typical in the modern Chinese script. The results suggest that L2 learners relied more heavily on visuospatial WM while processing characters with atypical, exaggerated strokes, and on verbal WM when processing regular and conventional Chinese characters.

The results raise two further questions. One involves what beginning L2 Chinese readers are aware of in learning to read Chinese, and the other involves visual processing in learning to read Chinese. First, it seems that even novice adult L2 Chinese readers are aware that an exaggerated stroke in a character makes no distinction in the Chinese writing system, and that they consider characters with a salient stroke as non-distinctive. This is congruent with previous studies showing that adult L2 learners learn very quickly which visual features and configurations are allowed in written Chinese (e.g. Liu et al., 2007; Wang et al., 2003), suggesting that sensitivity to conventional and legal forms of Chinese is developed quite early in L2 Chinese reading. Compared to the awareness of which form or position is allowed in the Chinese writing system, knowledge that stroke thickness does not make a character different is easier and more obvious. Thus, it is likely that the participants in this study who had learned approximately 60-750 characters perceived that a visually salient stroke in the distinctive type characters is extraordinary and unusual when they were learning 18 new characters in this experiment.

Why was the visuospatial WM measure not related to learning to read conventional characters in this study, but related to characters with a visually distinctive stroke? We suggest two possibilities. One possibility is that the naming task used in the character learning experiment could have demanded more verbal WM than visuospatial WM. In the experiment, participants learned to read characters by hearing and repeating the pronunciations, and learning was measured by the correct naming of the characters. Since the learning procedure is more closely related to the verbal domain of WM --- especially the function of the phonological loop and rehearsal --- than the visuospatial sketchpad, verbal WM may have been more activated and involved. A silent reading procedure in the character learning experiment could have made visuospatial WM more relevant, but we cannot be certain how much more visuospatial WM would be involved in a silent reading procedure, because previous studies suggest that phonological recoding occurs even in silent reading (Rayner & Pollasek, 1989), and phonological involvement is thought to be inevitable in reading regardless of how transparent or opaque a writing system is (Perfetti, 2003).

Another possible reason why visuospatial WM effects were not found in reading regular characters in this study is that the letter rotation task --- the visuospatial WM measure used in this study --- may not be a sensitive measure for reading ordinary Chinese characters. A recent empirical study based on L1 Chinese kindergarteners and Grade 1-2 elementary students revealed that a traditional visual skill measure using geometric figures (geometric-figure processing task) only predicted Chinese reading among kindergarteners, but that a visual judgment measure using real and pseudo Chinese characters (character-configuration processing task) was a stronger predictor of character reading for both kindergartners and early elementary students (Luo et al., 2013). It is possible that a Chinese character-related visuospatial task may have been a more appropriate task, and that visuospatial WM effects were not found for regular characters in this study because the letter rotation task used four

Roman alphabet letters (F, J, L, P, and R) not related to the character-configuration of Chinese. In future examinations of visuospatial WM in reading Chinese, this issue needs to be considered.

Implications for Learning to Read Chinese Characters

The results of this study suggest that better verbal WM is significantly helpful for students learning simple characters that do not contain a pronunciation cue, regardless of character knowledge. However, the situation changes once they start learning the majority of Chinese characters called “phonetic compound characters” containing a pronunciation cue, which represent more than 80% of modern Chinese characters. When the same participants of this study learned another set of 18 characters in a separate study in which simple and phonetic compound characters were mixed and the simple characters served as the pronunciation part of the phonetic compounds, WM effects were not seen while character knowledge significantly helped the participants learn phonetic compounds whose pronunciation cues were consistent (Kim, 2010). Since frequent characters are usually simpler in form and are taught in the earlier stage of Chinese education than less frequent ones (She et al., 2003), higher verbal WM capacity would presumably have benefits in the very beginning stage of learning to read Chinese. However, once learners have learned a number of phonetic compound characters and developed an awareness of how phonetics play a role in characters containing them, character knowledge becomes a more significant contributor to reading and character learning. This implies that learning simple characters that do not contain any pronunciation cue depends on mnemonic techniques that vary among individual readers, whereas learning to read phonetic compound characters is based on finding and utilizing the systematic principles underlying the phonetic components of characters which are inherent in the Chinese writing system.

Hierarchical View of Working Memory

This study also has implications for the general WM system. Researchers disagree whether WM is a unitary system (i.e., domain-general) or consists of multiple separate sub-units (i.e., domain-specific), and what kind of subsystems constitute WM if it is not unitary (see Shah & Miyake [1996] for a review). The domain-general and domain-specific views recently seem to agree that the complex nature of WM tasks yields results consistent with both domain-specificity and domain-generality, which can be summarized in the hierarchical view of WM suggested by Engle et al. (1999; Miyake, 2001). According to Engle et al., WM consists of a storage component and a controlled attention component. The storage component is the same as the traditional concept of STM and is domain-specific, whereas controlled attention is the ability to maintain relevant task information in the midst of distraction or interference, inhibit irrelevant information for the task, and switch attention between these, which is related to higher-order cognition and is domain-general. For example, memorizing the last letter presented after the sentence in a reading span task or the direction of the top of the letter in a letter rotation task are tests of the storage component, while judging whether the read sentence is semantically valid in a reading span task or whether the letter on the screen is a normal or a mirrored image are examples of the controlled attention component. This hierarchical view incorporates both the domain-general and the domain-specific views by considering WM as ‘hierarchical structure with a general domain-free factor overarching several subordinate domain-specific factors’ (Engle et al., 1999: 125). In other words, both domain-specific and domain-general systems exist, and the domain-general system is hierarchically above the domain-specific systems.

The findings from the two WM tasks in the current study support the hierarchical view containing a domain-general as well as a domain-specific effect. A significant correlation ($r = 0.31$, $p < 0.05$) was found only between the letter rotation and the reading span tasks, and not with any other variables[2]. This moderate correlation implies that the visuospatial and the phonological domains of the WM share some cognitive resources. At the same time, the results from the mixed logit analysis found that letter rotation scores significantly predicted participants' performance in learning visually distinctive characters and reading span scores significantly predicted better learning of regular characters without any artificial visual-enhancement. Yet, there were no interaction effects between the two measures. This pattern suggests that the visuospatial and the verbal domains are separate from each other. Taken together, these results imply that there are domain-general and domain-specific components in the WM system, consistent with the hierarchical view of WM (Engle, et al., 1999).

[2] In other studies, correlation effects between the letter rotation and the reading span scores are mixed. Friedman and Miyake (2000) found a significant correlation between them in their first experiment, but not in their second experiment using a reduced version of the reading span task. Shah and Miyake (1996) did not find correlation effects when they used the same measures as the first experiment of Friedman and Miyake (2000).

References

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews. Neuroscience, 4*(10), 829-839.

Chen, X. (2004). *Developmental stages in learning to read Chinese characters*. (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology, 32(5)*, 867-873.

Comprehensive Chinese Character Dictionary [漢語大字典], (1993). Chengdu, China: Sichuan Dictionary Publisher / Hubei Dictionary Publisher.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769-786.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59,* 447-456.

Engle, R., Tuholski, S.W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20*, 163-179.

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General, 129*, 61-83.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, immediate serial recall in adults. *Quarterly Journal of Experimental Psychology, 56A*, 1213-1236.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434-446.

Liu, Y., Wang, M., & Perfetti, C. A. (2007). Threshold-style processing of Chinese characters for adult second-language learners. *Memory & Cognition, 35*(3), 471-480.

Luo, Y. C., Chen, X., Deacon, S. H., Zhang, J., & Yin, L. (2013). The role of visual processing in learning to read chinese characters. *Scientific Studies of Reading, 17*(1), 22-40.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189-217.

Kim, J. (2008). *Working memory effects on bilingual sentence processing.* (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Kim, S. (2010). *Developmental stages in reading Chinese as a second language*. (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological STM and working memory in l 2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, 34(3).

McBride-Chang, C. (2004). *Children's literacy development*. London: Arnold.

McBride-Chang, C., Zhou, Y., Cho, J., Aram, D., Levin, I., & Tolchinsky, L. (2011). Visual spatial skill: A consequence of learning to read? *Journal of Experimental Child Psychology,109*(2), 256-262.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General, 130*(2), 163-168.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*(4), 621-640.

Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading, 7*, 3-24.

R Development Core Team. (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hillsdale, NJ: Lawrence Erlbaum.

Service, E. (1992). Phonology, working memory and foreign-language learning. *Quarterly Journal of Experimental Psychology, A 45 (1)*, 21-50.

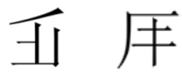
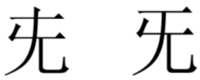
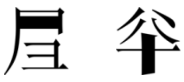
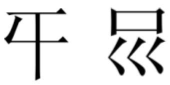
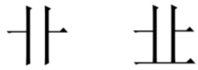
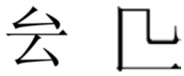
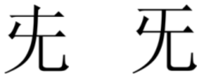
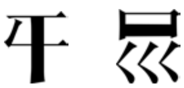
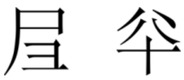
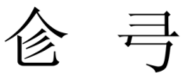
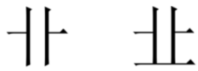
Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*(1), 4-27.

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development, 74*(1), 27-47.

Wang, M., Perfetti, C. A., & Liu, Y. (2003). Alphabetic readers quickly acquire orthographic structure in learning to read Chinese. *Scientific Studies of Reading, 7*(2), 183-208.

Appendix A

Materials for Experiment

List	Character Type: Distinctive	Character Type: Normal	Character Type: Similar
List 1			
	bai3 fang4	fa2 ju2	zan1 ji4
			
	liang2 qin4	kua4 pin3	gu3 mao3
			
	shi1 tan2	zai1 zhang3	kuang4 qiu1
List 2			
	fa2 ju2	bai3 fang4	zan1 ji4
			
	kua4 pin3	liang2 qin4	gu3 mao3
			
	zai1 zhang3	shi1 tan2	kuang4 qiu1